

Mathematik für Physiker

Gerhard Knieper
Ruhr-Universität Bochum
Fakultät für Mathematik
SS 2021

Version vom 11. Oktober 2024

Inhaltsverzeichnis

1 Grundlagen	5
1.1 Mengen und Abbildungen	5
1.2 Die Rechenregeln für die reellen Zahlen (Körperaxiome)	13
1.3 Die Anordnungsaxiome der reellen Zahlen	16
1.4 Die natürlichen Zahlen	20
1.5 Vollständigkeitsaxiom	25
1.6 Die komplexen Zahlen	29
2 Folgen und Reihen	33
2.1 Konvergenz von reellen Zahlenfolgen	33
2.2 Rechenregeln für konvergente Folgen	36
2.3 Konvergenzkriterien	37
2.4 Konvergenz von komplexen Zahlenfolgen	44
2.5 Konvergenz von Reihen	46
2.6 Umordnung von Reihen, absolute Konvergenz und Doppelreihen	55
2.7 Potenzreihen	68
3 Stetigkeit	73
3.1 Definition und grundlegende Eigenschaften	73
3.2 Stetige reelle Funktionen auf Intervallen und der Zwischenwertsatz	76
3.3 Kompakte Mengen und stetige Funktionen	81
3.4 Stetige Fortsetzbarkeit	85
3.5 Uneigentliche Grenzwerte	90
4 Differentialrechnung	93
4.1 Differenzierbarkeit von Funktionen	93
4.2 Kurven im \mathbb{R}^n	97
4.3 Rechenregeln für differenzierbare Funktionen	98
4.4 Extrema und Mittelwertsätze	101
4.5 Höhere Ableitungen und die Taylorsche Formel	105
4.6 Binomialreihen	113
5 Integralrechnung	119
5.1 Treppenfunktionen	119
5.2 Hauptsatz der Differential- und Integralrechnung	123
5.3 Integrationsregeln	125
5.4 Uneigentliche Integrale	126

6 Funktionenfolgen	131
6.1 Punktweise und gleichmäßige Konvergenz	131
6.2 Gleichmäßige Konvergenz und Vertauschung von Grenzprozessen	134
7 Vektorräume	139
7.1 Grundlegende Definitionen	139
7.2 Basis und Dimension	147
7.3 Lineare Komplemente und Dimensionsformel	153
8 Lineare Abbildungen	157
8.1 Grundlegendes über lineare Abbildungen	157
8.2 Erzeugendensysteme, Basissysteme und lineare Abbildungen	162
8.3 Vektorräume linearer Abbildungen und Dualräume	166
9 Lineare Differentialgleichungen I	169
9.1 Lineare Differentialgleichungen n -ter Ordnung	169
9.2 Homogene lineare Differentialgleichungen mit komplexen Koeffizienten	172
9.3 Homogene lineare Differentialgleichungen mit reellen Koeffizienten	175
9.4 Inhomogene lineare Differentialgleichungen mit reellen und komplexen Koeffizienten	178
10 Matrizen	183
10.1 Grundlegendes über Matrizen	183
10.2 Matrizen und lineare Abbildungen	186
10.3 Der Rang einer linearen Abbildung	189
10.4 Gaußscher Algorithmus	193
10.5 Lineare Gleichungssysteme	199
11 Determinanten	205
11.1 Die definierenden Eigenschaften der Determinante	205
11.2 Determinanten von Matrizen und Anwendungen	209
12 Eigenwerte und Klassifikation von Endomorphismen	217
12.1 Diagonalisierbarkeit und Eigenwerte	217
12.2 Polynome und der euklidische Algorithmus	220
12.3 Das charakteristische Polynom	224
12.4 Der Satz von Cayley-Hamilton	227
12.5 Invariante Unterräume	230
13 Euklidische und unitäre Vektorräume	237
13.1 Das skalare Produkt	237
13.2 Eigenschaften des Skalarproduktes	239
13.3 Orthonormalsysteme	242
13.4 Die adjungierte Abbildung	249
13.5 Normale Endomorphismen	254
13.6 Charakterisierung der selbstadjungierten, schiefadjungierten sowie der unitären und orthogonalen Endomorphismen	263
13.7 Symmetrische und Hermitesche Formen	269

14 Metrische Räume	277
14.1 Normierte und metrische Räume	277
14.2 Konvergenz und Stetigkeit in metrischen Räumen	279
14.3 Topologie in metrischen Räumen	286
15 Differentialrechnung auf Vektorräumen	295
15.1 Differenzierbarkeit	295
15.2 Rechenregeln für differenzierbare Abbildungen	304
15.3 Mittelwertsatz	306
15.4 Multilineare Abbildungen	313
15.5 Ableitungen höherer Ordnung	317
15.6 Taylorformel	327
15.7 Extremwerte	330
16 Der Banachsche Fixpunktsatz und seine Anwendungen	333
16.1 Banachscher Fixpunktsatz	333
16.2 Lokaler Umkehrsatz	334
16.3 Implizite Abbildungen und das Lösen von nicht linearen Gleichungen.	338
16.4 Differenzierbare Untermannigfaltigkeiten des \mathbb{R}^n	341
16.5 Tangentialraum und Normalraum	350
16.6 Lokale Extrema und Lagrangesche Multiplikatoren	353
16.7 Gewöhnliche Differentialgleichungen und der Satz von Picard-Lindelöf.	357
16.8 Lineare Differentialgleichungen II	365
16.9 Autonome Vektorfelder und dynamische Systeme.	368
16.10 Euler-Lagrange Gleichungen und Variationsrechnung.	372
17 Differentialformen 1. Grades (Pfaffsche Formen) und Kurvenintegrale	377
17.1 Kurven und ihre Länge	377
17.2 Differentialformen 1. Grades und Kurvenintegrale	379
17.3 Holomorphe Funktionen	390
17.4 Cauchysche Integralformel und Analytizität holomorpher Funktionen	397
17.5 Holomorphe Funktionen auf Kreisringen und ihre Laurententwicklung	401
17.6 Der Residuensatz und seine Anwendungen	405
18 Maß- und Integrationstheorie	413
18.1 Maße und messbare Mengen	413
18.2 Messbare Abbildungen	427
18.3 Translationsinvarianz des Lebesguemaßes	428
18.4 Messbare Funktionen	433
18.5 Das Lebesgue-Integral	441
18.6 Konvergenzsätze	450
18.7 Produkt-Maß und der Satz von Fubini	455
18.8 Transformationsformel	467
18.9 L^p -Räume	473

19 Differentialformen und Integralsätze	479
19.1 Differentialformen	479
19.2 Zurückholen (pull-back) von Differentialformen	487
19.3 Integration von Differentialformen	494
19.4 Integralsatz von Stokes	502
19.5 Die Volumenform und die klassischen Integralsätze	507

Kapitel 1

Grundlagen

1.1 Mengen und Abbildungen

Bevor wir mit den grundlegenden Begriffen der Analysis beginnen, müssen wir kurz auf die wesentlichen Begriffe der naiven Mengenlehre eingehen. Auch den Begriff der Abbildung zwischen Mengen, der in allen Teilen der Mathematik und auch der Physik eine fundamentale Rolle spielt, werden wir definieren.

Die für unsere Zwecke ausreichende "Definition" einer Menge geht auf Cantor (1895) zurück.

Definition 1.1.1. Eine *Menge* ist eine Zusammenfassung von wohlbestimmten und wohlunterschiedenen Objekten zu einem Ganzen. Die Objekte nennt man Elemente.

Wohlbestimmt bedeutet: von jedem Element steht fest, ob es zur Menge gehört oder nicht.

Wohlunterschieden bedeutet: jedes Element kommt nur einmal in der Menge vor.

(z.B. schreiben wir $\{1\}$ statt $\{1, 1\}$)

Wir benutzen folgende Symbole: $x \in M$ für x ist Element von M und $x \notin M$ für x ist kein Element von M .

Beispiele.

(a) $\{1, 2, 3\}$.

(b) $\mathbb{N} := \{1, 2, 3, \dots\}$ Menge der natürlichen Zahlen.

(c) $\mathbb{Z} := \{0, \pm 1, \pm 2, \dots\}$ Menge der ganzen Zahlen.

(d) $2\mathbb{Z} := \{0, \pm 2, \pm 4, \dots\}$ Menge der geraden Zahlen.

(a),(b),(c),(d) sind Beispiele für Aufzählungen. Meist werden jedoch Mengen mit sogenannten Aussageformen beschrieben.

Definition 1.1.2. Eine *Aussageform* A ordnet jedem Element x einer Menge G (G heißt *Grundbereich*) eine Aussage $A(x)$ zu. Dabei ist eine *Aussage* ein Satz, der entweder richtig oder falsch ist.

Beispiele. $G = \mathbb{N}$, so können wir zum Beispiel jedem $x \in \mathbb{N}$ die Aussage

$$A(x) := x \text{ ist eine Primzahl}$$

zuordnen. Die Aussage ist $A(x)$ genau dann richtig, falls $x \in \{2, 3, 5, 7, 11, 13 \dots\}$.
 Ein weiteres Beispiel einer Aussagenform mit Grundbereich $G = \mathbb{N}$ ist

$$A(x) := x > 5.$$

Diese ist richtig, falls $x \in \{6, 7, 8, 9, 10, 11, 12 \dots\}$

Ist nun A eine Aussagenform mit Grundbereich G , so verstehen wir unter

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

die Menge aller $x \in G$, für die $A(x)$ wahr ist.

Beispiele.

(a) $M = \{x \in \mathbb{Z} \mid x > 5\} = \{6, 7, 8, 9, 10, 11, 12 \dots\}$.

(b) $M = \{x \in \mathbb{Z} \mid \frac{1}{2}x \in \mathbb{Z}\} = 2\mathbb{Z}$.

(c) $M = \{x \in \mathbb{Z} \mid 3x = 2\} = \emptyset$.

Das Symbol \emptyset bezeichnet die *leere Menge*, d.h. die Menge, die kein Element enthält.

Wie viele Objekte in der Mathematik, so können auch Mengen verknüpft werden.

Definition 1.1.3. Es seien M, N Mengen. Dann heißen:

(a) $M \cup N = \{x \mid x \in M \text{ oder } x \in N\}$ die *Vereinigung* von M und N .

(b) $M \cap N = \{x \mid x \in M \text{ und } x \in N\}$ der *Durchschnitt* von M und N .

(c) $M \setminus N = \{x \mid x \in M \text{ und } x \notin N\}$ die *Differenz* von M und N .

Bemerkung. Die oben definierten Verknüpfungen zwischen Mengen, hängen eng mit entsprechenden Verknüpfungen zwischen Aussagen zusammen. Seien A und B Aussagen. Dann versteht man unter der Konjunktion von A und B die Aussage $A \wedge B$ (A und B), die genau dann wahr ist, falls A und B gleichzeitig wahr sind.

Die Disjunktion von A und B ist die Aussage $A \vee B$ (A oder B), die genau dann wahr ist, falls A oder B wahr ist (es handelt sich nicht um ein ausschließendes Oder, d.h. $A \vee B$ ist auch dann wahr, wenn beide Aussagen A, B wahr sind).

Die Negation von A ist die Aussage $\neg A$ (nicht A), die genau dann wahr ist, falls A falsch ist. Die Wahrheitswerte dieser Aussagen lassen sich sehr gut durch Wahrheitstabellen darstellen. Man erhält:

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

für $A \wedge B$ bzw. $A \vee B$ und

A	$\neg A$
w	f
f	w

für $\neg A$.

Damit gilt:

$$M \cup N = \{x \mid x \in M \vee x \in N\} \text{ und } M \cap N = \{x \mid x \in M \wedge x \in N\}$$

sowie

$$M \setminus N = \{x \mid x \in M \wedge x \notin N\}$$

Für Mengen führen wir die folgende wichtige Relation “ \subset ” ein.

Definition 1.1.4. Es seien M, N Mengen.

Dann heißt M *Teilmenge* von N ($M \subset N$), falls jedes Element aus M auch in N enthalten ist.

M und N heißen *gleich* ($M = N$), falls gilt: ($M \subset N$) und ($N \subset M$).

Bemerkung. Die Teilmengenrelation ist ebenfalls eng mit der Aussagenlogik verknüpft. Sind A, B Aussagen, so versteht man unter der Implikation $A \Rightarrow B$ (A impliziert B) die Aussage: wenn A gilt, so auch B .

Genauer:

A	B	$A \Rightarrow B$
w	w	w
w	f	f
f	w	w
f	f	w

Hier ist zu bemerken, dass $A \Rightarrow B$ auch richtig ist, falls A falsch ist. Aus einer falschen Prämisse läßt sich alles folgern. Unter einer Äquivalenz $A \Leftrightarrow B$ verstehen wir die Aussage $(A \Rightarrow B) \wedge (B \Rightarrow A)$. Sie ist genau dann wahr, falls A und B beide wahr oder beide falsch sind. Damit lässt sich die Teilmengenrelation wie folgt beschreiben: ($M \subset N$), falls

$$\text{Für alle } x \text{ gilt : } x \in M \Rightarrow x \in N$$

In der Aussagenlogik ist es üblich “für alle” durch das Zeichen \forall abzukürzen. Also können wir auch schreiben:

$$M \subset N \Leftrightarrow \forall x : (x \in M \Rightarrow x \in N).$$

Insbesondere ist \emptyset Teilmenge jeder Menge, denn

$$\forall x : (x \in \emptyset \Rightarrow x \in M)$$

ist wahr, da $x \in \emptyset$ falsch ist.

Das Symbol \forall nennt man auch Allquantor. Daneben gibt es auch noch den Existenzquantor \exists . Diese treten in folgender Form immer mit Aussageformen auf.

Definition 1.1.5. Sei $A(x)$ eine Aussageform mit Grundbereich G

(a) $\forall x \in G : A(x)$

Dies bedeutet: für alle $x \in G$ ist die Aussage $A(x)$ wahr.

(b) $\exists x \in G : A(x)$

Dies bedeutet: es existiert ein $x \in G$, so dass $A(x)$ wahr ist.

Bemerkung. Sowohl “ $(\forall x \in G: A(x))$ ” als auch “ $(\exists x \in G: A(x))$ ” sind Aussagen. Die Aussage “ $(\forall x \in G: A(x))$ ” ist genau dann falsch, falls ein $x_0 \in G$ existiert, so dass $A(x_0)$ falsch ist. In diesem Falle ist also “ $(\exists x \in G: \neg A(x))$ ” wahr.

Zum Beispiel ist die Aussage “ $\forall x \in \mathbb{N}: x$ ist eine gerade Zahl” offensichtlich falsch. Die Aussage “ $(\exists x \in \mathbb{N}: x$ ist keine gerade Zahl)” ist hingegen eine wahre Aussage.

Auch Mengen können wieder als Elemente einer Menge aufgefaßt werden.

Definition 1.1.6. Sei M eine Menge, so heißt

$$\mathcal{P}(M) = \{A \mid A \subset M\}$$

die *Potenzmenge* von M . $\mathcal{P}(M)$ ist die Menge aller Teilmengen von M .

Beispiel. Ist zum Beispiel $M = \{1, 2, 3\}$, so folgt:

$$\mathcal{P}(M) = \{\{1, 2, 3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1\}, \{2\}, \{3\}, \emptyset\}.$$

Satz 1.1.7. Es seien L, M, N Mengen. Dann gilt:

$$(1a) \quad L \cup M = M \cup L$$

Kommutativgesetze

$$(1b) \quad L \cap M = M \cap L$$

$$(2a) \quad L \cup (M \cap N) = (L \cup M) \cap (L \cup N)$$

Assoziativgesetze

$$(2b) \quad L \cap (M \cup N) = (L \cap M) \cup (L \cap N)$$

$$(3a) \quad L \cup (M \cap N) = (L \cup M) \cap (L \cup N)$$

Distributivgesetze

$$(3b) \quad L \cap (M \cup N) = (L \cap M) \cup (L \cap N)$$

Ist $M, N \subset L$, so gilt:

$$(4a) \quad M^c \cup N^c = (M \cap N)^c$$

de Morgansche Regeln

$$(4b) \quad M^c \cap N^c = (M \cup N)^c$$

Beweis. Wir beweisen (3a). Alle anderen Regeln beweist man mit analogen Argumenten. Nach Definition der Gleichheit von Mengen müssen dazu folgende Aussagen bewiesen werden.

$$\text{Beh.1: } L \cup (M \cap N) \subset (L \cup M) \cap (L \cup N)$$

$$\text{Beh.2: } (L \cup M) \cap (L \cup N) \subset L \cup (M \cap N)$$

Zu 1:

Ist $x \in L \cup (M \cap N)$, so folgt: $x \in L$ oder $x \in M \cap N$.

Ist $x \in L$, so ist x erst recht in $L \cup M$ als auch in $L \cup N$ enthalten und somit auch im Durchschnitt. Ist x nicht in L enthalten so ist $x \in N$ als auch in M enthalten und somit auch im Durchschnitt von $L \cup M$ und $L \cup N$.

Zu 2:

Ist $x \in (L \cup M) \cap (L \cup N)$, so ist x in $L \cup M$ und in $L \cup N$ enthalten. Also ist x in L oder aber in M als auch in N enthalten. Also ist $x \in L \cup (M \cap N)$: \square

In manchen Fällen muss man den Durchschnitt und die Vereinigung über eine beliebige Anzahl von Mengen bilden.

Definition 1.1.8. Eine Familie von Mengen $(M_\alpha)_{\alpha \in I}$ ordnet jedem α einer Menge I (*Indexmenge*) eine Menge M_α zu. Die Vereinigung bzw. der Durchschnitt einer solchen Familie ist wie folgt definiert:

$$\bigcup_{\alpha \in I} M_\alpha := \{x \mid \exists \alpha \in I: x \in M_\alpha\}$$

$$\bigcap_{\alpha \in I} M_\alpha := \{x \mid \forall \alpha \in I: x \in M_\alpha\}$$

Beispiel. Sei $M_n = \{k \in \mathbb{N} \mid k \geq n\}$ eine Familie von Teilmengen natürlicher Zahlen mit Indexmenge $I := \mathbb{N}$. Dann folgt: $\bigcup_{n \in \mathbb{N}} M_n = \mathbb{N}$, denn ist $k \in \mathbb{N}$, so ist $k \in M_1 = \mathbb{N}$, d.h.

$\mathbb{N} \subset \bigcup_{n \in \mathbb{N}} M_n$. Die Umkehrung folgt, denn $M_n \subset \mathbb{N}$ für jedes n .

$\bigcap_{n \in \mathbb{N}} M_n = \emptyset$, denn ist $k \in \bigcap_{n \in \mathbb{N}} M_n$, so ist $k \in M_n$ für alle $n \in \mathbb{N}$; aber $k \notin M_{k+1} = \{j \in \mathbb{N} \mid j \geq k+1\}$.

Eine weitere wichtige Methode, neue Mengen aus gegebenen zu konstruieren, ist das kartesische Produkt.

Definition 1.1.9. Es seien M, N Mengen. Dann heißt die Menge

$$M \times N = \{(x, y) \mid x \in M \text{ und } y \in N\}$$

das *kartesische Produkt* von M und N .

Beispiele.

a) $M = \{1, 2\}$, $N = \{0, 3\}$, $M \times N = \{(1, 0), (1, 3), (2, 0), (2, 3)\}$

b) $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R} = \{(x, y) \mid x \in \mathbb{R} \text{ und } y \in \mathbb{R}\}$
Den \mathbb{R}^2 stellt man sich als Ebene vor.

c) Sind M_1, \dots, M_n Mengen, und ist $n \in \mathbb{N}$, so definiert

$$M_1 \times \dots \times M_n = \{(x_1, \dots, x_n) \mid \forall i \in \{1, \dots, n\} : x_i \in M_i\}$$

das *n-fache kartesische Produkt*.

Elemente verschiedener Mengen sind oft durch eine Vorschrift einander zugeordnet. Die in der Mathematik wichtigste Zuordnung ist die Abbildung.

Definition 1.1.10. Es seien M, N nicht-leere Mengen. Eine *Abbildung*

$$f: M \rightarrow N$$

ist eine Vorschrift, die jedem $x \in M$ genau ein $y \in N$ zuordnet. Man schreibt: $f(x) = y$. Zwei Abbildungen $f, g: M \rightarrow N$ heißen *gleich* ($f = g$), falls

$$f(x) = g(x)$$

für alle $x \in M$ gilt. Die Menge M heißt *Definitionsbereich* von f ; die Menge N heißt *Wertebereich* von f .

Bemerkungen.

- (a) Jedem $x \in M$ wird ein Element $f(x) \in N$ zugeordnet. Dieses Element ist eindeutig bestimmt. Es können jedoch mehrere Elemente auf das gleiche Element in N abgebildet werden.
- (b) Manchmal nennt man Abbildungen auch Funktionen. Insbesondere dann, wenn der Wertebereich aus den reellen Zahlen besteht.

Beispiele. In der Schule sind meistens Abbildungen $f: \mathbb{R} \rightarrow \mathbb{R}$ betrachtet worden. Einfache Abbildungen von diesem Typ sind z.B.:

(1) lineare Abbildungen, d.h.

$$f(x) = m \cdot x, \quad m \in \mathbb{R}.$$

(2) Polynome, d.h.

$$f(x) = x^n + a_{n-1}x^{n-1} \dots + a_1x + a_0$$

mit $a_j \in \mathbb{R}$.

Von großer Bedeutung sind auch die Abbildungen

$$a: \mathbb{N} \rightarrow \mathbb{R}.$$

Sie heißen Folgen. In diesem Falle benutzt man auch die Bezeichnungen: a_n statt $a(n)$ und $(a_n)_{n \in \mathbb{N}}$ oder a_1, a_2, \dots statt $a: \mathbb{N} \rightarrow \mathbb{R}$. Beispiele sind: $a_n = \frac{1}{n}, a_n = \frac{5-n}{n^2}$ usw (diese werden in Kapitel 2 eine wichtige Rolle spielen).

Definition 1.1.11. Es sei $f: M \rightarrow N$ eine Abbildung und $A \subset M$. Dann heißt:

$$f(A) = \{f(x) \mid x \in A\}$$

das *Bild* von A unter f . Sei $B \subset N$, so heißt

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

das *Urbild* von B .

Zwischen Mengen und Abbildungen gelten die folgenden Beziehungen.

Satz 1.1.12. *Es seien M, N Mengen und $f: M \rightarrow N$ eine Abbildung.*

- (1) Für alle $A, B \subset M$ gilt:

- (a) $f(A \cup B) = f(A) \cup f(B)$
- (b) $f(A \cap B) \subset f(A) \cap f(B)$
- (c) $A \subset f^{-1}(f(A))$

(2) Für alle $A, B \subset N$ gilt:

- (a) $f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B)$
- (b) $f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B)$

Beweis. Wir zeigen nur die Beziehung (1a): Es seien $A, B \subset M$.

Es ist zu zeigen: für alle y : $y \in f(A \cup B) \Leftrightarrow y \in f(A) \cup f(B)$.

$$\begin{aligned} y \in f(A \cup B) &\Leftrightarrow \text{es existiert ein } x \in A \cup B \text{ mit } f(x) = y \\ &\Leftrightarrow \text{es existiert ein } x \in A \text{ mit } f(x) = y \text{ oder ein } x \in B \text{ mit } f(x) = y \\ &\Leftrightarrow y \in f(A) \text{ oder } y \in f(B) \Leftrightarrow y \in f(A) \cup f(B) \end{aligned}$$

Der Beweis der restlichen Aussagen folgt mit ähnlichen Argumenten und ist daher als Übung überlassen. \square

Bemerkung. Im allgemeinen gilt nur “ \subset ” in (1b), wie das folgende Beispiel zeigt. Wähle $f: \mathbb{R} \rightarrow \mathbb{R}$ mit $f(x) = x^2$, $A = \{x \mid -1 \leq x \leq 0\}$ und $B = \{x \mid 0 \leq x \leq 1\}$. Dann folgt $f(A \cap B) = f(\{0\}) = \{0\}$ und $f(A) = f(B) = B$.

Abbildungen kann man unter gewissen Voraussetzungen miteinander verknüpfen.

Definition 1.1.13. Sind $f: M \rightarrow N$ und $g: N \rightarrow L$ Abbildungen, so heißt $g \circ f: M \rightarrow L$ mit

$$(g \circ f)(x) := g(f(x))$$

die *Komposition* von g und f .

Beispiel. $f: \mathbb{R} \rightarrow \mathbb{R}$ mit $f(x) = x^2 + x$, $g: \mathbb{R} \rightarrow \mathbb{R}$ mit $g(x) = x^2$. Dann gilt: $g \circ f(x) = g(x^2 + x) = (x^2 + x)^2 = x^4 + 2x^3 + x^2$.

Von großer Bedeutung sind folgende Typen von Abbildungen.

Definition 1.1.14. Eine Abbildung $f: M \rightarrow N$ heißt

injektiv: $\forall x, y \in M: x \neq y \Rightarrow f(x) \neq f(y)$,

surjektiv: $\forall y \in N \exists x \in M: f(x) = y$,

bijektiv: f ist injektiv und surjektiv.

Bemerkung. Es gilt auch: f ist injektiv $\Leftrightarrow \forall x, y \in M: f(x) = f(y) \Rightarrow x = y$.

Dies folgt aus dem *Kontrapositionsgesetz*, was für Aussagen A, B besagt:

$$(A \Rightarrow B) \Leftrightarrow (\neg B \Rightarrow \neg A)$$

Dieses Gesetz lässt sich leicht mit Hilfe einer Wahrheitstabelle beweisen.

Jede bijektive Abbildung besitzt eine Umkehrabbildung.

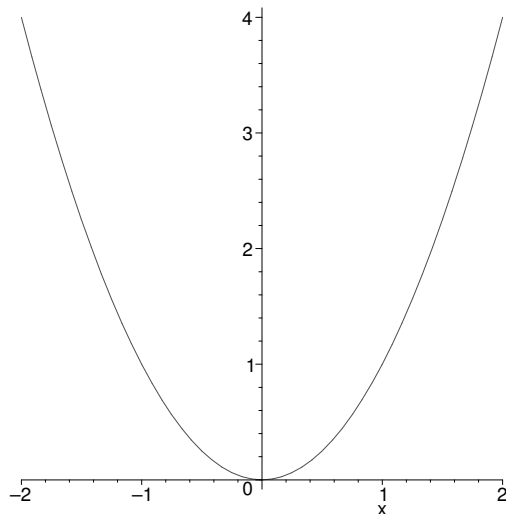


Abbildung 1.1: Der Graph von $f: \mathbb{R} \rightarrow \mathbb{R}$ mit $f(x) = x^2$

Satz 1.1.15. *Es sei $f: M \rightarrow N$ eine bijektive Abbildung. Dann existiert genau eine Abbildung $g: N \rightarrow M$ mit*

$$g \circ f(x) = x \text{ und } f \circ g(y) = y$$

für alle $x \in M$ und alle $y \in N$.

Bemerkung. Wir schreiben f^{-1} für g und nennen f^{-1} die *inverse Abbildung* zu f .

Warnung: Verwechseln Sie die Abbildung f^{-1} nicht mit der Urbildmenge $f^{-1}(B)$. Für $B \subset N$ ist die Menge $f^{-1}(B)$ für *beliebige* Abbildungen f definiert.

Beweis. *Eindeutigkeit:* Seien g_1, g_2 zwei Abbildungen mit obiger Eigenschaft, so gilt

$$g_1 \circ f(x) = g_2 \circ f(x)$$

für alle $x \in M$. Ist $y \in N$, so existiert $x \in M$ mit $f(x) = y$ (denn f ist surjektiv). Daraus folgt: $g_1(y) = g_2(y)$.

Existenz: Sei $y \in N$. Da f bijektiv ist, gibt es ein eindeutig bestimmtes $x \in M$ mit $f(x) = y$. Definieren wir $g(y) = x$, so gilt:

$$g \circ f(x) = g(y) = x \quad \text{und} \quad f \circ g(y) = f(x) = y.$$

□

Jeder Abbildung $f: M \rightarrow N$ ist eine Menge zugeordnet, die wir den Graph von f nennen. Insbesondere eignet sich diese Menge dazu, Abbildungen zu veranschaulichen (siehe Abbildung 1.1)

Definition 1.1.16. Sei $f: M \rightarrow N$ eine Abbildung, so nennt man

$$G_f = \{(x, y) \in M \times N \mid f(x) = y\}$$

den *Graph* von f .

1.2 Die Rechenregeln für die reellen Zahlen (Körperaxiome)

Die wichtigste Menge, die wir in diesem Semester betrachten werden, ist die Menge der reellen Zahlen. Anschaulich stellen wir uns diese Menge als Punkte auf einer Geraden "lückenlos" angeordnet vor (Zahlengerade). Man kann mit etwas größerem und zeitintensivem Aufwand die reellen Zahlen aus den natürlichen Zahlen konstruieren. Wir werden den etwas pragmatischeren und bequemeren Weg gehen, indem wir die *reellen Zahlen axiomatisch definieren*. Wir werden also festlegen, welche Eigenschaften (Axiome) die reellen Zahlen erfüllen sollen. Aus den Axiomen lassen sich dann alle weiteren Eigenschaften ableiten.

Beginnen werden wir mit den *Körperaxiomen*. Diese legen die grundlegenden Rechenregeln fest.

Die reellen Zahlen \mathbb{R} bilden eine Menge mit zwei Verknüpfungen $(+, \cdot)$, d.h. zwei Abbildungen

$$+ : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}, \quad (x, y) \rightarrow +(x, y) =: x + y$$

und

$$\cdot : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}, \quad (x, y) \rightarrow \cdot(x, y) =: x \cdot y.$$

Bezüglich der Addition bildet \mathbb{R} eine kommutative Gruppe, d.h. es gelten die *Axiome der Addition*:

- A1: Für alle $x, y, z \in \mathbb{R}$ gilt: $(x + y) + z = x + (y + z)$ *Assoziativgesetz*;
- A2: Für alle $x, y \in \mathbb{R}$ gilt: $x + y = y + x$ *Kommutativgesetz*;
- A3: Es gibt ein *neutrales Element* $0 \in \mathbb{R}$,
d.h. für alle $x \in \mathbb{R}$ gilt: $x + 0 = x$;
- A4: Für alle $x \in \mathbb{R}$ existiert ein $y \in \mathbb{R}$ mit $x + y = 0$.
 y heißt das zu x *inverse Element*.

Die Axiome A1 - A4 besagen, dass $(\mathbb{R}, +)$ eine *abelsche (kommutative) Gruppe* darstellt.

Bevor wir die weiteren Axiome behandeln, wollen wir an dieser Stelle schon einige Schlußfolgerungen ziehen.

Bemerkungen.

- (a) Das neutrale Element ist eindeutig bestimmt, denn ist $0'$ ein weiteres neutrales Element, so gilt:

$$0' \stackrel{(1)}{=} 0' + 0 \stackrel{(2)}{=} 0 + 0' \stackrel{(3)}{=} 0.$$

((1) folgt aus der Neutralität von 0, (3) folgt aus der Neutralität von $0'$ und (2) ist das Kommutativgesetz A2 für $y + x = x + y = 0$).

- (b) Zu jedem $x \in \mathbb{R}$ gibt es genau ein inverses Element, denn ist $x + y = x + y' = 0$, so folgt:

$$y = y + (x + y') \stackrel{(1)}{=} (y + x) + y' \stackrel{(2)}{=} y'.$$

(In (1) benutzen wir das Assoziativgesetz A3, in (2) das Kommutativgesetz A2).

(c) Das zu x inverse Element bezeichnen wir mit $-x$ und setzen

$$x - y := x + (-y)$$

Bemerkung. Man kann die Eindeutigkeit des neutralen und inversen Elementes auch für beliebige (nicht kommutative) Gruppen beweisen.

Nun wollen wir Existenz und Eindeutigkeit der Lösung einer Gleichung beweisen, die wir aus der Schule längst kennen. Wir wollen allerdings nur die uns gegebenen Rechengesetze verwenden.

Satz 1.2.1. *Seien $a, b \in \mathbb{R}$. Dann hat die Gleichung*

$$a + x = b$$

genau eine Lösung, nämlich $x = b - a$.

Beweis. Wir werden die Existenz und Eindeutigkeit der Lösung simultan zeigen, in dem wir äquivalente Umformungen durchführen.

Sei $x \in \mathbb{R}$, so gilt:

$$a + x = b \stackrel{(1)}{\Leftrightarrow} (-a) + (a + x) = (-a) + b$$

((1) Addition von $(-a)$ bzw. a auf beiden Seiten)

$$\stackrel{(2)}{\Leftrightarrow} ((-a) + a) + x = (-a) + b$$

((2) benutzt das Assoziativgesetz)

$$\Leftrightarrow 0 + x = b - a \Leftrightarrow x = b - a$$

(Die restlichen Äquivalenzen benutzen das Kommutativgesetz, und die Definition des inversen und neutralen Elementes.)

Dieser Beweis zeigt: Ist x eine Lösung von $a + x = b$, so ist $x = b - a$, d.h. die Lösung ist eindeutig. Ist umgekehrt $x = b - a$, so ist $a + x = b$, d.h. die Lösung existiert. Dies ist genau das, was im Satz behauptet wurde. \square

Satz 1.2.2. (a) *Für alle $x \in \mathbb{R}$ gilt: $-(-x) = x$.*

(b) *Für alle $x, y \in \mathbb{R}$ gilt: $-(x + y) = (-x) + (-y)$.*

Beweis.

(a) Nach Definition gilt: $-(-x)$ ist das zu $(-x)$ inverse Element.

Da $(-x) + x = x + (-x) = 0$, ist x invers zu $(-x)$. Damit folgt aus der Eindeutigkeit des inversen Elementes: $x = -(-x)$.

$$\begin{aligned} \text{(b)} \quad (x + y) + ((-x) + (-y)) &\stackrel{(1)}{=} ((x + y) + (-x)) + (-y) \stackrel{(2)}{=} ((y + x) + (-x)) + (-y) \stackrel{(3)}{=} \\ &(y + (x + (-x))) + (-y) \stackrel{(4)}{=} (y + 0) + (-y) \stackrel{(5)}{=} y + (-y) = 0 \end{aligned}$$

(1),(3) folgt aus dem Assoziativgesetz, (2) aus dem Kommutativgesetz, (4) aus A4, (5) aus (A3). \square

Nun betrachten wir die *Axiome der Multiplikation*.

M1: Für alle $x, y, z \in \mathbb{R}$ gilt: $(x \cdot y) \cdot z = x \cdot (y \cdot z)$

Assoziativgesetz;

M2: Für alle $x, y \in \mathbb{R}$ gilt: $x \cdot y = y \cdot x$

Kommutativgesetz;

M3: Es existiert ein *neutrales Element* $1 \in \mathbb{R} \setminus \{0\}$,
mit $x \cdot 1 = x$ für alle $x \in \mathbb{R}$

M4: Für alle $x \in \mathbb{R} \setminus \{0\}$ existiert ein $y \in \mathbb{R}$ mit $x \cdot y = 1$.
 y heißt das zu x *inverse Element*.

D: Für alle $x, y, z \in \mathbb{R}$ gilt: $x \cdot (y + z) = x \cdot y + x \cdot z$

Distributivgesetz

Bemerkung. Die Assoziativgesetze M1, A1 ermöglichen es, Klammern wegzulassen. Deshalb können wir definieren:

$$x \cdot y \cdot z := x \cdot (y \cdot z) = (x \cdot y) \cdot z, \quad x + y + z := x + (y + z) = (x + y) + z.$$

Statt $x \cdot y$ schreiben wir auch xy .

Genau wie im Falle der Addition verifiziert man folgende

Bemerkungen.

(a) Das neutrale Element bezüglich der Multiplikation ist eindeutig bestimmt.

(b) Zu jedem $x \in \mathbb{R} \setminus \{0\}$ gibt es genau ein inverses Element. Es wird mit x^{-1} bezeichnet.

Satz 1.2.3. Für alle $x \in \mathbb{R}$ gilt: $x \cdot 0 = 0$.

Beweis. Sei $x \in \mathbb{R}$, so folgt: $x \cdot 0 = x \cdot (0 + 0) = x \cdot 0 + x \cdot 0 \Rightarrow x \cdot 0 + (-(x \cdot 0)) = x \cdot 0$.
Dies impliziert: $0 = x \cdot 0$. □

Bemerkung. Ist $x \neq 0$, so folgt $x^{-1} \neq 0$. Denn ist $x^{-1} = 0$, so gilt wegen Satz 1.2.3: $1 = x \cdot x^{-1} = 0$, aber $1 \neq 0$ wegen M3.

Analog zu Satz 1.2.1 und Satz 1.2.2 gilt:

Satz 1.2.4. Seien $a, b \in \mathbb{R}$ und $a \neq 0$. Dann hat die Gleichung $a \cdot x = b$ genau eine Lösung, nämlich $x = b \cdot a^{-1}$.

Beweis. Man zeigt analog zu Satz 1.2.1: $x = b \cdot a^{-1}$ ist die eindeutig bestimmte Lösung. □

Bemerkung. Man schreibt auch: $\frac{b}{a} = b \cdot a^{-1}$ und nennt $\frac{b}{a}$ den Quotienten von b und a .

Satz 1.2.5.

(a) $\forall x \in \mathbb{R} \setminus \{0\} : (x^{-1})^{-1} = x$.

(b) $\forall x, y \in \mathbb{R} \setminus \{0\} : (x \cdot y)^{-1} = x^{-1} \cdot y^{-1}$.

Das Produkt zweier reeller Zahlen ist ungleich Null, falls alle ihre Faktoren ungleich Null sind.

Satz 1.2.6. Für alle $x, y \in \mathbb{R}$ gilt: Ist $x \cdot y = 0$, so ist $x = 0$ oder $y = 0$.

Beweis. Seien $x, y \in \mathbb{R}$ und $x \cdot y = 0$. Wir können annehmen: $x \neq 0$. Dann folgt aus Satz 1.2.3: $0 = x^{-1} \cdot 0 = x^{-1}(x \cdot y) = y$. \square

Der nächste Satz verknüpft multiplikative mit additiven Strukturen.

Satz 1.2.7. $\forall a, b \in \mathbb{R} : (-a)b = -(ab)$.

Beweis. Wegen Satz 1.2.1 hat die Gleichung $ab + x = 0$ genau eine Lösung, nämlich $x = -(a \cdot b)$. Wir zeigen: $(-a)b$ ist ebenfalls eine Lösung.

$$ab + (-a)b = (a + (-a)) \cdot b = 0 \cdot b = 0.$$

\square

Korollar 1.2.8. $\forall a, b \in \mathbb{R} : (-a)(-b) = a \cdot b$.

Beweis. $(-a)(-b) \stackrel{(1)}{=} -(a(-b)) \stackrel{(2)}{=} -(-(ab)) \stackrel{(3)}{=} ab$. (1) und (2) folgen aus Satz 1.2.7 und (3) aus Satz 1.2.2. \square

Bemerkung. Insbesondere gilt: $(-1)(-1) = +1$.

Definition 1.2.9. Eine Menge K mit Operationen $(+, \cdot)$ für die die Axiome A1 - A4 sowie M1 - M4, D gelten, heißt *Körper*.

Bemerkung. Die Aussagen, die wir bisher abgeleitet haben, gelten für beliebige Körper. Die Körperaxiome charakterisieren die reellen Zahlen keineswegs. Zum Beispiel bilden die rationalen Zahlen einen Körper. Es existieren sogar endliche Körper, d.h. Körper mit endlich vielen Elementen.

Die reellen Zahlen sind auf einer Zahlengeraden angeordnet. Dies läßt sich durch die Anordnungsaxiome formalisieren.

1.3 Die Anordnungsaxiome der reellen Zahlen

Auf \mathbb{R} existiert eine Ordnungsrelation " \leq ". Allgemein versteht man unter einer Ordnungsrelation das Folgende.

Definition 1.3.1. Sei M eine Menge. Dann heißt " \leq " eine *Ordnungsrelation* auf M , falls

- R1: Für alle $x \in M : x \leq x$ *Reflexivität*
- R2: Für alle $x, y \in M : x \leq y$ und $y \leq x \Rightarrow x = y$ *Antisymmetrie*
- R3: Für alle $x, y, z \in M : x \leq y; y \leq z \Rightarrow x \leq z$ *Transitivität*
- R4: Eine Ordnung heißt *total*, falls für alle $x, y \in M : x \leq y$ oder $y \leq x$.

Beispiel. Ist X eine Menge, so definiert “ \subset ” eine Ordnung auf $M = \mathcal{P}(X)$. Diese ist jedoch nicht total.

Nun verlangen wir, dass die reellen Zahlen die folgenden Anordnungsaxiome erfüllen sollen. Sie bestehen aus einer totalen Ordnung, die kompatibel zur additiven und multiplikativen Struktur der reellen Zahlen ist.

Anordnungsaxiome:

Auf \mathbb{R} ist eine totale Ordnung “ \leq ” mit folgenden Eigenschaften gegeben:

O1: Für alle $x, y, z \in \mathbb{R} : x \leq y \Rightarrow x + z \leq y + z$
(*Monotonie der Addition*)

O2: Für alle $x, y, z \in \mathbb{R} : x \leq y$ und $z \geq 0 \Rightarrow x \cdot z \leq y \cdot z$
(*Monotonie der Multiplikation*)

Bemerkung. Wir schreiben $x < y$ genau dann, falls $x \leq y$ und $x \neq y$. Dann folgt in Ergänzung zu O1, O2 und R3:

O1': Für alle $x, y, z \in \mathbb{R} : x < y \Rightarrow x + z < y + z$,
denn aus $x + z = y + z$ folgt $x = y$.

O2': Für alle $x, y, z \in \mathbb{R} : x < y$ und $z > 0 \Rightarrow xz < yz$,
denn $xz = yz$ impliziert $xzz^{-1} = yzz^{-1}$ und somit $x = y$.

R3': $x \leq y$ und $y < z \Rightarrow x < z$
denn wäre $x = z$ so folgt aus $(x \leq y$ und $y < z)$ mit R2: $x = y = z$.

Folgerungen aus den Anordnungsaxiomen:

Satz 1.3.2.

(1) Für alle $a, b, c, d \in \mathbb{R}$ gilt:

$$a \leq b \text{ und } c \leq d \Rightarrow a + c \leq b + d$$

$$a \leq b \text{ und } c < d \Rightarrow a + c < b + d$$

(2) Für alle $a, b, c \in \mathbb{R}$ folgt:

$$(i) c \leq 0 \Rightarrow -c \geq 0, \quad c < 0 \Rightarrow -c > 0$$

$$(ii) a \leq b \text{ und } c \leq 0 \Rightarrow ac \geq bc$$

$$(iii) a < b \text{ und } c < 0 \Rightarrow ac > bc$$

(3) Für alle $a \in \mathbb{R}$ gilt: $a^2 \geq 0$. Ist $a \neq 0$ so ist $a^2 > 0$.

(*Insbesondere gilt: $1 = 1 \cdot 1 > 0$*)

(4) Für alle $a \in \mathbb{R}$ folgt: $a > 0 \Rightarrow \frac{1}{a} > 0$.

(5) Für alle $a, b \in \mathbb{R} : 0 < a \leq b \Rightarrow \frac{1}{a} \geq \frac{1}{b}$ sowie $0 < a < b \Rightarrow \frac{1}{a} > \frac{1}{b}$.

Beweis.

(1) $a \leq b \stackrel{(O1)}{\Rightarrow} a + c \leq b + c$. Aus $c \leq d \stackrel{(O1)}{\Rightarrow} b + c \leq b + d$. Damit folgt aus (R3) $a + c \leq b + d$.

Ist zusätzlich $c < d \stackrel{(O1')}{\Rightarrow} b + c < b + d$ und damit $a + c < b + d$ wegen (R3').

- (2) (i) Ist $c \leq 0 \stackrel{(O1)}{\Rightarrow} c + (-c) \leq -c \Rightarrow 0 \leq -c$.
Ist $c < 0$, so folgt wegen Satz 1.2.7 $-c = (-1)c$. Da $-1 \neq 0$ und $c \neq 0$, folgt wegen Satz 1.2.6 $-c \neq 0$. Da $-c \geq 0$, folgt damit: $-c > 0$.
- (ii) Ist $a \leq b$ und $c \leq 0 \stackrel{(O2),(2i)}{\Rightarrow} a(-c) \leq b(-c) \stackrel{(\text{Satz 1.2.7})}{\Leftrightarrow} -ac \leq -bc$. Addieren wir auf beiden Seiten ac , so erhalten wir aus O1: $0 \leq -bc + ac$. Nochmalige Addition von bc auf beiden Seiten ergibt: $bc \leq ac$.
- (iii) Ist $a < b$ und $c < 0$, so folgt $bc < ac$, denn ist $bc = ac$, so ist $a = b$, da $c \neq 0$.
- (3) Ist $a \geq 0$, so folgt dies direkt aus O2.
Ist $a \leq 0$ so folgt aus (2ii) nach Multiplikation mit a :
 $a^2 = a \cdot a \geq a \cdot 0 = 0$.
Ist zusätzlich $a \neq 0$, so ist $a^2 \neq 0$ wegen Satz 1.2.6. Also gilt: $a^2 > 0$.
- (4) Sei $a > 0$.
Wäre $\frac{1}{a} \leq 0$, so würde aus O2 folgen:
 $1 = \frac{1}{a}a \leq 0 \cdot a = 0$. Diese Aussage ist wegen (3) falsch (Widerspruch).
- (5) Ist $0 < a \leq b$. Da wegen (4) $\frac{1}{a}, \frac{1}{b} > 0$, folgt aus O2 durch Multiplikation mit $\frac{1}{a}$: $0 < 1 \leq \frac{b}{a}$.
Nochmalige Multiplikation mit $\frac{1}{b}$ ergibt: $0 < \frac{1}{b} \leq \frac{b}{a} \cdot \frac{1}{b} = ba^{-1}b^{-1} = \frac{1}{a}$.
Sei $0 < a < b$. Dann ist $\frac{1}{a} \neq \frac{1}{b}$, denn aus $\frac{1}{a} = \frac{1}{b}$ würde $a = b$ folgen. Somit gilt auch die zweite Ungleichung.

□

Definition 1.3.3. Sei $D \subset \mathbb{R}$ und $f : D \rightarrow \mathbb{R}$ eine Funktion. $f : D \rightarrow \mathbb{R}$ heißt *monoton wachsend* (*streng monoton wachsend*), falls

$$\forall x, y \in D : x < y \Rightarrow f(x) \leq f(y) \quad (f(x) < f(y)).$$

$f : D \rightarrow \mathbb{R}$ heißt *monoton fallend* (*streng monoton fallend*), falls

$$\forall x, y \in D : x < y \Rightarrow f(x) \geq f(y) \quad (f(x) > f(y)).$$

Bemerkungen.

- (1) Streng monotone (streng monoton wachsende oder fallende) Funktionen sind offensichtlich injektiv. Sie besitzen daher eine inverse Abbildung $f^{-1} : f(D) \rightarrow \mathbb{R}$. Diese Abbildung ist ebenfalls streng monoton und hat denselben Monotonietyp wie f . Denn ist z.B. f monoton wachsend und $y_1, y_2 \in f(D)$ mit $y_1 < y_2$, so existieren $x_1, x_2 \in D$ mit $f(x_1) = y_1$ und $f(x_2) = y_2$. Dann ist $f^{-1}(y_1) = x_1 < x_2 = f^{-1}(y_2)$, denn wäre $x_2 \leq x_1$, so würde, da f monoton wachsend ist, $y_2 = f(x_2) \leq f(x_1) = y_1$ folgen.
- (2) Ist $\mathbb{R}_+ = \{x \in \mathbb{R} \mid x > 0\}$, so besagt 1.3.2(5), dass die Funktion $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ mit $f(x) = \frac{1}{x}$ streng monoton fallend ist.

Unmittelbar aus den Anordnungsaxiomen erhalten wir:

Satz 1.3.4. Die Abbildung $f : \mathbb{R}_+ \cup \{0\} \rightarrow \mathbb{R}_+ \cup \{0\}$ mit $f(x) = x^2$ ist streng monoton wachsend.

Beweis. Es sei $0 \leq x_1 < x_2$. Multiplizieren wir diese Ungleichung mit x_1 und x_2 , so folgt aus O2': $x_1^2 < x_2x_1$ und $x_2x_1 < x_2^2$. Damit folgt die Behauptung aus R3'. □

Bemerkungen.

- (1) Damit ist f insbesondere eine injektive Abbildung. Der Beweis der Surjektivität verlangt das Vollständigkeitsaxiom (s. Abschnitt 2.4).
- (2) Die Umkehrabbildung von $f(x) = x^2$ heißt *Quadratwurzel* ($\sqrt{\quad}$). Da wir noch nicht gezeigt haben, dass f surjektiv ist, ist sie zunächst nur auf $f(\mathbb{R}_+ \cup \{0\})$ definiert. Auf dieser Menge ist wegen der obigen Bemerkung die Wurzel streng monoton wachsend.

Mit Hilfe der Anordnungsaxiome definiert man den Betrag einer reellen Zahl.

Definition 1.3.5. Sei $a \in \mathbb{R}$, so heißt

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

der *Absolutbetrag* (*Betrag*) von a .

Bemerkung. $|a|$ beschreibt den ‘‘Abstand’’ von a zur Null. Da $-a > 0$ falls $a < 0$, gilt: $|a| \geq 0$. Des weiteren folgt: $-|a| \leq a \leq |a|$, denn $a \in \{\pm|a|\}$.

Satz 1.3.6. Sei $\epsilon \geq 0$, so gilt für alle $a \in \mathbb{R}$:

$$|a| \leq \epsilon \Leftrightarrow -\epsilon \leq a \leq \epsilon.$$

Beweis. Betrachte zunächst den Fall $a \geq 0$. Dann gilt: $a = |a| \leq \epsilon \Leftrightarrow -\epsilon \leq a \leq \epsilon$. Ist $a < 0$, so ist $-a = |a| \leq \epsilon$ ebenfalls äquivalent zu $-\epsilon \leq a \leq \epsilon$. □

Bemerkung. Ist $\epsilon > 0$, so gilt für alle $a \in \mathbb{R}$: $|a| < \epsilon \Leftrightarrow -\epsilon < a < \epsilon$.

Satz 1.3.7. Für alle $a, b \in \mathbb{R}$ gilt:

- (1) $|a| = 0 \Leftrightarrow a = 0$.
- (2) $|ab| = |a||b|$.
- (3) $|a + b| \leq |a| + |b|$ (Dreiecksungleichung).

Bemerkung. Die Dreiecksungleichung ist, wie wir sehen werden, von fundamentaler Bedeutung für die Analysis.

Beweis.

- (1) $|a| = 0 \Leftrightarrow a = |a| = 0$.
- (2) Da $a = |a|$ oder $-a = |a|$, folgt:
 - (i) $|a||b| = ab$ oder
 - (ii) $|a||b| = -ab$.

Gilt (i), so ist $ab \geq 0$ und somit $|a \cdot b| = a \cdot b$. Gilt (ii), so folgt $a \cdot b \leq 0$ und somit $-ab = |a \cdot b|$.

- (3) Wegen Satz 1.3.6 gilt: $|a + b| \leq |a| + |b| \Leftrightarrow -(|a| + |b|) \leq a + b \leq (|a| + |b|)$.
Da wegen obiger Bemerkung $-|a| \leq a \leq |a|$ und $-|b| \leq b \leq |b|$, folgt aus Satz 1.3.2(1):
 $-|a| - |b| \leq a + b \leq |a| + |b|$.

□

Die Anordnungs- und Körperaxiome reichen noch nicht aus, um die reellen Zahlen zu kennzeichnen. Bevor wir die axiomatische “Definition” der reellen Zahlen fortsetzen, wollen wir zunächst zeigen, wie wir die natürlichen Zahlen in dem angeordneten Körper \mathbb{R} definieren können. Man stellt sich vor, dass die natürlichen Zahlen, ausgehend von der 1, durch fortgesetzte Addition der 1 erzeugt werden. Wir werden dies im nächsten Abschnitt formalisieren.

1.4 Die natürlichen Zahlen

Eine wichtige Eigenschaft der natürlichen Zahlen ist die Induktivität.

Definition 1.4.1. Eine Teilmenge $M \subset \mathbb{R}$ heißt *induktiv*, falls

- (1) $1 \in M$,
- (2) $\forall x \in \mathbb{R} : (x \in M \Rightarrow x + 1 \in M)$.

Beispiel. Die Menge $\mathbb{R}_+ = \{x \in \mathbb{R} \mid x > 0\}$ ist induktiv, da $1 > 0$ und $x > 0 \stackrel{O1'}{\Rightarrow} x + 1 > 1 \stackrel{R3}{\Rightarrow} x + 1 > 0$.

Diese Menge ist natürlich zu groß. Daher definieren wir \mathbb{N} als die kleinste induktive Menge, genauer:

Definition 1.4.2.

$$\begin{aligned} \mathbb{N} &:= \cap \{M \mid M \subset \mathbb{R}, M \text{ induktiv} \} \\ &:= \{x \in \mathbb{R} \mid x \in M \text{ für alle induktiven Mengen } M\} \end{aligned}$$

Bemerkung. \mathbb{N} ist also die kleinste induktive Menge, d.h. \mathbb{N} ist induktiv und in jeder anderen induktiven Menge enthalten.

Eine Konsequenz ist das folgende wichtige Induktionsprinzip.

Satz 1.4.3. (Prinzip der vollständigen Induktion)

Sei $M \subset \mathbb{N}$ und es gelte

- (1) $1 \in M$.
- (2) $\forall n \in \mathbb{N} : (n \in M \Rightarrow n + 1 \in M)$.

Dann folgt $M = \mathbb{N}$.

Beweis. Aus (1) und (2) folgt, dass M induktiv ist. Da nach Definition 1.4.2 \mathbb{N} in jeder induktiven Menge enthalten ist, folgt: $\mathbb{N} \subset M$. Auf der anderen Seite gilt nach Voraussetzung an M : $M \subset \mathbb{N}$, also folgt $M = \mathbb{N}$. \square

Anwendung findet dieses Prinzip in der Beweismethode der vollständigen Induktion.

Sei A eine Aussageform mit Grundbereich \mathbb{N} . Betrachte die Menge $M = \{n \in \mathbb{N} \mid A(n)\}$, d.h. die Menge aller natürlichen Zahlen $n \in \mathbb{N}$, für die die Aussage $A(n)$ wahr ist.

Korollar 1.4.4. Sei

- (1) $A(1)$ wahr,
- (2) $\forall n \in \mathbb{N} : (A(n) \Rightarrow A(n + 1))$.

Dann ist $A(n)$ eine wahre Aussage für alle $n \in \mathbb{N}$.

Beweis. Sei $M = \{n \in \mathbb{N} \mid A(n) \text{ wahr}\}$. Aus (1) und (2) folgt, dass die Menge M induktiv ist. Da $M \subset \mathbb{N}$, folgt aus dem Prinzip der vollständigen Induktion: $M = \mathbb{N}$, und damit ist die Aussage wahr für alle $n \in \mathbb{N}$. \square

Bemerkung. Ein Induktionsbeweis gliedert sich also in die folgenden zwei Schritte:

- (1) Man zeigt, dass $A(1)$ wahr ist (*Induktionsanfang*),
- (2) für jedes $n \in \mathbb{N}$ beweist man die Aussage $A(n) \Rightarrow A(n+1)$ (*Induktionsschritt*). Dabei heißt $A(n)$ die *Induktionsvoraussetzung* und $A(n+1)$ die *Induktionsbehauptung*.

Nach unserer Definition der natürlichen Zahlen müssen die folgenden, offensichtlich erscheinenden Aussagen bewiesen werden. Die Beweismethode hierfür ist, die vollständige Induktion beweisen.

Satz 1.4.5. Für alle $m, n \in \mathbb{N}$ gilt:

- (a) $m + n \in \mathbb{N}$.
- (b) $m \cdot n \in \mathbb{N}$.
- (c) $n \geq 1$.
- (d) $n > m \Rightarrow n - m \in \mathbb{N}$.

Beweis. Wir beweisen nur die Aussage (a). Die Aussagen (b), (c) und (d) seien als Übung überlassen.

Sei $m \in \mathbb{N}$ fest gewählt.

Definiere $A(n) := m + n \in \mathbb{N}$.

Induktionsanfang: $A(1)$ ist wahr, denn $m + 1 \in \mathbb{N}$, da $m \in \mathbb{N}$ und \mathbb{N} eine induktive Menge ist.

Induktionsschritt: Sei $A(n)$ wahr, d.h. $m + n \in \mathbb{N}$. Aus dem Assoziativgesetz der Addition und der Induktivität von \mathbb{N} folgt

$$m + (n + 1) = (m + n) + 1 \in \mathbb{N},$$

d.h. $A(n + 1)$ ist eine wahre Aussage. \square

Aus den natürlichen Zahlen lassen sich die ganzen und rationalen Zahlen wie folgt konstruieren.

Definition 1.4.6. Die Menge

$$\mathbb{Z} := \{x \in \mathbb{R} \mid x \in \mathbb{N}\} \cup \{x \in \mathbb{R} \mid -x \in \mathbb{N}\} \cup \{0\}$$

ist die Menge der ganzen Zahlen. Die Menge

$$\mathbb{Z}_- := \{x \in \mathbb{R} \mid -x \in \mathbb{N}\} = \{x \in \mathbb{Z} \mid x \leq -1\}$$

ist die Menge der negativen ganzen Zahlen.

$$\mathbb{Q} := \left\{ \frac{m}{n} \mid m \in \mathbb{Z} \text{ und } n \in \mathbb{N} \right\}$$

ist die Menge der rationalen Zahlen.

Bemerkung. $\mathbb{Q} \subset \mathbb{R}$ ist ebenfalls ein angeordneter Körper. Wenn wir die axiomatische Definition von \mathbb{R} beendet haben, sehen wir, dass $\mathbb{Q} \neq \mathbb{R}$.

Manchmal möchten wir den Induktionsanfang bei einer beliebigen ganzen Zahl beginnen.

Satz 1.4.7. (verallgemeinerte Induktion)

Sei $k \in \mathbb{Z}$ und $\mathbb{N}_k := \{m \in \mathbb{Z} \mid m \geq k\}$. Sei A eine Aussageform mit Grundbereich \mathbb{N}_k . Sei

- (1) $A(k)$ wahr,
- (2) $\forall m \in \mathbb{N}_k : (A(m) \Rightarrow A(m+1))$.

Dann ist $A(m)$ wahr für alle $m \in \mathbb{N}_k$.

Beweis. Definiere die Aussageform $B(n) :\Leftrightarrow A(n+k-1)$ über \mathbb{N} . Dann folgt:

- (1) $B(1)$ ist wahr,
- (2) $\forall n \in \mathbb{N} : (B(n) \Rightarrow B(n+1))$.

Nach Korollar 1.4.4 ist $B(n)$ wahr für alle $n \in \mathbb{N}$. Da die Abbildung $f : \mathbb{N} \rightarrow \mathbb{N}_k$ mit $n \rightarrow n+k-1$ bijektiv ist, gilt $A(m)$ für alle $m \in \mathbb{N}_k$. \square

Sind $a_1, \dots, a_n \in \mathbb{R}, n \in \mathbb{N}$, so definieren wir rekursiv

$$\begin{aligned} \sum_{i=1}^1 a_i &:= a_1; & \sum_{i=1}^n a_i &= a_1 + \dots + a_n := \left(\sum_{i=1}^{n-1} a_i \right) + a_n \\ \prod_{i=1}^1 a_i &:= a_1; & \prod_{i=1}^n a_i &= a_1 \cdot \dots \cdot a_n := \left(\prod_{i=1}^{n-1} a_i \right) \cdot a_n. \end{aligned}$$

Ist $a_1 = a_2 = \dots = a_n = a$, so schreiben wir: $a^n := \prod_{i=1}^n a_i$ (n -te Potenz).

Für negative ganzzahlige Exponenten definieren wir: $a^0 := 1$ und $a^{-n} = \frac{1}{a^n}$ für $n \in \mathbb{N}$ falls $a \neq 0$.

Ein weiteres Beispiel zur vollständigen Induktion ist die folgende Ungleichung von Jakob Bernoulli (1654 - 1705).

Satz 1.4.8. (Bernoullische Ungleichung)

Sei $a \in \mathbb{R}, a \geq -1$. Dann gilt:

$$(1+a)^n \geq 1+na$$

für alle $n \in \mathbb{N}_0$.

Beweis. Für $n \in \mathbb{N}_0$ betrachte die Aussage

$$A(n) :\Leftrightarrow (1+a)^n \geq 1+na.$$

- (1) Induktionsanfang: $A(0)$ ist wahr, denn $(1+a)^0 = 1 = 1+0a$.
- (2) Induktionsschritt: $\forall n \in \mathbb{N}_0 : (A(n) \Rightarrow A(n+1))$.

Sei $n \in \mathbb{N}_0$ und es gelte: $(1+a)^n \geq 1+na$. Wegen $a \geq -1$ folgt aus O1: $1+a \geq 0$. O2 und die Induktionsvoraussetzung implizieren:

$$\begin{aligned} (1+a)^{n+1} &= (1+a)^n(1+a) \geq (1+na)(1+a) \\ &= (1+na)1 + (1+na)a = 1+na+a+na^2 \\ &\geq 1+(n+1)a. \end{aligned}$$

Die letzte Ungleichung folgt, da $a^2 \geq 0$ und $n \geq 0$. \square

Mit Hilfe der natürlichen Zahlen lassen sich endliche Mengen definieren. Eine Menge $M \neq \emptyset$ heißt *endlich*, falls es ein $n \in \mathbb{N}$ und eine bijektive Abbildung $f : \{1, \dots, n\} \rightarrow M$ gibt. Man kann mittels Induktion zeigen, dass die Zahl $n \in \mathbb{N}$ eindeutig bestimmt ist. Sie heißt die *Mächtigkeit* oder *Kardinalität* von M und sie beschreibt die Anzahl ihrer Elemente. Wir schreiben: $\#M := n$ oder auch $\text{card}(M)$ bzw. $|M|$ statt $\#M$. Man definiert $\#\emptyset := 0$.

Sei $M_n = \{1, \dots, n\}$ und $\mathcal{P}(M_n)$ die Menge aller Teilmengen von M_n .

Definition 1.4.9. Für $k \in \mathbb{N}_0$ und $0 \leq k \leq n$ definiere

$$\binom{n}{k} := \# \{A \in \mathcal{P}(M_n) \mid \# A = k\}.$$

Die Zahlen $\binom{n}{k}$ heißen *Binomialkoeffizienten*. Sie beschreiben die Anzahl aller k -elementigen Teilmengen von $\{1, \dots, n\}$. Dies lässt sich auch interpretieren als die Anzahl der Möglichkeiten aus der Zahlenmenge $\{1, \dots, n\}$ eine k -zählige Teilmenge auszuwählen.

Bemerkung. Offensichtlich gilt: $\binom{n}{0} = \binom{n}{n} = 1$, denn M_n enthält nur eine 0-elementige Menge (die leere Menge) und eine n -elementige Menge (die Menge M_n).

Satz 1.4.10. Es gilt für alle $n \in \mathbb{N}$ und $k \in \mathbb{N}$ mit $1 \leq k \leq n$

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

Beweis.

$$\begin{aligned} \{A \in \mathcal{P}(M_{n+1}) \mid \# A = k\} &= \{A \in \mathcal{P}(M_{n+1}) \mid \# A = k \text{ und } n+1 \in A\} \\ &\cup \{A \in \mathcal{P}(M_{n+1}) \mid \# A = k \text{ und } (n+1) \notin A\}. \end{aligned}$$

Diese Vereinigung ist disjunkt, d.h. ihr Durchschnitt ist leer.

Daher gilt:

$$\begin{aligned} \binom{n+1}{k} &= \# \{A \in \mathcal{P}(M_{n+1}) \mid \# A = k \text{ und } n+1 \in A\} + \\ &\# \{A \in \mathcal{P}(M_{n+1}) \mid \# A = k \text{ und } n+1 \notin A\}. \end{aligned}$$

Da die Abbildung

$$f : \{A \in \mathcal{P}(M_{n+1}) \mid \# A = k \text{ und } (n+1) \in A\} \rightarrow \{B \in \mathcal{P}(M_n) \mid \# B = k-1\}$$

mit $A \mapsto A \setminus \{n+1\}$ bijektiv ist, ist der erste Summand $\binom{n}{k-1}$.

Im zweiten Fall ist $(n+1) \notin A$, d.h. A kann als k -elementige Teilmenge von M_n aufgefaßt werden, d.h. der zweite Summand ist $\binom{n}{k}$. \square

Aus Satz 1.4.10 läßt sich das *Pascalsche Dreieck* ableiten.

$$\begin{array}{cccccccc} n=0 & & & & \binom{0}{0} = 1 & & & & \\ n=1 & & & & \binom{1}{0} = 1 & & \binom{1}{1} = 1 & & \\ n=2 & & & & \binom{2}{0} = 1 & & \binom{2}{1} = 2 & & \binom{2}{2} = 1 \\ n=3 & & & & \binom{3}{0} = 1 & & \binom{3}{1} = 3 & & \binom{3}{2} = 3 & & \binom{3}{3} = 1 \end{array}$$

$$\begin{array}{cccccc}
 & & & & & 1 \\
 & & & & & 1 & 1 \\
 & & & & 1 & 2 & 1 \\
 & & 1 & 3 & 3 & 1 \\
 1 & 4 & 6 & 4 & 1
 \end{array}$$

Definiere $0! = 1$ und $n! = 1 \cdot 2 \cdot \dots \cdot n$ für $n \in \mathbb{N}$.

Dann gilt:

Satz 1.4.11. Für $n \in \mathbb{N}_0$ und $k \in \mathbb{N}_0$ mit $0 \leq k \leq n$ folgt:

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

Beweis. Durch Induktion: Für $n \in \mathbb{N}_0$ definiere die Aussage

$$A(n) : \Leftrightarrow \forall k \in \{0, 1, \dots, n\} : \binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Induktionsanfang: $A(0)$ ist wahr, denn $\binom{0}{0} = 1$ und $\frac{0!}{0!0!} = 1$

Induktionsschritt: Für $n \in \mathbb{N}$ sei $A(n)$ wahr.

Zu zeigen ist: $A(n+1)$ ist wahr, d.h. $\binom{n+1}{k} = \frac{(n+1)!}{k!(n+1-k)!}$ für alle $k \in \{0, 1, \dots, n+1\}$.

Da $\binom{n+1}{0} = 1$ und $\frac{(n+1)!}{0!(n+1)!} = 1$ gilt diese Gleichung für $k=0$ und $k=n+1$.

Sei $1 \leq k \leq n$. Wegen Satz 1.4.10 und der Induktionsvoraussetzung gilt:

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

□

Satz 1.4.12. (Binomische Formel)

Sind $a, b \in \mathbb{R}$, so gilt für alle $n \in \mathbb{N}_0$

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

Beweis. Der Beweis wird mit Hilfe der vollständigen Induktion geführt und sei als Übung überlassen. □

Bemerkung.

Sei $M = \{1, \dots, n\}$, so folgt: $2^{\#M} = 2^n = (1+1)^n = \sum_{k=0}^n \binom{n}{k} = \# \mathcal{P}(M)$.

1.5 Vollständigkeitsaxiom

Die Körperaxiome und Anordnungsaxiome gelten auch für die rationalen Zahlen \mathbb{Q} . Diese haben jedoch nicht alle Eigenschaften, die wir von den reellen Zahlen verlangen wollen; z.B. hat die Gleichung $x^2 = 2$ keine Lösung in \mathbb{Q} , wie der folgende Widerspruchsbeweis zeigt. Sei $x \in \mathbb{Q}$ mit $x^2 = 2$. Dann existieren teilerfremde Zahlen $n \in \mathbb{Z}, m \in \mathbb{N}$ mit $x = \frac{n}{m}$. Dann folgt: $\frac{n^2}{m^2} = 2$, d.h. $n^2 = 2m^2$. Folglich ist n^2 eine gerade Zahl und somit ist auch n gerade. Damit existiert ein $k \in \mathbb{Z}$ mit $n = 2k$ und somit gilt $4k^2 = 2m^2$. Daraus folgt wiederum, dass m^2 und somit m gerade sind. Dann haben aber n und m den gemeinsamen Teiler 2, und dies steht im Widerspruch zur Annahme der Teilerfremdheit.

Bemerkung. Allgemein ist ein Beweis durch Widerspruch immer von der folgenden Form. Wir wollen aus einer Aussage A eine Aussage B herleiten, d.h. wir wollen zeigen, dass die Implikation $A \Rightarrow B$ wahr ist. Dazu müssen wir überprüfen: ist A wahr so auch B . In allen anderen Fällen ist $A \Rightarrow B$ trivialerweise erfüllt (siehe Bemerkung nach der Definition 1.1.4). Angenommen, wir können mit Hilfe von A aus $\neg B$ eine falsche Aussage C (Widerspruch) herleiten. Dann ist $\neg B$ ebenfalls falsch, und somit ist B wahr.

Definition 1.5.1. Sei $M \subset \mathbb{R}$.

(a) Dann heißt M *nach oben beschränkt*, falls ein $c \in \mathbb{R}$ existiert mit

$$x \leq c \text{ für alle } x \in M.$$

Man nennt c eine *obere Schranke* von M .

(b) s heißt *kleinste obere Schranke* von M , falls

- (i) s eine obere Schranke von M ist, und
- (ii) für alle oberen Schranken c von M gilt: $s \leq c$.

Außer den Körperaxiomen und den Anordnungsaxiomen verlangen wir nun für \mathbb{R} noch das wichtige Vollständigkeitsaxiom.

Vollständigkeitsaxiom

V: Jede nicht leere, nach oben beschränkte Teilmenge von \mathbb{R} besitzt eine kleinste obere Schranke.

Bemerkung. Die kleinste obere Schranke einer nach oben beschränkten Menge $M \subset \mathbb{R}$ ist eindeutig bestimmt, denn sind s_1, s_2 kleinste obere Schranken von M , so folgt: $s_1 \leq s_2$ und $s_2 \leq s_1$. Aus dem Anordnungsaxiom R2 folgt: $s_1 = s_2$.

Definition 1.5.2. Sei $M \subset \mathbb{R}$ nach oben beschränkt. Dann nennen wir die kleinste obere Schranke auch das *Supremum* von M . Wir bezeichnen diese Zahl mit $\sup M$.

Die *axiomatische Definition der reellen Zahlen* läßt sich nun wie folgt zusammenfassen. Die reellen Zahlen sind eine Menge \mathbb{R} mit folgenden Eigenschaften.

- (1) Es gibt Verknüpfungen $+, \cdot$, so dass $(\mathbb{R}, +, \cdot)$ ein Körper ist, d.h. es gelten die Körperaxiome A1, ..., A4, M1, ..., M4, D.
- (2) Es gibt eine totale Ordnung " \leq " auf \mathbb{R} , so dass der Körper $(\mathbb{R}, +, \cdot, \leq)$ angeordnet ist, d.h. es gelten die Axiome (R1, ..., R4) sowie O1 und O2.

(3) \mathbb{R} ist vollständig, d.h. es gilt das Vollständigkeitsaxiom V.

Bemerkung. Man kann zeigen, dass die reellen Zahlen bis auf "Äquivalenz" durch diese Axiome gekennzeichnet sind. Alles, was wir jemals mit reellen Zahlen machen werden, läßt sich auf diese Axiome zurückführen.

Nun können wir beweisen, dass die Gleichung $x^2 = 2$ eine Lösung in \mathbb{R} besitzt. Wir zeigen sogar:

Satz 1.5.3. Zu jedem $a \geq 0$ existiert genau eine reelle Zahl $x \geq 0$ mit

$$x^2 = a.$$

Wir nennen diese Zahl x die Quadratwurzel von a und schreiben $x = \sqrt{a}$ oder $x = a^{1/2}$.

Beweis.

(1) *Eindeutigkeit:*

Seien $x_1, x_2 \geq 0$ mit $x_1^2 = x_2^2 = a$. Dann folgt:

$$x_1^2 - x_2^2 = (x_1 - x_2)(x_1 + x_2) = 0.$$

Aus Satz 1.2.6 folgt: $x_1 - x_2 = 0$ oder $x_1 + x_2 = 0$.

Aus der ersten Gleichung folgt unmittelbar $x_1 = x_2$.

Da $x_2 \geq 0$, ist wegen Satz 1.3.2(i) $-x_2 \leq 0$. Aus der zweiten Gleichung folgt dann:

$0 \leq x_1 = -x_2 \leq 0$ und wegen des Anordnungsaxiom's R2 folgt $x_1 = 0$ und damit $x_2 = 0$.

(2) *Existenz:* (hier benötigen wir das Vollständigkeitsaxiom V)

Sei

$$M = \{x \geq 0 \mid x^2 \leq a\}.$$

Die Menge M ist nicht leer, denn sie enthält die 0. Außerdem ist $a + 1$ eine obere Schranke von M . Denn ist $x > a + 1$, so folgt: $x^2 > a^2 + 2a + 1 > 2a \geq a$ und daher $x \notin M$.

Daher existiert wegen des Vollständigkeitsaxiomes V eine kleinste obere Schranke $s = \sup M$. Da $0 \in M$, folgt $s \geq 0$.

Behauptung: $s^2 = a$.

Ist $s^2 \neq a$, so können zwei Fälle auftreten:

$$(1) s^2 < a \quad \text{oder} \quad (2) s^2 > a.$$

Wir zeigen nun, dass beide Fälle im Widerspruch zur Definition von s stehen.

(1) Sei $s^2 < a$.

Wie wir gleich sehen werden, existiert dann ein $h > 0$ mit $(s + h)^2 \leq a$.

Damit ist $s + h \in M$ und s keine obere Schranke.

Gesucht ist also ein $h > 0$ mit $(s + h)^2 = s^2 + 2sh + h^2 \leq a$. Dies ist äquivalent zu

$$h(2s + h) = 2sh + h^2 \leq c := a - s^2$$

Definiere h als das Minimum der Zahlen 1 und $\frac{c}{2s+1}$. Dann ist die obige Ungleichung erfüllt, denn:

$$h(2s + h) \leq h(2s + 1) \leq c.$$

(2) Sei $s^2 > a$.

Wir zeigen die Existenz von $h \in \mathbb{R}$ mit $0 < h < s$ und $(s-h)^2 \geq a$. Im Widerspruch zur Definition von s ist dann $s-h$ eine kleinere obere Schranke von M . Denn ist $x \in M$, so ist $x \geq 0$ und $(s-h)^2 \geq a \geq x^2$. Dann folgt wegen Satz 1.3.4 auch $s-h \geq x$.

Da $(s-h)^2 = s^2 - 2sh + h^2 > s^2 - 2sh$, erfüllt $h = \frac{s^2-a}{2s} \leq \frac{s^2}{2s} = \frac{1}{2}s$ die verlangte Eigenschaft. Denn dann ist $0 < h < s$ sowie $(s-h)^2 \geq s^2 - (s^2-a) = a$. □

Bemerkungen.

(a) Genauso kann man zeigen: Ist $n \in \mathbb{N}$ und $a \geq 0$, so existiert genau ein $x \geq 0$ mit $x^n = a$. Man nennt $x = \sqrt[n]{a} = a^{1/n}$ die n -te Wurzel aus a .

(b) Allgemeiner definiert man für $a \geq 0$ und $n, m \in \mathbb{N}$

$$a^{n/m} := (a^{1/m})^n.$$

Für $a > 0$ definiert man

$$a^{-n/m} := (a^{-1})^{n/m}.$$

Setzt man $a^0 := 1$, so erhält man die folgenden Rechenregeln:

$$a^{r+s} = a^r a^s; \quad a^{rs} = (a^r)^s; \quad \text{und} \quad a^r b^r = (ab)^r$$

für alle $r, s \in \mathbb{Q}$ sowie $a > 0$ und $b > 0$. Mit Hilfe der Exponentialfunktion werden wir später a^r für jedes $a > 0$ und $r \in \mathbb{R}$ definieren.

Definition 1.5.4. Eine Menge $M \subset \mathbb{R}$ heißt *nach unten beschränkt*, falls ein $c \in \mathbb{R}$ existiert mit $c \leq x$ für alle $x \in M$. Dann heißt c eine *untere Schranke* von M .

s heißt *größte untere Schranke*, falls s untere Schranke ist und für alle unteren Schranken c von M gilt: $c \leq s$.

Satz 1.5.5. Sei $M \neq \emptyset$ und M nach unten beschränkt. Dann existiert genau eine *größte untere Schranke*. Diese nennen wir auch das *Infimum* von M und wird mit $\inf M$ bezeichnet.

Beweis. Definiere $-M := \{-x \mid x \in M\}$.

Ist c untere Schranke von M , d.h. $c \leq x$ für alle $x \in M$, so folgt $-x \leq -c$ für alle $x \in M$ und somit ist $-M$ nach oben durch $(-c)$ beschränkt.

Wegen des Vollständigkeitsaxioms besitzt $-M$ eine kleinste obere Schranke $s := \sup -M$.

Behauptung: $-s$ ist größte untere Schranke von M .

(1) $-s$ ist untere Schranke von M , denn ist $x \in M$, so ist $-x \in -M$ und somit $-x \leq s$. Daher gilt: $x \geq -s$.

(2) $-s$ ist größte untere Schranke von M , denn ist c untere Schranke von M , so gilt für alle $x \in M$: $c \leq x \Leftrightarrow -c \geq -x$.

Daher ist $-c$ obere Schranke von $-M$. Da s kleinste obere Schranke ist, gilt $s \leq -c$ und somit $-s \geq c$, d.h. die untere Schranke c ist kleiner. □

Bemerkung. Der obige Beweis zeigt sogar: $-\sup(-M) = \inf M$.
Weitere Folgerungen aus dem Vollständigkeitsaxiom:

Satz 1.5.6. *Die reellen Zahlen sind archimedisch angeordnet, d.h. \mathbb{N} ist nicht nach oben beschränkt.*

Beweis. Annahme: \mathbb{N} wäre nach oben beschränkt. Aus der Vollständigkeit folgt dann die Existenz des Supremums $s := \sup \mathbb{N}$.

Da s kleinste obere Schranke von \mathbb{N} ist, ist $s - 1$ keine obere Schranke von \mathbb{N} , d.h. es gibt $n \in \mathbb{N}$ mit $s - 1 < n$. Insbesondere ist $s < n + 1 \in \mathbb{N}$ und somit ist s keine obere Schranke im Widerspruch zur Existenz des Supremums. \square

Dies hat zur Konsequenz, dass die Kehrwerte der natürlichen Zahlen beliebig klein werden.

Korollar 1.5.7. *Für alle $\epsilon > 0$ existiert ein $n \in \mathbb{N}$ mit $\frac{1}{n} < \epsilon$.*

Bemerkung. Mittels der Quantoren \forall (für alle) und \exists (es existiert) lässt sich der Satz auch knapp wie folgt schreiben:

$$\forall \epsilon > 0 \exists n \in \mathbb{N} : \frac{1}{n} < \epsilon.$$

Beweis. Sei $\epsilon > 0$. Dann folgt aus Satz 1.5.6:

$$\exists n \in \mathbb{N} \text{ mit } n > \frac{1}{\epsilon} \Rightarrow \epsilon > \frac{1}{n}. \quad \square$$

Manchmal ist das Supremum bzw. Infimum einer beschränkten Menge wieder ein Element dieser Menge. Dann spricht man vom Maximum bzw. Minimum. Genauer definiert man:

Definition 1.5.8. Sei $M \subset \mathbb{R}$ eine nach oben (bzw. nach unten) beschränkte Menge. Dann heißt s *Maximum (Minimum)* von M , falls $s \in M$ und $s = \sup M$ ($s = \inf M$). Man schreibt: $s = \max M$ ($s = \min M$).

Beispiel. $M = \{x \in \mathbb{R} \mid x < 1\}$.

Es gilt: $\sup M = 1$, aber $1 \notin M$, d.h. das Maximum existiert nicht.

Hingegen ist 1 das Maximum von $M = \{x \in \mathbb{R} \mid x \leq 1\}$.

Satz 1.5.9. *Sei $M \subset \mathbb{Z}$. Ist M nach oben beschränkt, so besitzt M ein Maximum. Ist M nach unten beschränkt, so existiert das Minimum von M .*

Beweis. Sei $M \subset \mathbb{Z}$ nach oben beschränkt und $s = \sup M$. Betrachte $s' = s - \frac{1}{2}$.

Da s die kleinste obere Schranke ist, ist s' keine obere Schranke und daher existiert ein $x \in M$ mit $s - \frac{1}{2} < x \leq s$. Dann ist x obere Schranke und somit $x = s$. Denn sonst existiert $y \in M$ mit $x < y \leq s$ (s ist obere Schranke) und es folgt: $y - x \in \mathbb{N}$.

Dann wäre aber $y - x \leq s - (s - \frac{1}{2}) = \frac{1}{2}$ im Widerspruch dazu, dass $n \geq 1$ für alle $n \in \mathbb{N}$.

Also folgt: $s = \max M$. \square

Definition 1.5.10. (*Gaußklammer*)

Sei $x \in \mathbb{R}$. Dann heißt

$$[x] = \max\{k \in \mathbb{Z} \mid k \leq x\} \in \mathbb{Z}$$

die *Gaußklammer* von $[x]$. Sie ist die größte ganze Zahl kleiner gleich x .

Bemerkung. Da $M = \{k \in \mathbb{Z} \mid k \leq x\}$ nach oben beschränkt ist, existiert das Maximum wegen Satz 1.5.9. Es gilt: $[x] \leq x < [x] + 1$.

Die erste Ungleichung gilt, da $[x] \in M$ (denn $[x]$ ist Maximum). Falls $[x] + 1 \leq x$, so wäre $[x] + 1 \in M$ und somit wäre $[x]$ keine obere Schranke von M .

Satz 1.5.11. *Es seien $x < y$ reelle Zahlen. Dann existiert eine rationale Zahl $q \in \mathbb{Q}$ mit $x < q < y$.*

Beweis. Ist $[x] + 1 < y$, so setze $q = [x] + 1 > x$. Ist $[x] \leq x < y \leq [x] + 1$, so wähle $n \in \mathbb{N}$, so dass $y - x > \frac{1}{n}$ (wegen Korollar 1.5.7 ist dies möglich).

Setze $q = [x] + \frac{m}{n} \in \mathbb{Q}$, wobei $m = m_0 + 1$ und $m_0 = \max\{k \in \mathbb{Z} \mid [x] + \frac{k}{n} \leq x\}$. Dann folgt: $x < [x] + \frac{m}{n}$, denn sonst ist m_0 nicht das Maximum.

Des weiteren gilt: $[x] + \frac{m}{n} = [x] + \frac{m_0}{n} + \frac{1}{n} \leq x + \frac{1}{n} < y$. □

1.6 Die komplexen Zahlen

Die Gleichung $x^2 = -1$ besitzt keine Lösung in \mathbb{R} , denn aus den Anordnungsaxiomen folgt für alle $x \in \mathbb{R}$: $x^2 \geq 0 > -1$. Insbesondere ist $x^2 \neq -1$. Angenommen es wäre möglich den Zahlenbereich der reellen Zahlen so zu erweitern, dass auch diese Gleichung eine Lösung besitzt. Bezeichnen wir die Lösung mit i , so ist mit $x, y \in \mathbb{R}$ auch

$$x + iy$$

ein Element dieses größeren Zahlenbereiches. Da dieser Zahlenbereich die Körperaxiome erfüllen soll, erhalten wir für alle $x, y, u, v \in \mathbb{R}$:

$$(x + iy) + (u + iv) = x + u + i(y + v)$$

sowie

$$(x + iy)(u + iv) = xu + i(xv + yu) + i^2yv = xu - yv + i(xv + yu)$$

Um diese Verknüpfungen sich besser veranschaulichen zu können, identifizieren wir die Zahl $x + iy$ mit $(x, y) \in \mathbb{R}^2$. Man nennt $\mathbb{C} := \mathbb{R}^2$ zusammen mit der oben definierten Addition und Multiplikation die *Menge der komplexen Zahlen*. Dass diese Menge ein Körper ist, zeigt der folgende Satz.

Satz 1.6.1. *Die Menge $\mathbb{C} := \mathbb{R}^2$ zusammen mit der wie folgt erklärten Addition (+)*

$$(x, y) + (u, v) = (x + u, y + v)$$

und Multiplikation (\cdot)

$$(x, y) \cdot (u, v) = (xu - yv, xv + yu)$$

bildet einen Körper. Dabei ist $(0, 0)$ das neutrale Element der Addition und $(1, 0)$ das neutrale Element der Multiplikation. Das zu $z = (x, y) \neq 0$ inverse Element der Multiplikation ist durch

$$z^{-1} := \left(\frac{x}{x^2 + y^2}, \frac{-y}{x^2 + y^2} \right)$$

gegeben.

Bemerkungen.

- a) Schränken wir Addition und Multiplikation in \mathbb{C} auf Elemente der Form $(x, 0) \in \mathbb{R} \times \{0\}$ ein, so gilt:

$$(x, 0) + (u, 0) = (x + u, 0)$$

sowie

$$(x, 0) \cdot (u, 0) = (xu, 0)$$

Identifizieren wir also die Elemente $x \in \mathbb{R}$ mit $(x, 0) \in \mathbb{C}$, so entsprechen der Addition und Multiplikation in \mathbb{R} die Addition und Multiplikation in $\mathbb{R} \times \{0\} \subset \mathbb{C}$. Man nennt \mathbb{R} daher auch Unterkörper von \mathbb{C} . Da

$$x(u, v) := (x, 0) \cdot (u, v) = (xu, xv),$$

lässt sich die Multiplikation einer reellen Zahl x mit einer komplexen Zahl (u, v) als komponentenweise Multiplikation (skalare Multiplikation) deuten. Ist $x > 0$, so handelt es sich um eine Streckung des "Vektors" (u, v) mit Streckungsfaktor x .

- b) Die Zahl $i := (0, 1)$ heißt *imaginäre Einheit*. Die Zahl $(1, 0)$ heißt auch *reelle Einheit*. Da $(1, 0)$ auch das neutrale Element der Multiplikation in \mathbb{C} ist, schreibt man auch 1 statt $(1, 0)$.

Mit dieser Vereinbarung und der Identifikation von $x \in \mathbb{R}$ mit $(x, 0)$ gilt:

$$z := (x, y) = (x, 0) \cdot (1, 0) + (y, 0) \cdot (0, 1) = x1 + yi$$

Die reellen Zahlen x und y heißen *Realteil* und *Imaginärteil* von z . Sie werden mit $\operatorname{Re} z$ und $\operatorname{Im} z$ bezeichnet.

- c) Da $i^2 = -1$ folgt aus dem Distributivgesetz:

$$(z - i)(z + i) = z^2 + 1.$$

Damit bestehen die Lösungen der Gleichung $z^2 + 1 = 0$ genau aus der Menge $\{i, -i\}$.

Beweis. Die Gültigkeit der Assoziativgesetze und Kommutativgesetze überprüft man durch direktes Nachrechnen. Sie folgen aus den entsprechenden Rechengesetzen für die reellen Zahlen. Außerdem folgt aus Teil a) der obigen Bemerkung, dass $(1, 0)$ das neutrale Element ist. Die Gleichung

$$z \cdot z^{-1} := (x, y) \cdot \left(\frac{x}{x^2 + y^2}, \frac{-y}{x^2 + y^2} \right) = \left(\frac{x^2}{x^2 + y^2} + \frac{y^2}{x^2 + y^2}, \frac{-xy}{x^2 + y^2} + \frac{yx}{x^2 + y^2} \right) = (1, 0)$$

folgt aus der Definition der Multiplikation. □

Eine wichtige Operation in \mathbb{C} ist die komplexe Konjugation.

Definition 1.6.2. Sei $z = x + iy$ mit $x, y \in \mathbb{R}$ eine Komplexe Zahl. Dann heißt

$$\bar{z} := x - iy$$

die zu z *komplex konjugierte Zahl*.

Bemerkung. \bar{z} lässt sich als Spiegelung von z an der reellen Zahlengeraden $\mathbb{R} \times \{0\}$ auffassen.

Es gelten folgende Rechenregeln.

Satz 1.6.3. Seien $z, w \in \mathbb{C}$, so gilt:

$$(a) \begin{aligned} \overline{z + w} &= \bar{z} + \bar{w} \\ \overline{z \bar{w}} &= \bar{z} \cdot w \end{aligned}$$

$$(b) \begin{aligned} z + \bar{z} &= 2 \operatorname{Re} z \\ z - \bar{z} &= 2i \operatorname{Im} z \end{aligned}$$

$$(c) \bar{z} = z \quad \text{genau dann, falls } z \in \mathbb{R}$$

$$(d) z\bar{z} = x^2 + y^2.$$

Ähnlich wie in den reellen Zahlen lässt sich der Betrag einer komplexen Zahl definieren.

Definition 1.6.4. Der *Betrag* einer komplexen Zahl z ist die nicht negative Zahl

$$|z| = \sqrt{z\bar{z}} = \sqrt{x^2 + y^2}$$

Bemerkung. Ist $z \in \mathbb{R}$, so stimmt der Betrag mit dem im Abschnitt 1.3 eingeführten Betrag für die reellen Zahlen überein. Der Betrag bestimmt den Abstand der komplexen Zahl zur $0 \in \mathbb{C}$.

Es gelten die folgenden Rechenregeln:

Satz 1.6.5. Für alle $z, w \in \mathbb{C}$ gilt:

$$(a) |z| > 0 \quad \text{genau dann, falls } z \neq 0$$

$$(b) |\bar{z}| = |z|$$

$$(c) |\operatorname{Re} z| \leq |z| \quad \text{und} \quad |\operatorname{Im} z| \leq |z|$$

$$(d) |zw| = |z||w|$$

$$(e) |z + w| \leq |z| + |w| \quad (\text{Dreiecksungleichung})$$

Beweis. Die Aussagen (a), (b), (c) folgen unmittelbar aus der Definition. Die Aussage (d) folgt aus $|zw|^2 = zw\bar{z}\bar{w} = z\bar{z}w\bar{w} = |z|^2|w|^2$. Die Aussage (e) folgt, denn $|z + w|^2 = (z + w)(\bar{z} + \bar{w}) = z\bar{z} + z\bar{w} + w\bar{z} + w\bar{w}$. Da $2 \operatorname{Re}(z\bar{w}) = z\bar{w} + \bar{z}w$, erhalten wir

$$\begin{aligned} |z + w|^2 &= z\bar{z} + 2 \operatorname{Re}(z\bar{w}) + w\bar{w} \\ &\leq |z|^2 + 2|z\bar{w}| + |w|^2 = (|z| + |w|)^2 \end{aligned}$$

□

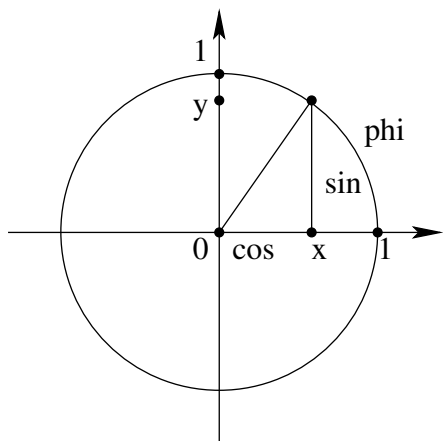


Abbildung 1.2:

Nun wollen wir kurz auf die geometrische Interpretation der Addition und Multiplikation in \mathbb{C} eingehen.

Die Addition in \mathbb{C} entspricht der Addition von Vektoren im \mathbb{R}^2 , denn ist $z = (x, y)$ und $w = (u, v)$, so ist $z + w = (x + u, y + v)$.

Etwas schwieriger ist die geometrische Interpretation der Multiplikation.

Dazu greifen wir auf die in der Schule eingeführte cosinus- und sinus-Funktion zurück. (Später werden wir diese Funktionen mittels der komplexen Exponentialfunktion rigoros definieren).

Ist $(x, y) = x + iy$ und $x^2 + y^2 = 1$, so existiert ein $\varphi \in \{x \in \mathbb{R} \mid 0 \leq x < 2\pi\}$ mit $\cos \varphi = x$ und $\sin \varphi = y$. φ entspricht der Länge des Kreisbogens von $(1, 0)$ bis (x, y) (siehe Abbildung 1.2).

Sind nun z, w komplexe Zahlen, so existieren $\varphi, \psi \in \{x \in \mathbb{R} \mid 0 \leq x < 2\pi\}$ mit

$$z = |z| \frac{z}{|z|} = |z|(\cos \varphi + i \sin \varphi) \quad \text{und} \quad w = |w|(\cos \psi + i \sin \psi).$$

Beachte, dass die Zahlen $\frac{z}{|z|}$ und $\frac{w}{|w|}$ den Betrag 1 haben. Dann gilt:

$$z \cdot w = |z||w|((\cos \varphi \cos \psi - \sin \varphi \sin \psi) + i(\cos \varphi \sin \psi + \sin \varphi \cos \psi))$$

Aus den in der Schule hergeleiteten Additionstheoremen folgt:

$$z \cdot w = |z||w| \cdot (\cos(\varphi + \psi) + i \sin(\varphi + \psi)).$$

Also kann für jedes gegebene $z = |z| \cdot (\cos \varphi + i \sin \varphi)$ die Abbildung

$$w \rightarrow z \cdot w$$

als Drehung des Vektors w um den Winkel φ und anschließender Streckung mit dem Faktor $|z|$ interpretiert werden. Die Additionstheoreme werden wir bald, mit Hilfe der Exponentialfunktion und ihrer Eigenschaften, herleiten.

Kapitel 2

Folgen und Reihen

2.1 Konvergenz von reellen Zahlenfolgen

Definition 2.1.1. Unter einer *Folge reeller Zahlen* $(a_n)_{n \in \mathbb{N}}$ verstehen wir eine Abbildung $\mathbb{N} \rightarrow \mathbb{R}$ mit $n \mapsto a_n$. Die Zahlen a_n heißen die *Glieder der Folge*. Manchmal schreibt man (a_n) oder auch (a_1, a_2, a_3, \dots) .

Beispiele.

(a) $(\frac{1}{n})$; $(1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \dots)$

(b) $((-1)^n)$; $(-1, +1, -1, \dots)$

(c) (n^2) ; $(1, 4, 9, 16, \dots)$

Nun werden wir den fundamentalen Begriff der Konvergenz einer Folge einführen. Anschaulich verstehen wir unter der Konvergenz einer Folge $(a_n)_{n \in \mathbb{N}}$, dass die Folgenglieder a_n sich mit wachsendem n einer reellen Zahl a nähern. Dies bedeutet, dass für jede "kleine Umgebung" von a die Folgenglieder a_n ab einem bestimmten n_0 alle in dieser Umgebung liegen. Was für Umgebungen betrachten wir?

Definition 2.1.2. Sei $\epsilon > 0$ und $a \in \mathbb{R}$. Unter einer ϵ -Umgebung von a verstehen wir die Menge

$$B(a, \epsilon) = \{x \in \mathbb{R} \mid |x - a| < \epsilon\}.$$

Bemerkung. Für gegebene reelle Zahlen $b < c$ heißt eine Menge der Form

$$(b, c) := \{x \in \mathbb{R} \mid b < x < c\}$$

offenes Intervall. Dann gilt:

$$B(a, \epsilon) = \{x \in \mathbb{R} \mid -\epsilon < x - a < \epsilon\} = (-\epsilon + a, +\epsilon + a).$$

Definition 2.1.3. Eine Folge $(a_n)_{n \in \mathbb{N}}$ ist *konvergent* mit *Grenzwert* a genau dann, wenn für alle $\epsilon > 0$ ein $n_0 \in \mathbb{N}$ existiert mit $a_n \in B(a, \epsilon)$ für alle $n \geq n_0$.

Unter Verwendung der Aussagenlogik lässt sich dies auch knapp wie folgt schreiben:

$$\forall \epsilon > 0 \exists n_0 \in \mathbb{N} \forall n \geq n_0 : a_n \in B(a, \epsilon).$$

Bemerkung. Konvergenz einer Folge a_n mit Grenzwert a bedeutet also, dass für jede vorgegebene ϵ -Umgebung von a die Folgenglieder ab einem bestimmten $n_0 = n_0(\epsilon)$ alle in dieser Umgebung liegen. Man kann die Konvergenz einer Folge auch so interpretieren: In jeder ϵ -Umgebung von a liegen alle Folgenglieder *bis auf endlich viele* (man sagt auch: es liegen *fast alle* Folgenglieder in der Umgebung).

Eine konvergente Folge hat natürlich, wie der folgende Satz zeigt, nur einen Grenzwert.

Satz 2.1.4. *Eine konvergente Folge hat genau einen Grenzwert.*

Beweis. Es sei $(a_n)_{n \in \mathbb{N}}$ Folge mit den Grenzwerten $a, b \in \mathbb{R}$. Dann existiert zu jedem $\epsilon > 0$ ein a_n mit $a_n \in B(a, \epsilon) \cap B(b, \epsilon)$. Also gilt:

$$|a - b| = |a - a_n + a_n - b| \leq |a - a_n| + |a_n - b| < 2\epsilon$$

für alle $\epsilon > 0$. Daraus folgt aber $a = b$, denn sonst wähle $\epsilon := \frac{|a-b|}{2}$ und wir erhalten mit $|a - b| < 2\epsilon = |a - b|$ einen Widerspruch. \square

Bemerkung. Falls $(a_n)_{n \in \mathbb{N}}$ konvergiert mit Grenzwert a , so schreiben wir: $\lim_{n \rightarrow \infty} a_n = a$. Folgen, die nicht konvergieren, heißen *divergent*.

Der folgende Satz folgt aus der Tatsache, dass die reellen Zahlen archimedisch angeordnet sind, d.h. die natürlichen Zahlen sind nicht nach oben beschränkt (siehe Satz 1.5.6)

Satz 2.1.5. *Die Folge $a_n = \frac{1}{n}$ konvergiert mit Grenzwert 0.*

Beweis. Zu zeigen: $\forall \epsilon > 0 \exists n_0 \in \mathbb{N} \forall n \geq n_0 : \left| \frac{1}{n} \right| < \epsilon$. Sei $\epsilon > 0$, so existiert wegen Korollar 1.5.7 ein $n_0 \in \mathbb{N}$ mit $\frac{1}{n_0} < \epsilon$. Ist $n \geq n_0$, so ist $\left| \frac{1}{n} \right| = \frac{1}{n} \leq \frac{1}{n_0} < \epsilon$. \square

Bemerkung.

- (a) Eine Folge mit Grenzwert 0 heißt *Nullfolge*.
- (b) Eine Folge $(a_n)_{n \in \mathbb{N}}$ hat genau dann a als Grenzwert, falls die Folge $b_n := a - a_n$ eine Nullfolge ist, denn $a_n \in B(a, \epsilon)$ ist äquivalent zu $a - a_n \in B(0, \epsilon)$.

Beispiele.

- (a) $a_n = \frac{n}{n+1}$, so gilt: $\lim_{n \rightarrow \infty} a_n = 1$. Zum Beweis bemerken wir, dass

$$\left| \frac{n}{n+1} - 1 \right| = \left| \frac{n - (n+1)}{n+1} \right| = \frac{1}{n+1} < \frac{1}{n}.$$

Ist $\epsilon > 0$, so wähle $n_0 \in \mathbb{N}$ mit $\frac{1}{n_0} < \epsilon$. Dann folgt für alle $n \geq n_0$:

$$\frac{1}{n+1} < \frac{1}{n} \leq \frac{1}{n_0} < \epsilon.$$

- (b) q^n ist Nullfolge, falls $|q| < 1$.

Da $|q^n| = \frac{1}{\left(\frac{1}{|q|}\right)^n}$ und $\frac{1}{|q|} = 1 + h$ mit $h > 0$, folgt aus der Bernoullischen Ungleichung

(siehe Satz 1.4.8):

$$|q^n| = \frac{1}{(1+h)^n} \leq \frac{1}{1+nh} < \frac{1}{nh}.$$

Ist $\epsilon > 0$, so wähle $n_0 \in \mathbb{N}$, so dass $\frac{1}{n_0 h} < \epsilon$, d.h. $\frac{1}{h\epsilon} < n_0$ (wieder existiert ein $n_0 \in \mathbb{N}$, da \mathbb{R} archimedisch angeordnet ist). Ist $n \geq n_0$, so folgt: $\frac{1}{nh} \leq \frac{1}{n_0 h} < \epsilon$.

(c) Für alle $k \in \mathbb{N}$ ist $\frac{1}{\sqrt[k]{n}}$ eine Nullfolge.

Sei $\epsilon > 0$, so existiert wegen Korollar 1.5.7 ein $n_0 \in \mathbb{N}$ mit $\frac{1}{n_0} < \epsilon^k$. Ist $n \geq n_0$, so gilt $\frac{1}{n} \leq \frac{1}{n_0} < \epsilon^k$ und da die Wurzelfunktion $\sqrt[k]{\cdot}$ streng monoton wachsend ist (siehe Bemerkung (1) nach Definition 1.3.3) folgt auch $\frac{1}{\sqrt[k]{n}} < \epsilon$ für alle $n \geq n_0$.

(d) Es gilt: $\lim_{n \rightarrow \infty} \sqrt[n]{n} = 1$.

Es sei $h_n := \sqrt[n]{n} - 1 \geq 0$. Aus der binomischen Formel 1.4.12 folgt für $n \geq 2$:

$$n = (h_n + 1)^n \geq 1 + \binom{n}{2} h_n^2 = 1 + \frac{n(n-1)}{2} h_n^2$$

und somit gilt: $h_n^2 \leq \frac{2}{n}$. Da wegen (c) $\frac{1}{\sqrt[n]{n}}$ eine Nullfolge ist, ist wegen $0 \leq h_n \leq \sqrt{\frac{2}{n}}$ auch h_n eine Nullfolge.

(e) Für alle $a \in \mathbb{R}$ ist $\frac{a^n}{n!}$ eine Nullfolge. Für $a = 0$ ist die Behauptung trivial. Sei $a \neq 0$ und $b_n = \frac{|a|^n}{n!}$. Dann gilt

$$\frac{b_{n+1}}{b_n} = \frac{|a|^{n+1} n!}{(n+1)! |a|^n} = \frac{|a|}{n+1}.$$

Daher existiert ein $n_0 \in \mathbb{N}$ mit $\frac{b_{n+1}}{b_n} \leq \frac{1}{2}$ für $n \geq n_0$. Insbesondere ist $b_{n_0+k} \leq \left(\frac{1}{2}\right)^k b_{n_0}$.

Da wegen (b) die Folge $\left(\frac{1}{2}\right)^k$ eine Nullfolge ist, ist auch b_n eine Nullfolge. Dann ist aber auch $\frac{a^n}{n!}$ wegen $\left|\frac{a^n}{n!}\right| \leq b_n$ eine Nullfolge.

Definition 2.1.6. Eine Folge $(a_n)_{n \in \mathbb{N}}$ heißt *beschränkt*, falls die Menge $\{a_n \mid n \in \mathbb{N}\}$ nach oben und unten beschränkt ist.

Bemerkung. $(a_n)_{n \in \mathbb{N}}$ ist genau dann beschränkt, falls ein $c > 0$ existiert mit $|a_n| \leq c$ für alle $n \in \mathbb{N}$.

Satz 2.1.7. *Konvergente Folgen sind beschränkt.*

Beweis. Sei $(a_n)_{n \in \mathbb{N}}$ konvergent mit Grenzwert a . Zu $\epsilon = 1$ existiert ein $n_0 \in \mathbb{N}$ mit $a_n \in B(a, 1)$ für alle $n \geq n_0$. Daraus folgt:

$$|a_n| \leq |a_n - a| + |a| \leq 1 + |a|$$

für $n \geq n_0$. Daher ist die Folge durch $c := \max\{1 + |a|, |a_1|, \dots, |a_{n_0-1}|\}$ beschränkt. \square

Bemerkung. Beschränkte Folgen müssen natürlich nicht konvergent sein, wie z.B. die Folge $a_n = (-1)^n$ zeigt.

Von großer Bedeutung ist der Begriff der Teilfolge. Eine Teilfolge von $(a_n)_{n \in \mathbb{N}}$ ist eine unendliche Folge, die man durch Fortlassen von Folgengliedern unter der Beibehaltung der Reihenfolge erhält. Genauer:

Definition 2.1.8. Sei $(a_n)_{n \in \mathbb{N}}$ eine Folge. Dann heißt die Folge $(b_n)_{n \in \mathbb{N}}$ *Teilfolge* von $(a_n)_{n \in \mathbb{N}}$, falls eine streng monoton steigende Abbildung $k: \mathbb{N} \rightarrow \mathbb{N}$ existiert mit $b_n = a_{k(n)}$.

Bemerkung. $k: \mathbb{N} \rightarrow \mathbb{N}$ monoton steigend bedeutet $k(n) < k(n+1)$ für alle $n \in \mathbb{N}$. Man schreibt auch k_n statt $k(n)$. Der Abbildung entspricht die monoton steigende Folge natürlicher Zahlen $k_1 < k_2 < \dots < k_n < \dots$.

Beispiel. Sei $a_n = \frac{(-1)^n}{n}$ und $k(n) = 2n$. Dann ist $b_n = a_{2n} = \frac{1}{2n}$ eine Teilfolge von (a_n) .

Satz 2.1.9. Sei $(a_n)_{n \in \mathbb{N}}$ eine konvergente Folge mit Grenzwert a . Dann konvergiert jede Teilfolge gegen den Grenzwert a .

Beweis. Sei $(b_n)_{n \in \mathbb{N}}$ Teilfolge von $(a_n)_{n \in \mathbb{N}}$, d.h. $b_n = a_{k(n)}$ für eine monoton steigende Funktion $k: \mathbb{N} \rightarrow \mathbb{N}$. Sei $\epsilon > 0$, so existiert $n_0 \in \mathbb{N}$, so dass $a_n \in B(a, \epsilon)$ für $n \geq n_0$. Da $k(n) \geq n$, folgt: $b_n = a_{k(n)} \in B(a, \epsilon)$ für $n \geq n_0$. \square

2.2 Rechenregeln für konvergente Folgen

Satz 2.2.1. Es seien (a_n) und (b_n) konvergente Folgen mit $\lim_{n \rightarrow \infty} a_n = a$ und $\lim_{n \rightarrow \infty} b_n = b$. Ist $c \in \mathbb{R}$, so sind die Folgen $(a_n + b_n)$ und (ca_n) konvergent mit $\lim_{n \rightarrow \infty} (a_n + b_n) = a + b$ und $\lim_{n \rightarrow \infty} (ca_n) = ca$.

Beweis. Sei $\epsilon > 0$. Da (a_n) und (b_n) konvergent sind, existieren $n_0, n_1 \in \mathbb{N}$ mit $|a_n - a| < \epsilon/2$ für $n \geq n_0$ und $|b_n - b| < \epsilon/2$ für $n \geq n_1$.

Ist $m_0 = \max\{n_0, n_1\}$, so folgt für $n \geq m_0$:

$$|a_n + b_n - (a + b)| = |a_n - a + b_n - b| \leq |a_n - a| + |b_n - b| < \epsilon/2 + \epsilon/2 = \epsilon.$$

Dies beweist die erste Behauptung.

Ist $c = 0$, so ist die zweite Behauptung trivial. Sei $c \neq 0$ und $\epsilon > 0$. Wähle $n_0 \in \mathbb{N}$, so dass $|a_n - a| < \epsilon/|c|$ für alle $n \in n_0$. Dann gilt:

$$|ca_n - ca| = |c| |a_n - a| < |c| \frac{\epsilon}{|c|} = \epsilon.$$

\square

Bevor wir den entsprechenden Satz für das Produkt zweier Folgen zeigen, beweisen wir:

Satz 2.2.2. Es sei (a_n) eine Nullfolge und (b_n) eine beschränkte Folge. Dann ist auch $(a_n \cdot b_n)$ eine Nullfolge.

Beweis. Da (b_n) beschränkt ist, existiert ein $c > 0$ mit $|b_n| \leq c$ für alle $n \in \mathbb{N}$. Sei $\epsilon > 0$, so existiert (zu $\frac{\epsilon}{c} > 0$) ein $n_0 \in \mathbb{N}$, so dass $|a_n| < \frac{\epsilon}{c}$ für alle $n \geq n_0$. Dann folgt für alle $n \geq n_0$:

$$|a_n \cdot b_n| = |a_n| |b_n| < \frac{\epsilon}{c} \cdot c = \epsilon$$

\square

Damit können wir beweisen:

Satz 2.2.3. Es seien (a_n) und (b_n) konvergent mit $\lim_{n \rightarrow \infty} a_n = a$ und $\lim_{n \rightarrow \infty} b_n = b$. Dann ist auch $(a_n \cdot b_n)$ konvergent mit $\lim_{n \rightarrow \infty} (a_n \cdot b_n) = a \cdot b$.

Beweis. Da $a_n - a$ eine Nullfolge ist und b_n als konvergente Folge beschränkt ist, folgt aus Satz 2.2.2, dass $((a_n - a)b_n) = (a_n b_n - ab_n)$ eine Nullfolge ist.

Da $a_n b_n = a_n b_n - ab_n + ab_n$, folgt aus Satz 2.2.1:

$$\lim_{n \rightarrow \infty} (a_n b_n) = \lim_{n \rightarrow \infty} (a_n b_n - ab_n) + \lim_{n \rightarrow \infty} ab_n = 0 + ab.$$

\square

Satz 2.2.4. *Es sei (a_n) eine konvergente Folge mit $a_n \neq 0$ und $\lim_{n \rightarrow \infty} a_n = a \neq 0$. Dann ist auch die Folge $\frac{1}{a_n}$ konvergent mit $\lim_{n \rightarrow \infty} \frac{1}{a_n} = \frac{1}{a}$.*

Beweis. Da

$$\frac{1}{a} - \frac{1}{a_n} = \frac{a_n - a}{a_n \cdot a} = (a_n - a) \frac{1}{a_n \cdot a},$$

genügt es zu zeigen: die Folge $\left(\frac{1}{a_n a}\right)$ ist beschränkt. Denn dann ist wegen Satz 2.2.2 $(a_n - a) \frac{1}{a_n a}$ eine Nullfolge.

Da $a - a_n$ eine Nullfolge ist, gilt für fast alle n (für alle bis auf endlich viele) $|a - a_n| < \frac{|a|}{2}$.

Da wegen der Dreiecksungleichung $|a_n| \geq |a| - |a - a_n|$, folgt $|a_n| \geq \frac{|a|}{2}$ für fast alle $n \in \mathbb{N}$.

Damit gilt: $\frac{1}{|a_n \cdot a|} = \frac{1}{|a_n| |a|} \leq \frac{2}{|a|^2}$ für fast alle $n \in \mathbb{N}$ und somit ist die Folge $\left(\frac{1}{a_n a}\right)$ beschränkt. \square

Bemerkung. Als Anwendung der obigen Rechenregeln erhalten wir für Polynome $p(x) = a_k x^k + \dots + a_0$, dass die Abbildung p konvergente Folgen auf konvergente Folgen abbildet. Genauer gilt: Ist (x_n) eine konvergente Folge mit Grenzwert x , so ist $(p(x_n))$ eine konvergente Folge mit

$$\lim_{n \rightarrow \infty} p(x_n) = \lim_{n \rightarrow \infty} a_k x_n^k + \lim_{n \rightarrow \infty} a_{k-1} x_n^{k-1} + \dots + a_0 = p(x).$$

Dies ist nichts anderes als die Stetigkeit von p . Diesen Begriff werden wir bald einführen.

2.3 Konvergenzkriterien

Definition 2.3.1. Eine Folge $(a_n)_{n \in \mathbb{N}}$ heißt *monoton wachsend* (*monoton fallend*), falls $a_n \leq a_{n+1}$ ($a_n \geq a_{n+1}$) für alle $n \in \mathbb{N}$.

Für monotone Folgen gilt der folgende fundamentale Satz.

Satz 2.3.2. *Sei $(a_n)_{n \in \mathbb{N}}$ eine monoton wachsende, nach oben beschränkte Folge. Dann ist $(a_n)_{n \in \mathbb{N}}$ konvergent mit*

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

Beweis. Da die Menge $A := \{a_n \mid n \in \mathbb{N}\}$ nach oben beschränkt ist, existiert wegen des Vollständigkeitsaxioms das Supremum von A . Sei $a = \sup A$. Wir zeigen, dass a der Grenzwert der Folge (a_n) ist. Sei $\epsilon > 0$, so ist $a - \epsilon$ keine obere Schranke von A , d.h. $\exists n_0 \in \mathbb{N}$ mit $a - \epsilon < a_{n_0} \leq a$. Da (a_n) monoton wachsend ist, folgt:

$$a - \epsilon < a_{n_0} \leq a_n \leq a$$

und somit gilt: $|a - a_n| = a - a_n < a - (a - \epsilon) = \epsilon$ für alle $n \geq n_0$. \square

Bemerkung. Ist (a_n) monoton fallend und nach unten beschränkt, so ist (a_n) ebenfalls konvergent. Dazu betrachte die Folge $b_n = -a_n$. Diese Folge ist monoton wachsend und nach oben beschränkt und somit konvergent. Dann konvergiert auch (a_n) mit Grenzwert $\inf\{a_n \mid n \in \mathbb{N}\}$. Denn mit Satz 1.5.5 folgt:

$$-\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} b_n = \sup\{-a_n \mid n \in \mathbb{N}\} = -\inf\{a_n \mid n \in \mathbb{N}\}.$$

Beispiel. Sei $a > 0$. Im Folgenden konstruieren wir eine monoton fallende Folge, die gegen \sqrt{a} konvergiert. Betrachte die Folge (a_n) mit

$$a_1 > 0 \text{ und } a_{n+1} = \frac{1}{2} \left(a_n + \frac{a}{a_n} \right).$$

Dann gilt:

- (1) Mit Hilfe der vollständigen Induktion zeigt man: $a_n > 0$, denn $a_1 > 0$ und wegen $a > 0$ ist auch $a_{n+1} > 0$, falls $a_n > 0$.
- (2) $a_{n+1}^2 \geq a$ für alle $n \in \mathbb{N}$, denn

$$\begin{aligned} a_{n+1}^2 &= \frac{1}{4} \left(a_n^2 + 2a + \left(\frac{a}{a_n} \right)^2 \right) \\ &= a + \frac{1}{4} \left(a_n^2 - 2a + \left(\frac{a}{a_n} \right)^2 \right) = a + \frac{1}{4} \left(a_n - \frac{a}{a_n} \right)^2. \end{aligned}$$

Insbesondere ist (a_n) für $n \geq 2$ nach unten durch \sqrt{a} beschränkt.

- (3) $a_{n+1} \leq a_n$ für $n \geq 2$:

$$a_n - a_{n+1} = a_n - \frac{1}{2}a_n - \frac{1}{2} \frac{a}{a_n} = \frac{1}{2} \left(a_n - \frac{a}{a_n} \right) = \frac{1}{2a_n} (a_n^2 - a) \geq 0.$$

- (4) Da a_n monoton fallend, existiert

$$\lim_{n \rightarrow \infty} a_n = c \geq \sqrt{a}.$$

Da (a_{n+1}) eine Teilfolge von (a_n) ist, hat sie den gleichen Limes und es folgt aus den Rechenregeln für konvergente Folgen mit (3):

$$0 = \lim_{n \rightarrow \infty} a_n - \lim_{n \rightarrow \infty} a_{n+1} = \lim_{n \rightarrow \infty} \frac{1}{2a_n} (a_n^2 - a) = \frac{c^2 - a}{2c}$$

Also ist $c = \sqrt{a}$ und somit besitzt (a_n) den Grenzwert \sqrt{a} . Insbesondere zeigt dies nochmals die Lösbarkeit der Gleichung $x^2 = a$. Im Gegensatz zu dem Beweis von Satz 1.5.3 ist dieser Beweis konstruktiv, d.h. der Beweis liefert ein Verfahren zur Berechnung von \sqrt{a} . Außerdem sind explizite Fehlerabschätzungen möglich. (Wir bemerken noch, dass die Konstruktion dieser Folge auf dem Newton-Verfahren beruht.)

Nun wollen wir uns mit dem Begriff der Intervallschachtelung beschäftigen. Ist $a \leq b$, so nennen wir $[a, b] = \{x \in \mathbb{R} \mid a \leq x \leq b\}$ ein *abgeschlossenes Intervall*.

Definition 2.3.3. Eine Mengenfamilie $(I_n)_{n \in \mathbb{N}}$ von abgeschlossenen Intervallen $I_n = [a_n, b_n]$ heißt *Intervallschachtelung*, falls gilt:

- (a) $I_{n+1} \subset I_n$ für alle $n \in \mathbb{N}$,
- (b) $b_n - a_n$ ist Nullfolge.

Satz 2.3.4. Sei (I_n) eine Intervallschachtelung. Dann existiert ein $c \in \mathbb{R}$ mit $\bigcap_{n \geq 1} I_n = \{c\}$.

Beweis.

(1) $\bigcap_{n \geq 1} I_n$ besteht aus höchstens einem Element, denn sind $c_1, c_2 \in \bigcap_{n \geq 1} I_n$ und $c_1 \leq c_2$, so folgt für alle $n \in \mathbb{N}$: $c_1, c_2 \in I_n \Rightarrow a_n \leq c_1 \leq c_2 < b_n \Rightarrow 0 \leq c_2 - c_1 \leq b_n - a_n$. Da $b_n - a_n$ eine Nullfolge ist, folgt: $c_2 = c_1$.

(2) Da (a_n) monoton steigend und nach oben beschränkt ist und b_n monoton fallend und nach unten beschränkt ist, folgt die Konvergenz von $(a_n), (b_n)$. Außerdem ist wegen Satz 2.2.1

$$\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} (b_n - a_n) + \lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} b_n =: c$$

denn $b_n - a_n$ ist eine Nullfolge. Des weiteren ist $a_n \leq c \leq b_n$ für alle $n \in \mathbb{N}$, denn $c = \sup\{a_n \mid n \in \mathbb{N}\} = \inf\{b_n \mid n \in \mathbb{N}\}$. \square

Wir haben gesehen, dass jede konvergente Folge beschränkt ist. Die Umkehrung gilt nicht (z.B. $a_n = (-1)^n$). Aber es gilt:

Satz 2.3.5. (Bolzano-Weierstraß)

Jede beschränkte Folge (a_n) besitzt eine konvergente Teilfolge.

Beweis. Da (a_n) beschränkt ist, existieren $x_1 \leq y_1$ mit $x_1 \leq a_n \leq y_1$ für alle $n \in \mathbb{N}$. Wir werden induktiv eine Intervallschachtelung $(I_n)_{n \in \mathbb{N}}$ definieren mit $I_n = [x_n, y_n]$, die unendlich viele Glieder der Folge (a_n) enthält.

Induktionsanfang: $I_1 = [x_1, y_1]$.

Es seien I_1, \dots, I_n schon definiert. $I_n = [x_n, y_n] = [x_n, M] \cup [M, y_n]$, wobei $M = \frac{x_n + y_n}{2}$ der Mittelpunkt von I_n ist. Definiere

$$[x_{n+1}, y_{n+1}] := I_{n+1} := \begin{cases} [x_n, M] & \text{falls in diesem Intervall unendlich} \\ & \text{viele Folgenglieder liegen} \\ [M, y_n] & \text{sonst} \end{cases}$$

Daher gilt: $I_{n+1} \subset I_n$ und $y_{n+1} - x_{n+1} = \frac{1}{2}(y_n - x_n)$.

Mittels Induktion zeigt man: $(y_n - x_n) = \left(\frac{1}{2}\right)^{n-1} (y_1 - x_1)$ und da $\left(\frac{1}{2}\right)^{n-1}$ eine Nullfolge ist, ist auch $y_n - x_n$ Nullfolge. Das heißt: I_n ist eine Intervallschachtelung und es gibt ein $c \in \mathbb{R}$ mit $c \in \bigcap_{n \in \mathbb{N}} I_n$.

Man definiert nun induktiv eine Teilfolge $a_{k_n} \in I_n$ von (a_n) .

Sei $k_1 = 1$. Seien $a_{k_1}, a_{k_2}, \dots, a_{k_n}$ schon definiert mit $k_1 < k_2 < \dots < k_n$. Dann existiert ein $m \in \mathbb{N}$ mit $m > k_n$ und $a_m \in I_{n+1}$ (denn I_{n+1} enthält unendlich viele Folgenglieder).

Setze $k_{n+1} = m$. Es gilt also: $x_n \leq a_{k_n} \leq y_n$ für alle $n \in \mathbb{N}$. Da $\lim_{n \rightarrow \infty} x_n = \lim_{n \rightarrow \infty} y_n = c$, folgt

$\lim_{n \rightarrow \infty} a_{k_n} = c$, denn wegen $a_{k_n}, c \in [x_n, y_n]$ für alle $n \in \mathbb{N}$ folgt

$$|a_{k_n} - c| \leq |y_n - x_n|.$$

\square

Definition 2.3.6. Eine Folge (a_n) heißt *Cauchyfolge*, wenn folgendes gilt:

$$\forall \epsilon > 0 \exists n_0 \in \mathbb{N} \forall n, m \geq n_0 : |a_n - a_m| < \epsilon.$$

Bemerkung.

- (a) In dieser Definition taucht nirgendwo der Begriff des Grenzwertes auf.
- (b) Es ist leicht zu sehen (s.u.), dass jede konvergente Folge auch eine Cauchyfolge ist. Der nächste Satz zeigt, dass auch die Umkehrung gilt.

Satz 2.3.7. *Eine Folge ist genau dann konvergent, wenn sie eine Cauchyfolge ist.*

Beweis.

- (1) Die offensichtliche Richtung ist: Jede konvergente Folge ist Cauchyfolge.
Sei (a_n) konvergent mit Grenzwert a und $\epsilon > 0$. Dann existiert ein $n_0 \in \mathbb{N}$ mit $|a_n - a| < \epsilon/2$ für alle $n \geq n_0$. Aus der Dreiecksungleichung folgt: $|a_n - a_m| \leq |a_n - a| + |a - a_m| < \epsilon$ für alle $n, m \geq n_0$.
- (2) Jede Cauchyfolge ist konvergent.
Wir zeigen zunächst: Jede Cauchyfolge ist beschränkt. Sei (a_n) Cauchyfolge und $\epsilon = 1$. Dann existiert ein $n_0 \in \mathbb{N}$ mit $|a_n - a_{n_0}| < 1$ für alle $n \geq n_0$, und somit gilt:

$$|a_n| \leq |a_n - a_{n_0}| + |a_{n_0}| \leq 1 + |a_{n_0}|.$$

Daraus folgt $|a_n| \leq \max\{|a_1|, \dots, |a_{n_0-1}|, 1 + |a_{n_0}|\}$ für alle $n \in \mathbb{N}$.

Nun zeigen wir die Konvergenz von (a_n) . Da (a_n) beschränkt ist, besitzt (a_n) wegen des Satzes 2.3.5 von Bolzano-Weierstraß eine konvergente Teilfolge (a_{k_n}) mit Grenzwert a . Sei $\epsilon > 0$. Dann existiert ein $n_1 \in \mathbb{N}$ mit $|a_{k_n} - a| < \epsilon/2$ für $n \geq n_1$ und es existiert ein $n_2 \in \mathbb{N}$ mit $|a_n - a_m| < \epsilon/2$ für $n, m \geq n_2$. Dann gilt für $n_0 = \max\{n_1, n_2\}$:

$$|a_n - a| \leq |a_n - a_{k_n}| + |a_{k_n} - a| < \epsilon$$

für $n \geq n_0$, denn $k_n \geq n \geq n_0$. □

Bemerkungen zur Axiomatik von \mathbb{R}

Die Existenz einer kleinsten oberen Schranke ist äquivalent dazu, dass \mathbb{R} archimedisch angeordnet ist und jede Cauchyfolge einen Grenzwert besitzt. Der Begriff der Vollständigkeit läßt sich auch auf "abstrakte" Räume (sogenannte metrische Räume) ausdehnen. Diese haben im allgemeinen keine totale Ordnung, aber der Begriff der Cauchyfolge macht immer noch Sinn. Vollständige Räume sind dann solche, für die jede Cauchyfolge einen Grenzwert besitzt.

Für den Grenzwert a einer Folge (a_n) verlangt man, dass in jeder ϵ -Umgebung von a alle Folgenglieder liegen, bis auf endlich viele. Unter dem Häufungspunkt einer Folge (a_n) versteht man eine Zahl a , so dass in jeder ϵ -Umgebung unendlich viele Folgenglieder liegen.

Definition 2.3.8. a heißt *Häufungspunkt* (HP) von (a_n) , falls für alle $\epsilon > 0$ die Menge $\{n \in \mathbb{N} \mid a_n \in B(a, \epsilon)\}$ unendlich ist.

Bemerkung. Es gilt:

a ist Häufungspunkt von $(a_n) \Leftrightarrow \forall \epsilon > 0 \quad \forall N \in \mathbb{N} \quad \exists n \geq N : a_n \in B(a, \epsilon)$.

“ \Rightarrow ” Denn ist a ein HP von (a_n) , $\epsilon > 0$ und $N \in \mathbb{N}$, so ist $\{n \in \mathbb{N} \mid n \geq N, a_n \in B(a, \epsilon)\}$ unendlich und insbesondere nicht leer.

“ \Leftarrow ” Wäre die Menge $\{n \in \mathbb{N} \mid a_n \in B(a, \epsilon)\}$ endlich, so gäbe es ein $N \in \mathbb{N}$ mit $a_n \notin B(a, \epsilon)$ für alle $n \geq N$ im Widerspruch zur rechten Seite der Äquivalenz.

Der Begriff des Häufungspunktes einer Folge hängt sehr eng mit konvergenten Teilfolgen zusammen.

Satz 2.3.9. Sei (a_n) eine Folge, so gilt:

a ist HP von $(a_n) \Leftrightarrow$ es gibt eine konvergente Teilfolge a_{k_n} mit $\lim_{n \rightarrow \infty} a_{k_n} = a$.

Bemerkung. Mit Bolzano-Weierstraß folgt:

Jede beschränkte Folge hat mindestens einen HP, denn jede beschränkte Folge hat eine konvergente Teilfolge.

Beweis. “ \Rightarrow ” Da für jedes $\epsilon > 0$ die Menge $\{n \in \mathbb{N} \mid a_n \in B(a, \epsilon)\}$ unendlich ist, existieren $k_1 < k_2 < \dots < k_n < \dots$ mit $|a_{k_n} - a| < \frac{1}{n}$. Denn sind $k_1 < \dots < k_n$ schon definiert, so betrachte die nicht leere Menge

$$\{m \in \mathbb{N} \mid m > k_n \text{ und } |a_m - a| < \frac{1}{n+1}\}.$$

Wähle ein Element aus dieser Menge und bezeichne es mit k_{n+1} . Insbesondere folgt:

$$k_{n+1} > k_n \quad \text{sowie} \quad |a_{k_{n+1}} - a| < \frac{1}{n+1}.$$

Dies zeigt, dass $a_{k_n} - a$ eine Nullfolge ist.

“ \Leftarrow ” Sei a_{k_n} konvergente Teilfolge mit $\lim_{n \rightarrow \infty} a_{k_n} = a$. Dann sind für jedes $\epsilon > 0$ fast alle a_{k_n} in der ϵ -Umgebung von a . Insbesondere ist $\{n \in \mathbb{N} \mid a_n \in B(a, \epsilon)\}$ unendlich für jedes $\epsilon > 0$. \square

Von besonderem Interesse sind die größten und kleinsten Häufungspunkte einer Folge (a_n) .

Definition 2.3.10. Sei (a_n) eine beschränkte Folge. Dann heißt

$$\limsup_{n \rightarrow \infty} a_n := \sup\{a \mid a \text{ HP von } (a_n)\}$$

der *Limes superior* und

$$\liminf_{n \rightarrow \infty} a_n := \inf\{a \mid a \text{ HP von } (a_n)\}$$

der *Limes inferior* der Folge (a_n) .

Bemerkung.

- (a) Wegen des Satzes von Bolzano-Weierstraß besitzt jede beschränkte Folge wenigstens eine konvergente Teilfolge. Ihr Limes ist wegen Satz 2.3.9 ein HP von (a_n) . Damit ist die Menge $\{a \mid a \text{ HP von } (a_n)\}$ nicht leer. Sie ist natürlich beschränkt, da (a_n) beschränkt ist.

(b) Wir schreiben auch häufig: $\overline{\lim}_{n \rightarrow \infty} a_n$ statt $\limsup_{n \rightarrow \infty} a_n$ bzw. $\underline{\lim}_{n \rightarrow \infty} a_n$ statt $\liminf_{n \rightarrow \infty} a_n$. Offensichtlich gilt:

$$\underline{\lim}_{n \rightarrow \infty} a_n \leq \overline{\lim}_{n \rightarrow \infty} a_n.$$

Der Limes superior (inferior) ist selbst wieder HP der Folge, d.h. das Supremum ist ein Maximum und das Infimum ist ein Minimum.

Satz 2.3.11. *Sei (a_n) eine beschränkte Folge. Dann gilt:*

$$\begin{aligned} \sup\{a \mid a \text{ HP von } (a_n)\} &= \max\{a \mid a \text{ HP von } (a_n)\} \\ \inf\{a \mid a \text{ HP von } (a_n)\} &= \min\{a \mid a \text{ HP von } (a_n)\}. \end{aligned}$$

Beweis. Wir zeigen nur den ersten Teil.

Sei $s = \overline{\lim}_{n \rightarrow \infty} a_n := \sup\{a \mid a \text{ HP von } (a_n)\}$ und $\epsilon > 0$. Wir zeigen, dass die Menge

$$M := \{n \in \mathbb{N} \mid |s - a_n| < \epsilon\}$$

unendlich ist. Aus der Definition des Supremums (siehe Definition 1.5.2) folgt die Existenz eines HP a von (a_n) mit $s - \epsilon/2 < a \leq s$. Insbesondere ist die Menge $\{n \in \mathbb{N} \mid |a - a_n| < \epsilon/2\}$ unendlich. Wegen $|a_n - s| \leq |a_n - a| + |a - s| \leq \epsilon/2 + \epsilon/2 = \epsilon$ ist auch die Menge M unendlich. \square

Sei nun (a_n) eine beschränkte Folge, so liegen fast alle Folgenglieder in jeder "Umgebung" des Intervalls $[\underline{\lim}_{n \rightarrow \infty} a_n, \overline{\lim}_{n \rightarrow \infty} a_n]$.

Genauer gilt:

Satz 2.3.12. *Sei (a_n) eine beschränkte Folge, so folgt für alle $\epsilon > 0$:*

$$-\epsilon + \underline{\lim}_{n \rightarrow \infty} a_n < a_n < \overline{\lim}_{n \rightarrow \infty} a_n + \epsilon$$

für fast alle $n \in \mathbb{N}$.

Beweis. Sei $\epsilon > 0$. Betrachte die Menge

$$M := \{n \in \mathbb{N} \mid a_n \geq \overline{\lim}_{n \rightarrow \infty} a_n + \epsilon\}.$$

Wäre diese Menge unendlich, so würde eine Teilfolge (a_{k_n}) von (a_n) existieren mit $a_{k_n} \geq \overline{\lim}_{n \rightarrow \infty} a_n + \epsilon$. Da diese Teilfolge beschränkt ist, hat sie einen HP $a \geq \overline{\lim}_{n \rightarrow \infty} a_n + \epsilon$. Dieser HP ist dann aber auch HP von $(a_n)_{n \in \mathbb{N}}$. Dies steht im Widerspruch zur Definition von $\overline{\lim}_{n \rightarrow \infty} a_n$ als größten HP der Folge (a_n) . Also ist die Menge M endlich. Genauso folgt, dass die Menge $\{n \in \mathbb{N} \mid a_n \leq \underline{\lim}_{n \rightarrow \infty} a_n - \epsilon\}$ endlich ist. \square

Korollar 2.3.13. *Sei (a_n) eine Folge, so gilt:*

(a_n) ist konvergent mit Grenzwert $a \Leftrightarrow (a_n)$ ist beschränkt und $\underline{\lim}_{n \rightarrow \infty} a_n = \overline{\lim}_{n \rightarrow \infty} a_n = a$.

Beweis. " \Rightarrow ": Sei (a_n) konvergent mit Grenzwert a . Aus Satz 2.1.7 folgt, dass a_n beschränkt ist. Die Konvergenz der Folge (a_n) impliziert, dass sie neben a keinen weiteren HP besitzt. Denn zu jedem weiteren HP a' existiert wegen Satz 2.3.9 eine konvergente Teilfolge mit Grenzwert a' und wegen Satz 2.1.9 ist $a' = a$.

" \Leftarrow ": Ist (a_n) beschränkt und gilt $\underline{\lim}_{n \rightarrow \infty} a_n = \overline{\lim}_{n \rightarrow \infty} a_n = a$, so sind wegen Satz 2.3.12 für jedes $\epsilon > 0$ fast alle Folgenglieder in der ϵ -Umgebung von a . \square

Bevor wir uns mit komplexen Folgen beschäftigen, wollen wir kurz auf den Begriff der Abzählbarkeit einer Menge eingehen, der eng mit dem Begriff der Folge verbunden ist.

Definition 2.3.14. Eine nichtleere Menge M heißt *abzählbar*, falls eine surjektive Abbildung $\mathbb{N} \rightarrow M$ mit $n \rightarrow a_n$ existiert. Die leere Menge bezeichnen wir auch als abzählbar. Die nicht abzählbaren Mengen heißen auch *überabzählbar*.

Bemerkung. Eine Menge M ist also abzählbar, wenn eine Folge $(a_n)_{n \in \mathbb{N}}$ existiert mit

$$M = \{a_n \mid n \in \mathbb{N}\}.$$

Jede endliche Menge $M = \{b_1, \dots, b_n\}$ ist offensichtlich abzählbar. Nicht endliche abzählbare Mengen heißen auch *abzählbar unendlich*.

Beispiele.

(a) \mathbb{N} ist abzählbar, denn die Abbildung $\mathbb{N} \rightarrow \mathbb{N}$ mit $n \rightarrow a_n := n$ ist bijektiv. Jede Teilmenge von $M \subset \mathbb{N}$ ist wieder abzählbar. Denn ist M unendlich (dies ist der einzige Fall den wir betrachten müssen) so existiert eine Teilfolge $a_{k_n} = k_n$ von (a_n) mit $M = \{k_n \mid k \in \mathbb{N}\}$. Allgemein ist jede Teilmenge einer abzählbaren Menge wieder abzählbar.

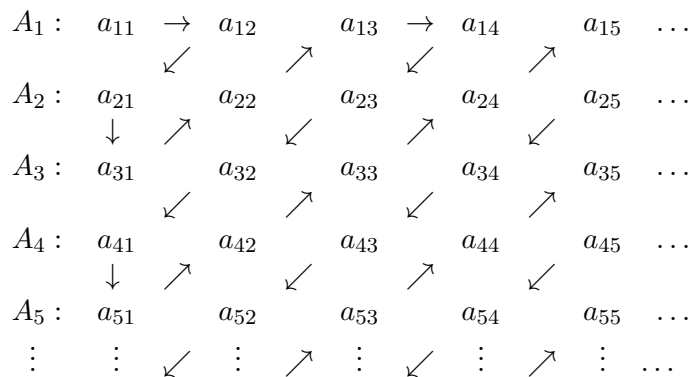
(b) Die Menge der ganzen Zahlen ist abzählbar. Betrachte die Folge $(a_n)_{n \in \mathbb{N}}$ mit

$$a_{2n-1} = -(n-1) \quad \text{und} \quad a_{2n} = n.$$

Diese Folge durchläuft alle ganzen Zahlen.

Satz 2.3.15. Die Vereinigung von abzählbar vielen abzählbaren Mengen A_n mit $n \in \mathbb{N}$ ist wieder abzählbar.

Beweis. Sei $A_n = \{a_{n1}, a_{n2}, a_{n3}, a_{n4}, \dots\}$ so definiert das folgende Diagramm eine Abzählung.



□

Korollar 2.3.16. Die Menge \mathbb{Q} der rationalen Zahlen ist abzählbar.

Beweis. Da die Menge \mathbb{Z} der ganzen Zahlen abzählbar ist, ist auch für jede natürliche Zahl $n \in \mathbb{N}$ die Menge

$$A_n = \left\{ \frac{k}{n} \mid k \in \mathbb{Z} \right\}$$

abzählbar. Also ist wegen Satz 2.3.15 auch

$$\mathbb{Q} = \bigcup_{n \in \mathbb{N}} A_n$$

wieder abzählbar.

□

Bemerkung. Es existiert also eine Folge, die alle rationalen Zahlen durchläuft. Jede reelle Zahl ist wegen Satz 1.5.11 ein Häufungspunkt einer solchen Folge.

Nun stellt sich die Frage, ob eine Folge existiert die alle reellen Zahlen durchläuft. Dies ist nicht der Fall, denn es gilt

Satz 2.3.17. *Die Menge der reellen Zahlen ist überabzählbar.*

Beweis. Wir zeigen sogar: Das Intervall $I_0 := [0, 1]$ ist nicht abzählbar.

Sei also $(a_n)_{n \in \mathbb{N}}$ eine beliebige Folge von Zahlen $a_n \in I_0$. Wir werden mit Hilfe einer Intervallschachtelung eine Zahl $a \in I_0$ konstruieren mit $a_n \neq a$ für alle $n \in \mathbb{N}$.

Im ersten Schritt zerlegen wir das Intervall $I_0 := [0, 1]$ in drei gleich große Intervalle $A_1 = [0, 1/3]$, $A_2 = [1/3, 2/3]$ und $A_3 = [2/3, 1]$. Wähle ein Intervall A_i mit $a_1 \notin A_i$ und setze $[x_1, y_1] := I_1 := A_i$. Die Länge $|I_1| := y_1 - x_1$ ist 3^{-1} .

Im zweiten Schritt zerlegen wir nun das Intervall I_1 wieder in drei gleich große Intervalle und wählen aus diesen Intervallen ein Intervall $I_2 = [x_2, y_2]$ mit $a_2 \notin I_2$. Außerdem ist die Länge $|I_2| := y_2 - x_2 = 3^{-2}$.

Fahren wir so fort, so erhalten wir eine Intervallschachtelung I_n mit

$$a_j \notin \bigcap_{n \in \mathbb{N}} I_n,$$

für alle $j \in \mathbb{N}$, denn nach Konstruktion der Intervalle ist $a_j \notin I_j$. Andererseits enthält jede Intervallschachtelung eine Zahl $a \in \bigcap_{n \in \mathbb{N}} I_n$. Diese Zahl ist also nicht in der Folge enthalten. \square

2.4 Konvergenz von komplexen Zahlenfolgen

Definition 2.4.1. Unter einer *Folge komplexer Zahlen* $(c_n)_{n \in \mathbb{N}}$ verstehen wir eine Abbildung $\mathbb{N} \rightarrow \mathbb{C}$ mit $n \mapsto c_n$. Die Zahlen c_n heißen die *Glieder der Folge*.

Bemerkung. Jede komplexe Zahlenfolge $(c_n)_{n \in \mathbb{N}}$ ist von der Form $c_n = a_n + ib_n$. Dabei sind $\operatorname{Re}(c_n) = a_n$ und $\operatorname{Im}(c_n) = b_n$ reelle Zahlenfolgen.

Beispiele.

(a) $c_n = \frac{1}{n} + in^2$

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

Ganz analog zu den reellen Zahlenfolgen lässt sich die Konvergenz der komplexen Zahlenfolgen definieren. Dabei muß nur der Betrag reeller Zahlen durch den Betrag komplexer Zahlen ersetzt werden.

Definition 2.4.2. Sei $\epsilon > 0$ und $c \in \mathbb{C}$. Unter einer ϵ -*Umgebung* von c verstehen wir die Menge

$$B(c, \epsilon) = \{z \in \mathbb{C} \mid |z - c| < \epsilon\}.$$

Bemerkung. Ist $c = a + ib$, so ist

$$B(c, \epsilon) = \{z = x + iy \in \mathbb{C} \mid \sqrt{(x - a)^2 + (y - b)^2} < \epsilon\}$$

die Menge aller Zahlen in einer offenen Kreisscheibe von Radius ϵ um c .

Definition 2.4.3. Eine Folge $(c_n)_{n \in \mathbb{N}}$ ist *konvergent* mit *Grenzwert* $c \Leftrightarrow$

$$\forall \epsilon > 0 \exists n_0 \in \mathbb{N} \forall n \geq n_0 : c_n \in B(c, \epsilon).$$

Bemerkung. Falls $(c_n)_{n \in \mathbb{N}}$ konvergiert mit Grenzwert c , so schreiben wir: $\lim_{n \rightarrow \infty} c_n = c$.
Folgen, die nicht konvergieren, heißen *divergent*.

Die Konvergenz von komplexer Zahlenfolgen lässt sich auf die Konvergenz reeller Zahlenfolgen zurückführen. Es gilt

Satz 2.4.4. Sei $(c_n)_{n \in \mathbb{N}}$ eine Folge komplexer Zahlen. Die Folge ist genau dann konvergent, wenn die reellen Zahlenfolgen $\operatorname{Re}(c_n) = a_n$ und $\operatorname{Im}(c_n) = b_n$ konvergieren. Im Falle der Konvergenz gilt:

$$\lim_{n \rightarrow \infty} c_n = \lim_{n \rightarrow \infty} a_n + i \lim_{n \rightarrow \infty} b_n$$

Beweis. Nehmen wir zunächst an, dass die Folge $c_n = a_n + ib_n$ gegen $c = a + ib$ konvergiert. Da $|c_n - c| = \sqrt{(a_n - a)^2 + (b_n - b)^2}$, gilt $|a_n - a| \leq |c_n - c|$ sowie $|b_n - b| \leq |c_n - c|$. Damit folgt auch die Konvergenz der Folgen a_n und b_n gegen a bzw. b .

Nehmen wir nun die Konvergenz der Folgen a_n und b_n gegen a bzw. b an. Es gilt:

$$|c_n - c| = |(a_n - a) + i(b_n - b)| \leq |(a_n - a)| + |(b_n - b)|.$$

Da $|(a_n - a)|$ und $|(b_n - b)|$ Nullfolgen sind, ist auch $|c_n - c|$ eine Nullfolge und somit konvergiert c_n gegen c . \square

Bemerkungen.

(a) Analog zu den Rechenregeln für reelle Zahlenfolgen ergeben sich die Rechenregeln für komplexe Zahlenfolgen:

Es seien (c_n) und (d_n) konvergent mit $\lim_{n \rightarrow \infty} c_n = c$ und $\lim_{n \rightarrow \infty} d_n = d$. Dann sind auch die Folgen $(c_n + d_n)$ sowie $(c_n \cdot d_n)$ konvergent und es gilt:

$$\lim_{n \rightarrow \infty} c_n + d_n = c + d,$$

sowie

$$\lim_{n \rightarrow \infty} c_n \cdot d_n = c \cdot d.$$

Ist darüberhinaus $d_n \neq 0$ und $\lim_{n \rightarrow \infty} d_n = d \neq 0$, so ist auch $\frac{c_n}{d_n}$ konvergent und es gilt:

$$\lim_{n \rightarrow \infty} \frac{c_n}{d_n} = \frac{c}{d}.$$

(b) Ist (c_n) eine konvergente Folge, so ist auch $(\overline{c_n})$ konvergent und es gilt:

$$\overline{\lim_{n \rightarrow \infty} c_n} = \lim_{n \rightarrow \infty} \overline{c_n}$$

Analog zu den reellen Zahlenfolgen definieren wir

Definition 2.4.5. Eine Folge $(c_n)_{n \in \mathbb{N}}$ heißt *beschränkt*, falls ein $r > 0$ existiert mit $|c_n| \leq r$ für alle $n \in \mathbb{N}$. Ist $k : \mathbb{N} \rightarrow \mathbb{N}$ eine streng monoton steigende Abbildung so heißt die Folge $d_n = c_{k_n}$ Teilfolge von (c_n) .

Genauso wie bei den reellen Zahlenfolgen ergibt sich:

Satz 2.4.6. (*Bolzano-Weierstraß für komplexe Folgen*)
Jede beschränkte Folge besitzt eine konvergente Teilfolge.

Beweis. Sei also $c_n = a_n + ib_n$ eine beschränkte Folge, so sind die reellen Zahlenfolgen (a_n) und (b_n) ebenfalls beschränkt. Wegen des Satzes 2.3.5 von Bolzano-Weierstraß können wir aus der Folge (a_n) eine konvergente Teilfolge (a_{k_n}) auswählen. Betrachte die Teilfolge $(c'_n) := (a'_n + ib'_n)$ von (c_n) mit $a'_n = a_{k_n}$ und $b'_n = b_{k_n}$. Nun wähle wieder mit dem Satz von Bolzano-Weierstraß eine konvergente Teilfolge (b'_{l_n}) von b'_n aus. Dann ist (c'_{l_n}) eine konvergente Teilfolge von c'_n und somit auch von c_n . \square

Die Definition 2.3.8 des Häufungspunktes für reelle Zahlenfolgen überträgt sich natürlich wörtlich auf komplexe Zahlenfolgen.

Definition 2.4.7. c heißt *Häufungspunkt* (HP) einer komplexen Zahlenfolge (c_n) , falls für alle $\epsilon > 0$ die Menge $\{n \in \mathbb{N} \mid c_n \in B(c, \epsilon)\}$ unendlich ist.

Bemerkung. Analog zu Satz 2.3.9 gilt: c ist genau dann ein Häufungspunkt (HP) von (c_n) , falls eine konvergente Teilfolge mit Grenzwert c existiert. Also hat jede beschränkte Folge einen Häufungspunkt.

Auch der Begriff der Cauchyfolge in Definition 2.3.6 überträgt sich wörtlich auf komplexe Zahlenfolgen, und analog zu Satz 2.3.7 erhalten wir:

Satz 2.4.8. *Eine komplexe Zahlenfolge ist genau dann konvergent, wenn sie eine Cauchyfolge ist.*

Beweis. Ist (c_n) eine Cauchyfolge, so sind ihre Realteile (a_n) und Imaginärteile (b_n) wegen $|a_n - a_m| \leq |c_n - c_m|$ sowie $|b_n - b_m| \leq |c_n - c_m|$ reelle Cauchyfolgen. Diese konvergieren wegen Satz 2.3.7 als reelle Zahlenfolgen. Damit ist wegen Satz 2.4.4 auch die komplexe Zahlenfolge konvergent. Wie schon im Beweis von Satz 2.3.7 bemerkt, ist die Umkehrung der Aussage eine direkte Konsequenz der Dreiecksungleichung.

2.5 Konvergenz von Reihen

Der Begriff der Konvergenz einer unendlichen Reihe wird auf den Begriff der konvergenten Zahlenfolgen zurückgeführt. Genauer definieren wir:

Definition 2.5.1. Sei $(c_k)_{k \in \mathbb{N}}$ eine komplexe Zahlenfolge, so heißt die Folge $(s_n)_{n \in \mathbb{N}}$ mit

$$s_n = \sum_{k=1}^n c_k = c_1 + \dots + c_n$$

die *Folge der Partialsummen* der Reihe $\sum_{k=1}^{\infty} c_k$. Die unendliche Reihe $\sum_{k=1}^{\infty} c_k$ heißt *konvergent* genau dann, falls die Folge der Partialsummen konvergiert. Wir schreiben dann:

$$\sum_{k=1}^{\infty} c_k := \lim_{n \rightarrow \infty} s_n.$$

Die Reihe $\sum_{k=1}^{\infty} c_k$ heißt *divergent*, falls sie nicht konvergent ist.

Die Zahlen c_k heißen die *Glieder der Reihe*, der Index k heißt *Summationsindex*. Wir nennen eine Reihe *reell*, wenn alle Glieder c_k reelle Zahlen sind.

Bemerkungen.

- (a) Eine Reihe $\sum_{k=1}^{\infty} c_k$ ist also nichts anderes als eine Folge von Partialsummen

$$s_n = \sum_{k=1}^n c_k = c_1 + \dots + c_n.$$

Ist diese Folge konvergent, so bezeichnen wir ihren Grenzwert ebenfalls mit dem Symbol $\sum_{k=1}^{\infty} c_k$.

- (b) Es ist oftmals zweckmäßig den Summationsindex mit einer von 1 verschiedenen ganzen Zahl $m \in \mathbb{Z}$ beginnen zu lassen. Ist zum Beispiel die Reihe $\sum_{k=1}^{\infty} c_k$ konvergent, so ist für jedes $m \in \mathbb{N}$ auch die Reihe $\sum_{k=m}^{\infty} c_k$ konvergent. Denn ist (s_n) die Partialsummenfolge der Reihe $\sum_{k=1}^{\infty} c_k$ und (t_n) die Partialsummenfolge für $\sum_{k=m}^{\infty} c_k$, so gilt für $n \geq m$

$$t_n := c_m + \dots + c_n = s_n - \sum_{k=1}^{m-1} c_k$$

Also ist auch die Folge (t_n) konvergent und wir erhalten:

$$\sum_{k=m}^{\infty} c_k := \lim_{n \rightarrow \infty} t_n = \lim_{n \rightarrow \infty} s_n - \sum_{k=1}^{m-1} c_k = \sum_{k=1}^{\infty} c_k - \sum_{k=1}^{m-1} c_k.$$

Wegen

$$\sum_{k=m+1}^{\infty} c_k = \sum_{k=1}^{\infty} c_k - \sum_{k=1}^m c_k$$

heißt $r_m := \sum_{k=m+1}^{\infty} c_k$ auch das *m-te Restglied der Reihe* $\sum_{k=1}^{\infty} c_k$. Wegen der Konvergenz der Reihe $\sum_{k=1}^{\infty} c_k$ bilden ihre Restglieder r_m insbesondere eine Nullfolge.

Ein wichtiges Beispiel für eine konvergente Reihe ist die geometrische Reihe.

Beispiel. Geometrische Reihe:

Sei $z \in \mathbb{C}$, $z \neq 1$, so gilt:

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

denn $(1 + z + \dots + z^n)(1 - z) = 1 - z^{n+1}$. Ist $|z| < 1$, so ist z^n eine Nullfolge, denn $|z^n| = |z|^n$ ist eine Nullfolge (siehe Beispiel (b) nach Satz 2.1.5). Daraus folgt aus den Rechenregeln für konvergente Folgen die Konvergenz von $s_n = \frac{1 - z^{n+1}}{1 - z}$ mit

$$\lim_{n \rightarrow \infty} s_n = \frac{1}{1 - z} \lim_{n \rightarrow \infty} (1 - z^{n+1}) = \frac{1}{1 - z}.$$

Zum Beispiel erhalten wir für

$$\begin{aligned} z = \frac{1}{2}: \quad 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \dots &= \frac{1}{1 - \frac{1}{2}} = 2 \\ z = -\frac{1}{2}: \quad 1 - \frac{1}{2} + \frac{1}{4} - \frac{1}{8} + \dots &= \frac{1}{1 + \frac{1}{2}} = \frac{2}{3}. \end{aligned}$$

Eine einfache notwendige Bedingung für die Konvergenz einer Reihe liefert der folgende Satz.

Satz 2.5.2. Sei $\sum_{k=1}^{\infty} c_k$ eine konvergente Reihe. Dann bilden die Glieder c_k der Reihe eine Nullfolge.

Beweis. Sei $s_n = \sum_{k=1}^n c_k$ eine konvergente Folge. Dann ist $s_{n+1} = \sum_{k=1}^{n+1} c_k$ eine Teilfolge mit gleichem Grenzwert, und somit ist $s_{n+1} - s_n = c_{n+1}$ eine Nullfolge. Dann ist aber auch c_n eine Nullfolge. \square

Bemerkung. Die Reihe $\sum_{k=0}^{\infty} z^k$ ist divergent, falls $|z| \geq 1$, denn die Glieder der Reihe bilden keine Nullfolge.

Aus den Rechenregeln für konvergente Folgen erhalten wir:

Satz 2.5.3. (Rechenregeln für konvergente Reihen)

Seien $\sum_{k=1}^{\infty} c_k$ und $\sum_{k=1}^{\infty} d_k$ konvergente Reihen mit den Grenzwerten $\sum_{k=1}^{\infty} c_k = c$ und $\sum_{k=1}^{\infty} d_k = d$. Dann folgt:

$$(1) \quad \sum_{k=1}^{\infty} (c_k + d_k) \text{ ist konvergent mit } \sum_{k=1}^{\infty} (c_k + d_k) = c + d.$$

$$(2) \quad \text{Ist } a \in \mathbb{C}, \text{ so ist } \sum_{k=1}^{\infty} (ac_k) \text{ konvergent mit } \sum_{k=1}^{\infty} (ac_k) = a \sum_{k=1}^{\infty} c_k.$$

Beweis. Da die Partialsummen $s_n = \sum_{k=1}^n c_k$ und $s'_n = \sum_{k=1}^n d_k$ konvergente Folgen mit Grenzwert c und Grenzwert d sind, erhalten wir aus den Rechenregeln für konvergente Folgen: $s_n + s'_n = \sum_{k=1}^n (c_k + d_k)$ ist konvergent mit Grenzwert $c + d$. Genauso folgt (2). \square

Nun wollen wir uns mit speziellen reellen Reihen beschäftigen, nämlich den alternierenden Reihen.

Definition 2.5.4. Eine reelle Reihe heißt *alternierend*, wenn sie von der Form

$$\sum_{k=1}^{\infty} (-1)^{k+1} a_k \quad \text{oder} \quad \sum_{k=1}^{\infty} (-1)^k a_k$$

ist, wobei $a_k \geq 0$ für alle $k \in \mathbb{N}$ gelten soll.

Alternierende Reihen sind schon konvergent, falls die Reihenglieder monoton gegen 0 konvergieren. Dies ist das Leibnizkriterium.

Satz 2.5.5. (Leibnizkriterium)

Eine alternierende Reihe $\sum_{k=1}^{\infty} (-1)^{k+1} a_k$ ist konvergent, falls a_k eine monoton fallende Nullfolge ist.

Beweis. Betrachte die Teilfolgen s_{2n} und s_{2n+1} von $s_n = \sum_{k=1}^n (-1)^{k+1} a_k$.

Da $s_{2n+2} = s_{2n} + (a_{2n+1} - a_{2n+2}) \geq s_{2n}$, ist die Folge (s_{2n}) monoton wachsend.

Da $s_{2n+3} = s_{2n+1} + (-a_{2n+2} + a_{2n+3}) = s_{2n+1} - (a_{2n+2} - a_{2n+3}) \leq s_{2n+1}$, ist die Folge s_{2n+1} monoton fallend.

Außerdem gilt: $s_{2n+1} - s_{2n} = (-1)^{2n+2} a_{2n+1} = a_{2n+1} \geq 0$. Daraus folgt:

$$s_2 \leq s_4 \leq \dots \leq s_{2n} \leq s_{2n+1} \leq \dots \leq s_3 \leq s_1.$$

Insbesondere sind beide Folgen beschränkt und wegen ihrer Monotonie konvergent. Da $s_{2n+1} - s_{2n}$ eine Nullfolge ist, haben beide Teilfolgen den gleichen Grenzwert s , d.h. für jedes $\epsilon > 0$ sind fast alle Folgenglieder von (s_{2n}) sowie von (s_{2n+1}) in der ϵ -Umgebung von s . Dann gilt dies aber auch für die Folge (s_n) . \square

Beispiel. Insbesondere ist die Reihe $\sum_{k=1}^{\infty} (-1)^{k+1} \frac{1}{k}$ konvergent. Das Leibnizkriterium (der ein reiner Existenzsatz ist) gibt uns nicht die Möglichkeit den Grenzwert zu berechnen. Der Grenzwert ist $\log 2$, wie wir später sehen werden.

Ein hinreichendes und notwendiges Kriterium für die Konvergenz einer Reihe ist das *Cauchy-Kriterium*. Es folgt aus dem entsprechenden Kriterium für Folgen.

Satz 2.5.6. (Konvergenzkriterium von Cauchy)

Eine Reihe $\sum_{k=1}^{\infty} c_k$ ist konvergent genau dann, falls gilt:

$$\forall \epsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n > m \geq n_0 : \left| \sum_{k=m+1}^n c_k \right| < \epsilon.$$

Beweis. Wegen Satz 2.4.7 ist die Folge $s_n = \sum_{k=1}^n c_k$ genau dann konvergent, falls sie eine

Cauchyfolge bildet. Da für $n > m$ gilt: $|s_n - s_m| = \left| \sum_{k=1}^n c_k - \sum_{k=1}^m c_k \right| = \left| \sum_{k=m+1}^n c_k \right|$, folgt die

Behauptung. \square

Reihen, deren Glieder eine Nullfolge bilden, sind nicht notwendigerweise konvergent. Ein wichtiges Beispiel dafür ist:

Beispiel. (*harmonische Reihe*)

Die harmonische Reihe $\sum_{n=1}^{\infty} \frac{1}{n}$ ist nicht konvergent, denn

$$\left| \sum_{n=n_0+1}^{2n_0} \frac{1}{n} \right| = \left| \frac{1}{n_0+1} + \dots + \frac{1}{2n_0} \right| \geq \frac{1}{2n_0} n_0 = \frac{1}{2}.$$

Damit erfüllt die harmonische Reihe nicht das Cauchysche Konvergenzkriterium.

Für spezielle reelle Reihen lassen sich noch folgende Konvergenzkriterien angeben.

Satz 2.5.7. Sei $\sum_{k=1}^{\infty} a_k$ eine reelle Reihe mit $a_k \geq 0$. Diese Reihe ist genau dann konvergent,

falls die Folge der Partialsummen $s_n = \sum_{k=1}^n a_k$ nach oben beschränkt ist.

Beweis. Da die Folge (s_n) monoton steigend ist, folgt ihre Konvergenz wegen Satz 2.3.2 aus der Beschränktheit der Folge (s_n) . \square

Beispiel. Die Reihe $\sum_{k=1}^{\infty} \frac{1}{k^2}$ ist konvergent, denn

$$\begin{aligned} s_n &= \sum_{k=1}^n \frac{1}{k^2} = 1 + \sum_{k=2}^n \frac{1}{k^2} \leq 1 + \sum_{k=2}^n \frac{1}{k(k-1)} = 1 + \sum_{k=2}^n \left(\frac{1}{k-1} - \frac{1}{k} \right) \\ &= 1 + \left(1 - \frac{1}{2} + \frac{1}{2} - \frac{1}{3} + \frac{1}{3} - \frac{1}{4} + \dots + \frac{1}{n-1} - \frac{1}{n} \right) = 1 + 1 - \frac{1}{n} \leq 2. \end{aligned}$$

Damit ist s_n beschränkt und die Reihe konvergiert wegen des obigen Satzes. Es ist jedoch nicht so einfach, ihren Grenzwert zu bestimmen; er ist $\frac{\pi^2}{6}$.

Definition 2.5.8. Eine (beliebige) Reihe $\sum_{k=1}^{\infty} c_k$ heißt *absolut konvergent*, wenn die Reihe $\sum_{k=1}^{\infty} |c_k|$ konvergiert.

Satz 2.5.9. Eine absolut konvergente Reihe ist konvergent.

Bemerkung. Die Umkehrung ist natürlich falsch, denn die alternierende Reihe $\sum_{k=1}^{\infty} (-1)^{k+1} \frac{1}{k}$ ist konvergent, aber wegen des Beispiels der harmonischen Reihe nicht absolut konvergent.

Beweis. Die Reihe $\sum_{k=1}^{\infty} |c_k|$ ist genau dann konvergent, wenn sie das Cauchy Kriterium erfüllt, d.h.

$$\forall \epsilon > 0 \exists n_0 \forall n > m \geq n_0 : \sum_{k=m+1}^n |c_k| < \epsilon.$$

Da $\left| \sum_{k=m+1}^n c_k \right| \leq \sum_{k=m+1}^n |c_k| < \epsilon$, erfüllt auch die Reihe $\sum_{k=1}^{\infty} c_k$ das Cauchy Kriterium. \square

Nun wollen wir einige Kriterien für die absolute Konvergenz von Reihen angeben. Sie beruhen auf *Vergleichskriterien* mit anderen Reihen.

Satz 2.5.10. (Majorantenkriterium)

Es seien $\sum_{k=1}^{\infty} c_k$ und $\sum_{k=1}^{\infty} d_k$ zwei Reihen mit $|c_k| \leq |d_k|$ für fast alle $k \in \mathbb{N}$.

(1) Sei $\sum_{k=1}^{\infty} |d_k|$ eine konvergente Reihe. Dann ist die Reihe $\sum_{k=1}^{\infty} c_k$ absolut konvergent.

(2) Sei $\sum_{k=1}^{\infty} c_k$ divergent. Dann divergiert auch die Reihe $\sum_{k=1}^{\infty} |d_k|$.

Beweis.

(1) Sei $|c_k| \leq |d_k|$ für $k \geq k_0$. Ist $n > k_0$, so folgt: die Summe

$$\sum_{k=1}^n |c_k| = \sum_{k=1}^{k_0} |c_k| + \sum_{k=k_0+1}^n |c_k| \leq \sum_{k=1}^{k_0} |c_k| + \sum_{k=k_0+1}^n |d_k| \leq \sum_{k=1}^{k_0} |c_k| + \sum_{k=k_0+1}^{\infty} |d_k|$$

ist beschränkt. Daher ist die Folge $s_n = \sum_{k=1}^n |c_k|$ konvergent.

(2) folgt aus (1). Denn ist $\sum_{k=1}^{\infty} c_k$ divergent so ist diese Reihe insbesondere nicht absolut konvergent. Wegen (1) muß daher die Reihe $\sum_{k=1}^{\infty} |d_k|$ divergieren. \square

Bemerkung. Wir nennen die Reihe $\sum_{k=1}^{\infty} d_k$ eine *Majorante* von $\sum_{k=1}^{\infty} c_k$. Umgekehrt nennen wir die Reihe $\sum_{k=1}^{\infty} c_k$ eine *Minorante* von $\sum_{k=1}^{\infty} d_k$.

Für das folgende Konvergenzkriterium benutzen wir die geometrische Reihe als Vergleichsreihe.

Satz 2.5.11. (Wurzelkriterium)

Sei $\sum_{k=1}^{\infty} c_k$ mit $\overline{\lim}_{k \rightarrow \infty} \sqrt[k]{|c_k|} = r$. Dann gilt:

(a) Ist $r < 1$, so ist die Reihe $\sum_{k=1}^{\infty} c_k$ absolut konvergent.

(b) Ist $r > 1$, so ist die Reihe $\sum_{k=1}^{\infty} c_k$ divergent.

Falls $\overline{\lim}_{k \rightarrow \infty} \sqrt[k]{|c_k|}$ nicht existiert, d.h. ist $\sqrt[k]{|c_k|}$ unbeschränkt, so ist die Reihe erst recht divergent.

Bemerkung. Für $r = 1$ kann die Reihe konvergieren oder auch divergieren. Beispiele hierfür ist die divergente Reihe $\sum_{k=1}^{\infty} \frac{1}{k}$ und die konvergente Reihe $\sum_k \frac{1}{k^2}$.

In beiden Fällen gilt wegen des Beispiels (d) nach Satz 2.1.5:

$$\lim_{k \rightarrow \infty} \sqrt[k]{\frac{1}{k}} = \lim_{k \rightarrow \infty} \frac{1}{\sqrt[k]{k}} = 1 \quad \text{und} \quad \lim_{k \rightarrow \infty} \sqrt[k]{\frac{1}{k^2}} = \lim_{k \rightarrow \infty} \frac{1}{\sqrt[k]{k} \sqrt[k]{k}} = 1.$$

Beweis. Sei $r < 1$. Wegen des Satzes 2.3.12 ist für alle $\epsilon > 0$ die Ungleichung $\sqrt[k]{|c_k|} \leq r + \epsilon$ für alle bis auf endlich viele $k \in \mathbb{N}$ erfüllt. Wähle $\epsilon > 0$, so dass $q = r + \epsilon < 1$ (z.B. $\epsilon = \frac{1-r}{2}$).

Dann ist $\sqrt[k]{|c_k|} \leq q$ und somit $|c_k| \leq q^k$ für fast alle $k \in \mathbb{N}$. Damit ist $\sum_{k=1}^{\infty} q^k$ konvergente

Majorante von $\sum_{k=1}^{\infty} c_k$. Wegen des Majorantenkriteriums konvergiert $\sum_{k=1}^{\infty} c_k$ absolut.

Sei $r > 1$. Da r Häufungspunkt der Folge $\sqrt[k]{|c_k|}$ ist, gilt $\sqrt[k]{|c_k|} \geq 1$ und somit $|c_k| \geq 1$ für unendlich viele $k \in \mathbb{N}$. Insbesondere ist c_k keine Nullfolge und daher ist $\sum_{k=1}^{\infty} c_k$ divergent.

Trivialerweise gilt dies auch, falls die Folge $\sqrt[k]{|c_k|}$ unbeschränkt ist. \square

Satz 2.5.12. (Quotientenkriterium)

Sei $c_k \neq 0$ für fast alle $k \in \mathbb{N}$. Dann ist die Reihe $\sum_{k=1}^{\infty} c_k$ absolut konvergent, falls

$$\overline{\lim}_{k \rightarrow \infty} \frac{|c_{k+1}|}{|c_k|} = r < 1.$$

Die Reihe $\sum_{k=1}^{\infty} c_k$ divergiert, falls $\frac{|c_{k+1}|}{|c_k|} \geq 1$ für fast alle $k \in \mathbb{N}$.

Bemerkung. Aus $\overline{\lim}_{k \rightarrow \infty} \frac{|c_{k+1}|}{|c_k|} > 1$ folgt nicht die Divergenz.

Beispiel. Betrachte eine Reihe $\sum_{k=1}^{\infty} c_k$ mit

$$c_{2k} = \left(\frac{1}{2}\right)^{2k} \quad \text{und} \quad c_{2k+1} = \left(\frac{3}{4}\right)^{2k+1}$$

Dann folgt:

$$\frac{c_{2k+1}}{c_{2k}} = \frac{3}{4} \left(\frac{\frac{3}{4}}{\frac{1}{2}}\right)^{2k} = \frac{3}{4} \left(\frac{6}{4}\right)^{2k},$$

und somit ist $\frac{c_{2k+1}}{c_{2k}}$ nicht einmal nach oben beschränkt.

Die Reihe $\sum_{k=1}^{\infty} c_k$ ist aber konvergent, denn mit $d_k = \left(\frac{3}{4}\right)^k$ ist $\sum_{k=1}^{\infty} d_k$ eine konvergente Majorante.

Beweis. Ist $\overline{\lim}_{k \rightarrow \infty} \frac{|c_{k+1}|}{|c_k|} = r < 1$, so gilt für $r < q < 1$ und fast alle $k \in \mathbb{N}$:

$|c_{k+1}| \leq q|c_k|$ (mit demselben Argument wie in Satz 2.5.11). Dann existiert ein $k_0 \in \mathbb{N}$ mit $|c_{k+1}| \leq q|c_k|$ für alle $k \geq k_0$ und deshalb folgt:

$$|c_{k_0+j}| \leq q^j |c_{k_0}| =: b_{j+k_0}$$

für alle $j \in \mathbb{N}$. Also ist, für beliebig gewählte b_1, \dots, b_{k_0} , die Reihe $\sum_{j=1}^{\infty} b_j$ eine konvergente Majorante. Gilt $|c_{k+1}| \geq |c_k|$ für fast alle $k \in \mathbb{N}$, so existiert ein $k_0 \in \mathbb{N}$ mit $c_{k_0} \neq 0$ und $|c_k| \geq |c_{k_0}|$ für alle $k \geq k_0$. Insbesondere ist c_k keine Nullfolge. \square

Bemerkung. Jede Reihe, die das Quotientenkriterium erfüllt, erfüllt auch das Wurzelkriterium. Dies folgt aus der Ungleichung

$$\overline{\lim}_{k \rightarrow \infty} \sqrt[k]{|c_k|} \leq \overline{\lim}_{k \rightarrow \infty} \frac{|c_{k+1}|}{|c_k|},$$

denn ist $\overline{\lim}_{k \rightarrow \infty} \frac{|c_{k+1}|}{|c_k|} = r$, so existiert für alle $\epsilon > 0$ ein k_0 mit $c_{k+1} \leq (r + \epsilon)c_k$ für alle $k \geq k_0$.

Also folgt mit Argumenten analog zu denen im obigen Beweis: $|c_{k_0+j}| \leq (r + \epsilon)^j |c_{k_0}|$ für alle $j \geq 0$ bzw. $|c_k| \leq (r + \epsilon)^{k-k_0} |c_{k_0}|$ für alle $k \geq k_0$. Dies impliziert für alle $\epsilon > 0$:

$$\overline{\lim}_{k \rightarrow \infty} \sqrt[k]{|c_k|} \leq r + \epsilon.$$

Da dies für alle $\epsilon > 0$ gilt, folgt $\overline{\lim}_{k \rightarrow \infty} \sqrt[k]{|c_k|} \leq r$.

Damit ist das Wurzelkriterium stärker als das Quotientenkriterium. Auf der anderen Seite ist das Quotientenkriterium oft leichter zu handhaben.

Wichtiges Beispiel für die Anwendung des Quotientenkriteriums ist der folgende Satz, mit dem wir die Exponentialfunktion einführen.

Satz 2.5.13. Für jedes $z \in \mathbb{C}$ konvergiert die Reihe

$$\sum_{k=0}^{\infty} \frac{1}{k!} z^k = 1 + z + \frac{1}{2!} z^2 + \frac{1}{3!} z^3 + \dots$$

absolut.

Beweis. Sei $c_k := \frac{z^k}{k!}$ für $k \in \mathbb{N}_0$. Dann folgt:

$$\frac{|c_{k+1}|}{|c_k|} = \frac{|z|^{k+1}}{(k+1)!} \frac{k!}{|z|^k} = \frac{1}{k+1} |z|.$$

Da $\lim_{k \rightarrow \infty} \frac{1}{k+1} |z| = 0$, konvergiert die Reihe wegen des Quotientenkriteriums absolut. \square

Definition 2.5.14. Die Funktion $\exp : \mathbb{C} \rightarrow \mathbb{C}$ mit

$$\exp(z) := \sum_{k=0}^{\infty} \frac{1}{k!} z^k$$

heißt *Exponentialfunktion*.

Bemerkung. Im nächsten Abschnitt (Korollar 2.6.5) werden wir das wichtige Additionstheorem

$$\exp(z + w) = \exp(z) \cdot \exp(w).$$

für alle $z, w \in \mathbb{C}$ beweisen. Dann werden wir auch diese fundamentale Funktion systematisch untersuchen. Es gibt noch eine weitere Darstellung der Exponentialfunktion, nämlich mittels einer Folge.

Satz 2.5.15. *Es gilt für alle $z \in \mathbb{C}$:*

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

Beweis. Es gilt

$$\begin{aligned} \left(1 + \frac{z}{n}\right)^n &= \sum_{k=0}^n \binom{n}{k} \frac{z^k}{n^k} = \sum_{k=0}^n \frac{n(n-1) \cdots (n-(k-1))}{n^k} \frac{z^k}{k!} \\ &= \sum_{k=0}^n \left(1 - \frac{1}{n}\right) \cdots \left(1 - \frac{k-1}{n}\right) \frac{z^k}{k!} \\ &= 1 + z + \left(1 - \frac{1}{n}\right) \frac{z^2}{2!} + \dots + \left(1 - \frac{1}{n}\right) \cdots \left(1 - \frac{n-1}{n}\right) \frac{z^n}{n!}. \end{aligned}$$

Zu $\epsilon > 0$ existiert ein n_0 mit $\sum_{k=n_0+1}^{\infty} \frac{|z|^k}{k!} \leq \frac{\epsilon}{3}$ (siehe Bemerkung (b) nach Definition 2.5.1).

Dann erhalten wir für alle $n \geq n_0$

$$\begin{aligned} \left| \left(1 + \frac{z}{n}\right)^n - \sum_{k=0}^{\infty} \frac{1}{k!} z^k \right| &\leq \sum_{k=0}^{n_0} \left| \left(1 - \frac{1}{n}\right) \cdots \left(1 - \frac{k-1}{n}\right) - 1 \right| \frac{|z|^k}{k!} \\ &\quad + \sum_{k=n_0+1}^n \underbrace{\left| \left(1 - \frac{1}{n}\right) \cdots \left(1 - \frac{k-1}{n}\right) - 1 \right|}_{\in [0,1]} \frac{|z|^k}{k!} + \sum_{k=n+1}^{\infty} \frac{|z|^k}{k!} \\ &\leq \sum_{k=0}^{n_0} \left| \left(1 - \frac{1}{n}\right) \cdots \left(1 - \frac{k-1}{n}\right) - 1 \right| \frac{|z|^k}{k!} + \sum_{k=n_0+1}^{\infty} \frac{|z|^k}{k!} + \sum_{k=n_0+1}^{\infty} \frac{|z|^k}{k!} \\ &\leq \sum_{k=0}^{n_0} \left| \left(1 - \frac{1}{n}\right) \cdots \left(1 - \frac{k-1}{n}\right) - 1 \right| \frac{|z|^k}{k!} + \frac{\epsilon}{3} + \frac{\epsilon}{3}. \end{aligned}$$

Die Koeffizienten der noch abzuschätzenden endlichen Summe konvergieren für $n \rightarrow \infty$ gegen Null. Also existiert ein $n_1 \geq n_0$, so dass diese Summe für alle $n \geq n_1$ sich ebenfalls durch $\frac{\epsilon}{3}$ abschätzen lässt. Also folgt:

Für jedes $\epsilon > 0$ existiert ein $n_1 \in \mathbb{N}$ mit

$$\left| \left(1 + \frac{z}{n}\right)^n - \sum_{k=0}^{\infty} \frac{1}{k!} z^k \right| \leq \epsilon \quad \text{für alle } n \geq n_1.$$

□

2.6 Umordnung von Reihen, absolute Konvergenz und Doppelreihen

Die Summe einer endlichen Reihe $\sum_{k=1}^n c_k$ ist wegen der Kommutativität der Addition in \mathbb{C} unabhängig von der Reihenfolge der Summanden, d.h. ist $\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ eine Permutation, so folgt: $\sum_{k=1}^n c_k = \sum_{k=1}^n c_{\sigma(k)}$.

Man kann sich die Frage stellen, ob sich dies auch auf unendliche Reihen übertragen läßt. Im allgemeinen lautet die Antwort nein.

Beispiel. $\sum_{k=1}^{\infty} (-1)^{k+1} \frac{1}{k}$ ist wegen des Leibnizkriteriums eine konvergente Reihe. Da

$$s_2 = 1 - \frac{1}{2} < s_{2n} < s_{2n+1} < s_{2n-1} < s_1 = 1$$

liegt ihr Grenzwert a zwischen $\frac{1}{2}$ und 1. Wie oben schon erwähnt, werden wir später sehen, dass $a = \log 2$ ist.

Wir ordnen nun die Reihe

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \dots$$

so um, dass auf jede positive Zahl zwei negative Zahlen folgen, d.h.

$$\begin{aligned} & 1 - \frac{1}{2} - \frac{1}{4} + \frac{1}{3} - \frac{1}{6} - \frac{1}{8} + \frac{1}{5} - \frac{1}{10} - \frac{1}{12} + \frac{1}{7} - \frac{1}{14} - \frac{1}{16} + \dots \\ &= \left(1 - \frac{1}{2}\right) - \frac{1}{4} + \left(\frac{1}{3} - \frac{1}{6}\right) - \frac{1}{8} + \left(\frac{1}{5} - \frac{1}{10}\right) - \frac{1}{12} + \left(\frac{1}{7} - \frac{1}{14}\right) - \frac{1}{16} + \dots \\ &= \frac{1}{2} - \frac{1}{4} + \frac{1}{6} - \frac{1}{8} + \frac{1}{10} - \frac{1}{12} + \frac{1}{14} - \frac{1}{16} + \dots \\ &= \frac{1}{2} \left(1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8} + \dots\right) = \frac{1}{2}a, \end{aligned}$$

d.h. der Grenzwert der umgeordneten Reihe ist nur halb so groß wie der Grenzwert der ursprünglichen Reihe.

Man kann die ursprüngliche Reihe sogar so umordnen, dass jeder beliebige Grenzwert vorkommt, ja sogar so, dass sie divergiert.

Definition 2.6.1. Sei eine Reihe und $\sigma : \mathbb{N} \rightarrow \mathbb{N}$ eine bijektive Abbildung (Permutation). Dann nennen wir die Reihe

$$\sum_{k=1}^{\infty} c_{\sigma(k)}$$

eine *Umordnung* von $\sum_{k=1}^{\infty} c_k$.

Die Klasse der absolut konvergenten Reihen behalten, wie der folgende Satz zeigt, nach Umordnung ihren Grenzwert bei.

Satz 2.6.2. *Es sei $\sum_{k=1}^{\infty} c_k$ absolut konvergent. Dann ist auch jede Umordnung $\sum_{k=1}^{\infty} c_{\sigma(k)}$ konvergent mit*

$$\sum_{k=1}^{\infty} c_{\sigma(k)} = \sum_{k=1}^{\infty} c_k.$$

Beweis. Da die Reihe $\sum_{k=1}^{\infty} c_k$ konvergent ist, konvergiert ihre Partialsummenfolge $s_n := \sum_{k=1}^n c_k$ gegen $c := \sum_{k=1}^{\infty} c_k$. Betrachte die Partialsummenfolge

$$t_n := \sum_{k=1}^n c_{\sigma(k)}.$$

Es genügt zu zeigen, dass $|t_n - s_n|$ eine Nullfolge ist, denn dann ist

$$|t_n - c| = |t_n - s_n + s_n - c| \leq |t_n - s_n| + |s_n - c|$$

ebenfalls eine Nullfolge und somit konvergiert t_n ebenfalls gegen c . Sei $\epsilon > 0$. Da die Reihe $\sum_{k=1}^{\infty} c_k$ absolut konvergent ist, existiert wegen des Cauchyschen Konvergenzkriteriums ein n_0 mit

$$\sum_{k=n_0+1}^n |c_k| < \epsilon$$

für alle $n \geq n_0$. Zu gegebenem $n \in \mathbb{N}$ betrachte die Mengen

$$M_n = \{1, \dots, n\} \quad \text{sowie} \quad T_n = \{k_1, \dots, k_n\},$$

mit $k_i = \sigma(i)$. Da $\sigma : \mathbb{N} \rightarrow \mathbb{N}$ bijektiv ist, existiert ein $n_1 \geq n_0$ mit:

$$M_{n_0} \subset T_{n_1} = \sigma(M_{n_1}).$$

Dann folgt für alle $n \geq n_1$:

$$\begin{aligned} |t_n - s_n| &= \left| \sum_{j \in T_n} c_j - \sum_{j \in M_n} c_j \right| = \left| \sum_{j \in T_n \setminus M_n} c_j - \sum_{j \in M_n \setminus T_n} c_j \right| \\ &\leq \sum_{j \in T_n \setminus M_n} |c_j| + \sum_{j \in M_n \setminus T_n} |c_j| \\ &= \sum_{j \in (T_n \setminus M_n) \cup (M_n \setminus T_n)} |c_j| < \epsilon, \end{aligned}$$

denn $(T_n \setminus M_n) \cup (M_n \setminus T_n) \subset (T_n \setminus M_{n_0}) \cup (M_n \setminus M_{n_0}) \subset \{j \in \mathbb{N} \mid j \geq n_0 + 1\}$. □

Bemerkung. Man kann sogar die Umkehrung des obigen Satzes zeigen:

Gilt für jede Umordnung σ : $\sum_{k=1}^{\infty} c_k = \sum_{k=1}^{\infty} c_{\sigma(k)}$. Dann ist $\sum_{k=1}^{\infty} c_k$ absolut konvergent.

Wir wollen nun den Umordnungssatz auf *Doppelreihen*

$$\sum_{i,k \in \mathbb{N}} a_{ik}$$

anwenden, wobei $a_{ik} \in \mathbb{C}$.

Dazu betrachten wir das quadratische Schema

$$\begin{array}{cccccc} a_{11} & a_{12} & a_{13} & a_{14} & \dots \\ a_{21} & a_{22} & a_{23} & a_{24} & \dots \\ a_{31} & a_{32} & a_{33} & a_{34} & \dots \\ a_{41} & a_{42} & a_{43} & a_{44} & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{array}$$

Zu diesem Schema betrachten wir die *Zeilenreihen* Z_i , *Spaltenreihen* S_k und die *Diagonalreihen* D_n gegeben durch

$$Z_i := \sum_{k=1}^{\infty} a_{ik}, \quad S_k = \sum_{i=1}^{\infty} a_{ik} \quad \text{und} \quad D_n = \sum_{i+k=n} a_{ik}.$$

Doppelreihen entstehen z.B. bei der Multiplikation von gewöhnlichen Reihen. Sind $\sum_{i=1}^{\infty} c_i$ und $\sum_{k=1}^{\infty} d_k$ Reihen, so gilt für das Produkt der n -ten mit der m -ten Partialsumme:

$$\left(\sum_{i=1}^m c_i \right) \left(\sum_{k=1}^n d_k \right) = \sum_{i=1}^m \sum_{k=1}^n c_i d_k.$$

Die Koeffizienten $a_{ik} := c_i d_k$ definieren dann eine Doppelreihe. Der folgende Satz folgt mit Hilfe des Umordnungssatzes.

Satz 2.6.3. Sei $\sum_{i,k \in \mathbb{N}} a_{ik}$ eine Doppelreihe und

$$\left\{ \sum_{k=1}^n \sum_{i=1}^m |a_{ik}| \mid n, m \in \mathbb{N} \right\}$$

eine beschränkte Menge mit Supremum s . Dann sind ihre Zeilen-, Spalten- und Diagonalreihen absolut konvergent, und es gilt:

$$\sum_{i=1}^{\infty} Z_i = \sum_{k=1}^{\infty} S_k = \sum_{n=2}^{\infty} D_n.$$

Beweis. Die Menge $\mathbb{N} \times \mathbb{N}$ ist abzählbar, d.h. es existiert eine Bijektion $\varphi : \mathbb{N} \rightarrow \mathbb{N} \times \mathbb{N}$. Wir wählen die Abzählung durch das schon in Beweis von Satz 2.3.15 beschriebene Diagonalverfahren, d.h. $1 \rightarrow (1, 1), 2 \rightarrow (1, 2), 3 \rightarrow (2, 1), 4 \rightarrow (3, 1), 5 \rightarrow (2, 2), 6 \rightarrow (1, 3) \dots$

Die Reihe $\sum_{k=1}^{\infty} a_{\varphi(k)}$ ist absolut konvergent, denn $\sum_{k=1}^n |a_{\varphi(k)}| \leq s$ für alle $n \in \mathbb{N}$. (In jedem Falle

hängt der Wert der Reihe $\sum_{k=1}^{\infty} a_{\varphi(k)}$ nicht von der Wahl von φ ab, denn eine andere Bijektion führt zu einer umgeordneten Reihe). Der Vorteil der Wahl der Abzählung durch das Diagonalverfahren besteht darin, dass dann die Partialsummenfolge der Diagonalreihe $\sum_{n=2}^{\infty} D_n$ mit

$D_n = \sum_{i+k=n} a_{ik}$ eine Teilfolge der Partialsummenfolge der Reihe $\sum_{k=1}^{\infty} a_{\varphi(k)}$ darstellt und somit den gleichen Grenzwert besitzen. Es gilt also bei dieser Wahl der Abzählung:

$$\sum_{k=1}^{\infty} a_{\varphi(k)} = \sum_{n=2}^{\infty} D_n \quad \text{mit} \quad D_n = \sum_{i+k=n} a_{ik}.$$

Außerdem überträgt sich die absolute Konvergenz von $\sum_{k=1}^{\infty} a_{\varphi(k)}$ auch auf die Diagonalreihe.

Es bleibt nun zu zeigen.

$$\sum_{k=1}^{\infty} a_{\varphi(k)} = \sum_{i=1}^{\infty} Z_i = \sum_{k=1}^{\infty} S_k$$

wobei $Z_i := \sum_{k=1}^{\infty} a_{ik}$ die Zeilenreihen und $S_k = \sum_{i=1}^{\infty} a_{ik}$ die Spaltenreihen bezeichnet. Die Reihen Z_i, S_k sind absolut konvergent, denn die zu den Absolutreihen gehörigen Partialsummenfolgen sind durch s beschränkt. Wir beweisen nur die erste Identität $\sum_{k=1}^{\infty} a_{\varphi(k)} = \sum_{i=1}^{\infty} Z_i$. Die zweite Identität folgt mit vollkommen analogen Argumenten.

Sei $\epsilon > 0$. Dann existiert ein $n_0 \in \mathbb{N}$ mit

$$\left| \sum_{k=1}^{\infty} a_{\varphi(k)} - \sum_{i=1}^m \sum_{k=1}^n a_{ik} \right| < \epsilon$$

für alle $n, m \geq n_0$. Diese Aussage (sie wurde in der Vorlesung nicht bewiesen) beweist man wie folgt:

Wegen der Definition von s existiert zu jedem $\epsilon > 0$ ein $n_0 \in \mathbb{N}$ mit

$$0 \leq s - \sum_{i=1}^{n_0} \sum_{k=1}^{n_0} |a_{ik}| < \epsilon.$$

Sind nun B_1, B_2 Teilmengen von $\mathbb{N} \times \mathbb{N}$ mit $M_{n_0} \times M_{n_0} \subset B_1 \subset B_2$, wobei $M_{n_0} = \{1, \dots, n_0\}$, so gilt:

$$\left| \sum_{(i,k) \in B_2} a_{ik} - \sum_{(i,k) \in B_1} a_{ik} \right| \leq \sum_{(i,k) \in B_2 \setminus B_1} |a_{ik}| = \sum_{(i,k) \in B_2} |a_{ik}| - \sum_{(i,k) \in B_1} |a_{ik}| \leq s - \sum_{i=1}^{n_0} \sum_{k=1}^{n_0} |a_{ik}| < \epsilon.$$

Zu $m, n \geq n_0$ und $B_1 = M_m \times M_n$ wähle $l_0 \in \mathbb{N}$ so groß, dass $\{\varphi(k) \mid k \in \{1, \dots, l_0\}\}$ die Menge B_1 umfasst. Also folgt für alle $l \geq l_0$:

$$\left| \sum_{k=1}^l a_{\varphi(k)} - \sum_{i=1}^m \sum_{k=1}^n a_{ik} \right| < \epsilon$$

und somit auch

$$\left| \sum_{k=1}^{\infty} a_{\varphi(k)} - \sum_{i=1}^m \sum_{k=1}^n a_{ik} \right| = \lim_{l \rightarrow \infty} \left| \sum_{k=1}^l a_{\varphi(k)} - \sum_{i=1}^m \sum_{k=1}^n a_{ik} \right| < \epsilon$$

Da die Reihen $\sum_{k=1}^{\infty} a_{ik}$ absolut konvergieren, also auch konvergieren, folgt

$$\lim_{n \rightarrow \infty} \left| \sum_{k=1}^{\infty} a_{\varphi(k)} - \sum_{i=1}^m \sum_{k=1}^n a_{ik} \right| = \left| \sum_{k=1}^{\infty} a_{\varphi(k)} - \sum_{i=1}^m \lim_{n \rightarrow \infty} \sum_{k=1}^n a_{ik} \right| = \left| \sum_{k=1}^{\infty} a_{\varphi(k)} - \sum_{i=1}^m \sum_{k=1}^{\infty} a_{ik} \right| < \epsilon$$

für alle $m \geq n_0$.

Also hat man für jedes $\epsilon > 0$ ein $n_0 \in \mathbb{N}$ mit $\left| \sum_{k=1}^{\infty} a_{\varphi(k)} - \sum_{i=1}^m \sum_{k=1}^{\infty} a_{ik} \right| < \epsilon$ für alle $m \geq n_0$;

damit konvergiert auch die Reihe $\sum_{i=1}^{\infty} \sum_{k=1}^{\infty} a_{ik}$ mit Grenzwert $\sum_{k=1}^{\infty} a_{\varphi(k)}$. \square

Wie oben erwähnt, besteht Unsere Hauptanwendung in der Multiplikation von Reihen. Multipliziert man die Partialsummen der Reihen $\sum_{i=1}^{\infty} c_i, \sum_{k=1}^{\infty} d_k$

$$\left(\sum_{i=1}^n c_i \right) \left(\sum_{k=1}^m d_k \right) = \sum_{i=1}^n \sum_{k=1}^m c_i d_k,$$

so führt dies auf eine Doppelreihe $a_{ik} = c_i d_k$ mit den Zeilenreihen Z_i , Spaltenreihen S_k und Diagonalreihen D_n gegeben durch

$$Z_i = \sum_{k=1}^{\infty} c_i d_k = c_i \sum_{k=1}^{\infty} d_k, \quad S_k = d_k \sum_{i=1}^{\infty} c_i, \quad D_n = \sum_{i+k=n} c_i d_k = c_1 d_{n-1} + c_2 d_{n-2} + \dots + c_{n-1} d_1.$$

Aus Satz 2.6.3 erhalten wir:

Satz 2.6.4. Seien $\sum_{i=1}^{\infty} c_i$ und $\sum_{k=1}^{\infty} d_k$ absolut konvergente Reihen, so gilt.

$$\left(\sum_{i=1}^{\infty} c_i \right) \left(\sum_{k=1}^{\infty} d_k \right) \stackrel{(1)}{=} \sum_{i=1}^{\infty} \left(\sum_{k=1}^{\infty} c_i d_k \right) \stackrel{(2)}{=} \sum_{k=1}^{\infty} \left(\sum_{i=1}^{\infty} c_i d_k \right) \stackrel{(3)}{=} \sum_{n=2}^{\infty} D_n := \sum_{n=2}^{\infty} \sum_{i+k=n} c_i d_k.$$

Außerdem sind alle vorkommenden Reihen absolut konvergent.

Bemerkungen.

(a) Für (1) und (2) benötigt man nur die Konvergenz der Reihen $\sum_{i=1}^{\infty} c_i$ und $\sum_{k=1}^{\infty} d_k$.

Für (3) muss man die absolute Konvergenz der Reihen $\sum_{i=1}^{\infty} c_i$ und $\sum_{k=1}^{\infty} d_k$ voraussetzen.

(b) Die Diagonalreihe $\sum_{n=2}^{\infty} D_n := \sum_{n=2}^{\infty} \sum_{i+k=n} c_i d_k$ heißt auch *Cauchyprodukt* der Reihen $\sum_{i=1}^{\infty} c_i$ und $\sum_{k=1}^{\infty} d_k$. Falls die Summationsindizes der Reihen $\sum c_i$ und $\sum d_k$ mit i_0 und k_0 beginnen, so beginnt der Summationsindex des Cauchyproduktes $\sum D_n$ mit $i_0 + k_0$.

Beweis. Da $\sum_{i=1}^{\infty} c_i$ und $\sum_{k=1}^{\infty} d_k$ absolut konvergent sind, ist

$$\sum_{i=1}^m \sum_{k=1}^n |c_i d_k| = \left(\sum_{i=1}^m |c_i| \right) \left(\sum_{k=1}^n |d_k| \right) \leq \left(\sum_{i=1}^{\infty} |c_i| \right) \left(\sum_{k=1}^{\infty} |d_k| \right).$$

Deshalb können wir Satz 2.6.3 auf die Doppelreihe $\sum_{i,k \in \mathbb{N}} c_i d_k$ anwenden. Insbesondere sind für jedes $i \in \mathbb{N}$ und $k \in \mathbb{N}$ die Zeilenreihe $\sum_{k=1}^{\infty} c_i d_k = c_i \sum_{k=1}^{\infty} d_k$ und Spaltenreihe $\sum_{i=1}^{\infty} d_k c_i = d_k \sum_{i=1}^{\infty} c_i$ absolut konvergent und es gilt:

$$\left(\sum_{k=1}^{\infty} d_k \right) \left(\sum_{i=1}^{\infty} c_i \right) = \sum_{i=1}^{\infty} \left(c_i \left(\sum_{k=1}^{\infty} d_k \right) \right) = \sum_{i=1}^{\infty} \left(\sum_{k=1}^{\infty} c_i d_k \right) = \sum_{k=1}^{\infty} \left(\sum_{i=1}^{\infty} c_i d_k \right) = \sum_{n=2}^{\infty} D_n.$$

Dabei folgt das erste Gleichheitszeichen aus den Rechenregeln für konvergente Reihen. \square

Wir benutzen nun das Cauchyprodukt, um das Additionstheorem der Exponentialfunktion zu beweisen.

Korollar 2.6.5.

$$\exp(z) \cdot \exp(w) = \exp(z + w)$$

für alle $z, w \in \mathbb{C}$.

Beweis. Die Reihen $\sum_{k=0}^{\infty} \frac{z^k}{k!}$ und $\sum_{j=0}^{\infty} \frac{w^j}{j!}$ sind absolut konvergent für jedes $z, w \in \mathbb{C}$. Dann folgt aus Satz 2.6.4:

$$\begin{aligned} \exp(z) \cdot \exp(w) &= \left(\sum_{k=0}^{\infty} \frac{z^k}{k!} \right) \left(\sum_{j=0}^{\infty} \frac{w^j}{j!} \right) = \sum_{n=0}^{\infty} \sum_{k+j=n} \frac{z^k w^j}{k! j!} \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k+j=n} \frac{n!}{k! j!} z^k w^j = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^n \frac{n!}{k!(n-k)!} z^k w^{n-k} \\ &= \sum_{k=0}^{\infty} \frac{1}{n!} (z + w)^n \\ &= \exp(z + w). \end{aligned}$$

□

Nun wollen wir erste Folgerungen aus diesem Additionstheorem herleiten.

Satz 2.6.6.

- (a) Für alle $z \in \mathbb{C}$ gilt: $\exp(-z) = \exp(z)^{-1}$. Insbesondere ist $\exp(z) \neq 0$.
- (b) Für alle $x \in \mathbb{R}$ ist $\exp(x) \in \mathbb{R}_+ := \{y \mid y > 0\}$. Außerdem ist die reelle Exponentialfunktion

$$\exp : \mathbb{R} \rightarrow \mathbb{R}_+$$

streng monoton steigend.

- (c) Für alle $x \in \mathbb{Q}$ ist $\exp(x) = e^x$, wobei $e = \exp(1)$ die Eulersche Zahl ist.
- (d) $|\exp(z)| = \sqrt{\exp(z) \cdot \overline{\exp(z)}} = \exp(\operatorname{Re}(z))$ für alle $z \in \mathbb{C}$. Insbesondere ist $|\exp(ix)| = 1$ für alle $x \in \mathbb{R}$.

Beweis.

- (a) Aus $\exp(-z) \exp(z) = \exp(-z + z) = \exp(0) = 1$ folgt (a).
- (b) Da $\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \dots$, folgt für $x > 0$: $\exp(x) > 1$ und $\exp(0) = 1$.
Wegen $\exp(-x) \cdot \exp(x) = \exp(x - x) = 1$ erhalten wir: $\exp(-x) = \frac{1}{\exp(x)} > 0$ für alle $x \geq 0$. Ist $y > x$ so folgt:

$$\exp(y) = \exp((y - x) + x) = \exp(y - x) \cdot \exp(x) > \exp(x).$$

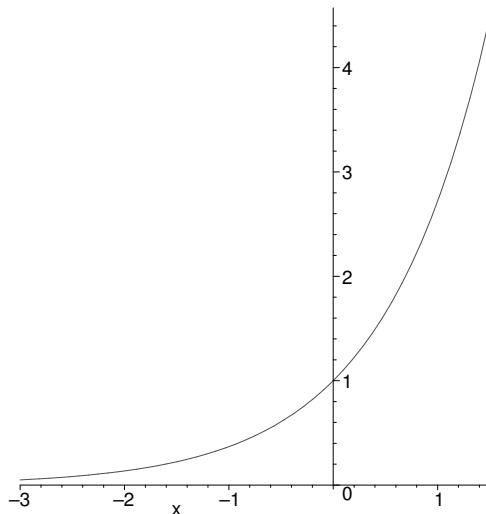


Abbildung 2.1: Die reelle Exponentialfunktion

- (c) Für $n \in \mathbb{N}$ gilt: $\exp(n) = (\exp(1))^n = e^n$ und wegen (a) auch $\exp(-n) = \frac{1}{e^n}$. Für rationale Zahlen $x = \frac{n}{m}$ mit $n \in \mathbb{Z}, m \in \mathbb{N}$ folgt

$$\left(\exp\left(\frac{n}{m}\right)\right)^m = \exp(n) = e^n \quad \text{und somit} \quad \exp(x) = (e^n)^{1/m} = e^{n/m} = e^x.$$

- (d) Wegen Bemerkung nach Satz 2.4.4 ist $\overline{\exp(z)} = \exp(\bar{z})$ und wir erhalten:

$$|\exp(z)| = \sqrt{\overline{\exp(z)} \cdot \exp(z)} = \sqrt{\exp(\bar{z} + z)} = \sqrt{\exp(2 \operatorname{Re}(z))} = \exp(\operatorname{Re}(z)).$$

Insbesondere ist $|\exp(ix)| = 1$ für alle $x \in \mathbb{R}$.

□

Bemerkungen.

- (1) Wir werden bald mit Hilfe des Zwischenwertsatzes zeigen: $\exp(\mathbb{R}) = \mathbb{R}_+$, d.h. die Abbildung $\exp : \mathbb{R} \rightarrow \mathbb{R}_+$ ist auch surjektiv. Die inverse Funktion $\log : \mathbb{R}_+ \rightarrow \mathbb{R}$ heißt *Logarithmus*. Aus dem Additionstheorem der Exponentialfunktion folgt dann

$$\log(x \cdot y) = \log(x) + \log(y).$$

- (2) Da die Beziehung $e^x = \exp(x)$ für alle rationalen Zahlen gilt, definieren wir $e^z := \exp(z)$ für beliebiges $z \in \mathbb{C}$.
- (3) Wegen (d) liegen die komplexen Zahlen $\exp(ix)$ für alle $x \in \mathbb{R}$ auf dem Einheitskreis $S^1 := \{z \in \mathbb{C} \mid |z| = 1\}$. Der Realteil von $\exp(ix)$ heißt *Cosinus* von x , der Imaginärteil heißt *Sinus* von x . Wir setzen also:

$$\cos(x) = \frac{e^{ix} + e^{-ix}}{2} \quad \text{und} \quad \sin(x) = \frac{e^{ix} - e^{-ix}}{2i}.$$

Somit gilt

$$e^{ix} = \cos(x) + i \sin(x).$$

Wegen $|e^{ix}| = 1$ folgt:

$$\cos^2(x) + \sin^2(x) = 1.$$

Ist $z = x + iy$, so erhalten wir die wichtige Darstellung

$$e^z = e^{x+iy} = e^x(\cos(y) + i \sin(y)).$$

Die Cosinus- und Sinusfunktion lassen sich auf \mathbb{C} fortsetzen.

Definition 2.6.7. Für beliebige $z \in \mathbb{C}$ setze

$$\cos(z) = \frac{e^{iz} + e^{-iz}}{2} \quad \text{und} \quad \sin(z) = \frac{e^{iz} - e^{-iz}}{2i}.$$

Die daraus resultierende Beziehung

$$e^{iz} = \cos(z) + i \sin(z)$$

heißt *Eulersche Formel*.

Cosinus und Sinus haben folgende Eigenschaften.

Satz 2.6.8.

(a) $\cos(-z) = \cos(z)$ und $\sin(-z) = -\sin(z)$.

(b) *Es gelten folgende Additionstheoreme:*

$$\begin{aligned} \cos(z+w) &= \cos(z)\cos(w) - \sin(z)\sin(w) \\ \sin(z+w) &= \sin(z)\cos(w) + \cos(z)\sin(w). \end{aligned}$$

(c)

$$\cos^2(z) + \sin^2(z) = 1.$$

(d) *Es gelten die folgenden Reihendarstellungen:*

$$\begin{aligned} \cos(z) &= \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k}}{(2k)!} = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \frac{z^6}{6!} + \dots \\ \sin(z) &= \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k+1}}{(2k+1)!} = z - \frac{z^3}{3!} + \frac{z^5}{5!} - \frac{z^7}{7!} + \dots \end{aligned}$$

Beweis.

(a) Folgt unmittelbar aus der Definition.

(b)

$$\begin{aligned}
\cos(z+w) &= \frac{1}{2} \left(e^{i(z+w)} + e^{-i(z+w)} \right) = \frac{1}{2} \left(e^{iz} e^{iw} + e^{-iz} e^{-iw} \right) \\
&= \frac{1}{2} \left((\cos(z) + i \sin(z))(\cos(w) + i \sin(w)) + (\cos(z) - i \sin(z))(\cos(w) - i \sin(w)) \right) \\
&= \cos(z) \cos(w) - \sin(z) \sin(w).
\end{aligned}$$

Das Additionstheorem für den Sinus beweist man analog.

(c) Diese Identität haben wir für reelle Zahlen schon bewiesen. Sie gilt aber auch für beliebige komplexe Zahlen, denn

$$\cos^2(z) + \sin^2(z) = \frac{1}{4}(e^{iz} + e^{-iz})^2 - \frac{1}{4}(e^{iz} - e^{-iz})^2 = \frac{1}{4}(4e^{iz}e^{-iz}) = 1$$

(d) Die Reihendarstellung des Cosinus ergibt sich aus

$$\begin{aligned}
\cos(z) &= \frac{1}{2} (e^{iz} + e^{-iz}) = \frac{1}{2} \left(\sum_{k=0}^{\infty} \frac{(iz)^k}{k!} + \sum_{k=0}^{\infty} \frac{(-1)^k (iz)^k}{k!} \right) \\
&= \sum_{k=0}^{\infty} \frac{(iz)^{2k}}{(2k)!} = \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k}}{(2k)!}.
\end{aligned}$$

Die Reihendarstellung des Sinus erhält man analog.

□

Aus den Additionstheoremen ergeben sich weitere nützliche Identitäten. Aus den Reihendarstellungen erhält man Abschätzungen für $\cos(x)$ und $\sin(x)$.

Satz 2.6.9.

(a) Für alle $z \in \mathbb{C}$ gilt:

$$\begin{aligned}
\cos(2z) &= \cos^2(z) - \sin^2(z) \\
\sin(2z) &= 2 \sin(z) \cos(z).
\end{aligned}$$

(b) Für alle $z, w \in \mathbb{C}$ gilt:

$$\begin{aligned}
\cos(z) - \cos(w) &= -2 \sin\left(\frac{z+w}{2}\right) \sin\left(\frac{z-w}{2}\right) \\
\sin(z) - \sin(w) &= 2 \cos\left(\frac{z+w}{2}\right) \sin\left(\frac{z-w}{2}\right)
\end{aligned}$$

(c) Es gelten folgende Abschätzungen für den reellen Cosinus und Sinus. Ist

$$\cos(x) - \sum_{k=0}^n (-1)^k \frac{x^{2k}}{(2k)!} =: R_{2n+2}^c(x)$$

und

$$\sin(x) - \sum_{k=0}^n (-1)^k \frac{x^{2k+1}}{(2k+1)!} =: R_{2n+3}^s(x)$$

so folgt:

$$\begin{aligned} |R_{2n+2}^c(x)| &\leq \frac{|x|^{2n+2}}{(2n+2)!} \quad \text{für } |x| \leq 2n+3 \\ |R_{2n+3}^s(x)| &\leq \frac{|x|^{2n+3}}{(2n+3)!} \quad \text{für } |x| \leq 2n+4. \end{aligned}$$

Beweis.

(a) Folgt unmittelbar aus den Additionstheoremen in Satz 2.6.8 (b).

(b) Betrachte $z = \frac{z+w}{2} + \frac{z-w}{2}$ und $w = \frac{z+w}{2} + \frac{w-z}{2}$. Dann folgt aus den Additionstheoremen

$$\begin{aligned} \cos(z) &= \cos\left(\frac{z+w}{2}\right) \cos\left(\frac{z-w}{2}\right) - \sin\left(\frac{z+w}{2}\right) \sin\left(\frac{z-w}{2}\right) \\ \cos(w) &= \cos\left(\frac{z+w}{2}\right) \cos\left(\frac{w-z}{2}\right) - \sin\left(\frac{z+w}{2}\right) \sin\left(\frac{w-z}{2}\right) \\ &= \cos\left(\frac{z+w}{2}\right) \cos\left(\frac{z-w}{2}\right) + \sin\left(\frac{z+w}{2}\right) \sin\left(\frac{z-w}{2}\right) \end{aligned}$$

und somit ist

$$\cos(z) - \cos(w) = -2 \sin\left(\frac{z+w}{2}\right) \sin\left(\frac{z-w}{2}\right).$$

Analog beweist man die zweite Formel.

(c) Es gilt

$$\begin{aligned} |R_{2n+2}^c(x)| &= \left| \sum_{k=n+1}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} \right| \\ &= |(-1)^{n+1} \frac{x^{2n+2}}{(2n+2)!} \left(1 - \frac{x^2}{(2n+3)(2n+4)} + \frac{x^4}{(2n+3)(2n+4)(2n+5)(2n+6)} - \dots \right) | \\ &= \frac{|x|^{2n+2}}{(2n+2)!} |a_0 - a_1 + a_2 - a_3 + \dots| \end{aligned}$$

wobei $a_0 = 1$ und

$$a_k = a_{k-1} \frac{x^2}{(2n+2k+1)(2n+2k+2)}.$$

Also ist die Folge (a_k) für $|x| \leq 2n+3$ monoton fallend und aus dem Leibnizkriterium folgt die Konvergenz der zugehörigen alternierenden Reihe

$$a_0 - a_1 + a_2 - a_3 + \dots = \sum_{k=0}^{\infty} (-1)^k a_k.$$

Ihr Grenzwert liegt zwischen 0 und 1, denn ist $s_n = \sum_{k=0}^n (-1)^k a_k$, so gilt

$$0 < 1 - a_1 = s_1 < s_3 < \dots < s_{2n+1} < s_{2n} < \dots < s_2 < s_0 = 1.$$

Die zweite Abschätzung beweist man analog.

□

Korollar 2.6.10.

- (a) Für alle $x \in (0, 2]$ ist $\sin(x) > 0$.
- (b) $\cos : [0, 2] \rightarrow \mathbb{R}$ ist streng monoton fallend und es gilt: $\cos(0) = 1$ und $\cos(2) \leq -\frac{1}{3}$. Insbesondere hat \cos auf dem Intervall $[0, 2]$ genau eine Nullstelle.

Beweis.

- (a) Nach Satz 2.6.9(c) gilt für $x \in (0, 2]$:

$$\sin(x) = x + R_3^s(x) = x \left(1 + \frac{R_3^s(x)}{x} \right), \quad \text{wobei} \quad \left| \frac{R_3^s(x)}{x} \right| \leq \frac{x^2}{6} \leq \frac{2}{3}.$$

Also ist

$$\sin(x) = x \left(1 + \frac{R_3^s(x)}{x} \right) \geq x \left(1 - \frac{2}{3} \right).$$

- (b) Ist $0 \leq x < y \leq 2$, so folgt aus Satz 2.6.9(b):

$$\cos(y) - \cos(x) = -2 \sin\left(\frac{y+x}{2}\right) \sin\left(\frac{y-x}{2}\right) < 0,$$

denn $0 < \frac{y-x}{2} < \frac{y+x}{2} < 2$. Nach Satz 2.6.9(c) gilt für $|x| \leq 4$:

$$\cos(x) = 1 - \frac{x^2}{2} + R_4^s(x) \quad \text{wobei} \quad |R_4^s(x)| \leq \frac{x^4}{24}.$$

Insbesondere ist $|R_4^s(2)| \leq \frac{16}{24} = \frac{2}{3}$ und somit

$$\cos(2) \leq 1 - \frac{4}{2} + \frac{2}{3} = -\frac{1}{3}.$$

Aus der strengen Monotonie folgt die Injektivität von $\cos : [0, 2] \rightarrow \mathbb{R}$. Insbesondere existiert maximal eine Nullstelle im Intervall $[0, 2]$. Da $\cos(0) = 1$, wechselt $\cos : [0, 2] \rightarrow \mathbb{R}$ das Vorzeichen. Wie wir im nächsten Kapitel sehen werden, haben auf einem Intervall stetige reellwertige Funktionen, die das Vorzeichen wechseln, eine Nullstelle (*Zwischenwertsatz*).

□

Wir benutzen nun diese Nullstelle um die Zahl π zu definieren.

Definition 2.6.11. Sei $x_0 \in [0, 2]$ die eindeutig bestimmte Nullstelle von $\cos : [0, 2] \rightarrow \mathbb{R}$. Dann definieren wir $\pi = 2x_0$.

Bemerkung. Da $1 = |\cos(\frac{\pi}{2}) + i \sin(\frac{\pi}{2})| = |\sin(\frac{\pi}{2})|$ und wegen Korollar 2.6.10 und Definition 2.6.11 $\sin(\frac{\pi}{2}) > 0$ ist, folgt $\sin(\frac{\pi}{2}) = 1$.

Wir erhalten nun die folgenden Periodizitätseigenschaften der Exponentialfunktion und somit auch von Cosinus und Sinus

Korollar 2.6.12.

(a) Für alle $z \in \mathbb{C}$ gilt:

$$e^{z+\frac{\pi i}{2}} = ie^z, \quad e^{z+\pi i} = -e^z, \quad e^{z+2\pi i} = e^z.$$

(b) Für alle $z \in \mathbb{C}$ gilt:

$$\cos\left(z + \frac{\pi}{2}\right) = -\sin(z), \quad \cos(z + \pi) = -\cos(z), \quad \cos(z + 2\pi) = \cos(z)$$

$$\sin\left(z + \frac{\pi}{2}\right) = \cos(z), \quad \sin(z + \pi) = -\sin(z), \quad \sin(z + 2\pi) = \sin(z).$$

Beweis.

(a) Dies folgt aus dem Additionstheorem von e^z zusammen mit der Identität $e^{\frac{\pi i}{2}} = \cos(\frac{\pi}{2}) + i \sin(\frac{\pi}{2}) = i$.

(b) Folgt aus Teil (a) und der Definition von Cosinus und Sinus. Zum Beispiel ist

$$\begin{aligned} \cos\left(z + \frac{\pi}{2}\right) &= \frac{1}{2} \left(e^{iz} e^{\frac{\pi i}{2}} + e^{-iz} e^{-\frac{\pi i}{2}} \right) = \frac{1}{2} \left(e^{iz} i + e^{-iz} \frac{1}{i} \right) \\ &= \frac{1}{2i} (-e^{iz} + e^{-iz}) = -\sin(z) \end{aligned}$$

□

Daraus ergeben sich sofort für die Nullstellen des Cosinus und Sinus:

Korollar 2.6.13.

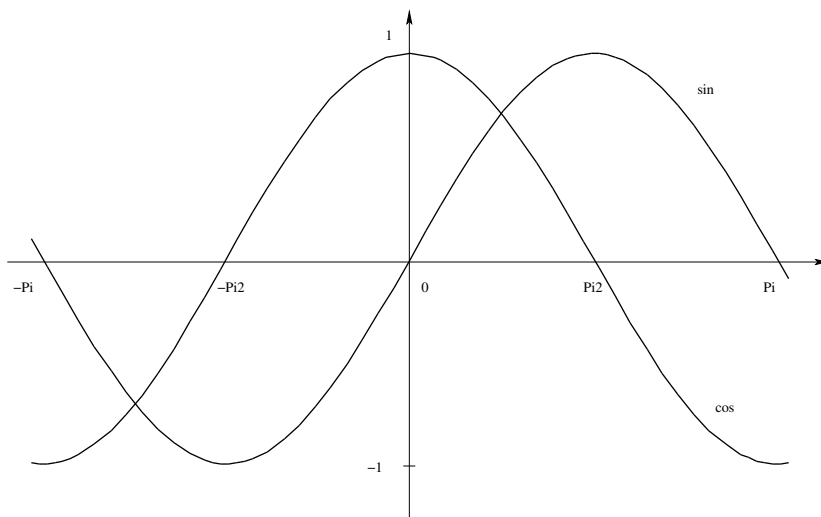
(a) $\{x \in \mathbb{R} \mid \sin x = 0\} = \{k\pi \mid k \in \mathbb{Z}\}$.

(b) $\{x \in \mathbb{R} \mid \cos x = 0\} = \{\frac{\pi}{2} + k\pi \mid k \in \mathbb{Z}\}$.

(c) $\{z \in \mathbb{C} \mid e^z = 1\} = \{2k\pi i \mid k \in \mathbb{Z}\}$.

(d) Sinus und Cosinus haben nur reelle Nullstellen.

Beweis.



- (a) Wegen Korollar 2.6.12 ist $|\sin(x + k\pi)| = |\sin(x)|$ und wegen $\sin(0) = 0$ folgt somit $\sin(k\pi) = 0$ für alle $k \in \mathbb{Z}$.

Die Umkehrung erhält man wie folgt. Da $\sin(x) = \cos(\frac{\pi}{2} - x) = \cos(-\frac{\pi}{2} + x)$, folgt aus der Definition von $\frac{\pi}{2}$:

$$\sin(x) > 0 \quad \text{für alle } 0 < x < \pi.$$

Ist nun $x \in \mathbb{R}$ eine beliebige reelle Zahl mit $\sin(x) = 0$, so existiert ein $k \in \mathbb{Z}$ mit $x = k\pi + y$ wobei $0 \leq y < \pi$. Dann folgt aus Korollar 2.6.12:

$$|\sin(y)| = |\sin(x - k\pi)| = |\sin(x)| = 0$$

und somit ist $y = 0$, d.h. $x = k\pi$.

- (b) Teil (b) folgt aus (a), denn $\cos(x) = -\sin(x - \frac{\pi}{2})$.

- (c) Es ist $e^{2k\pi i} = (e^{2\pi i})^k = 1$.

Sei nun umgekehrt $z = x + iy$ und $e^z = e^{x+iy} = 1$. Dann ist $e^x e^{iy} = 1$ und somit $|e^x e^{iy}| = |e^x| |e^{iy}| = e^x = 1$. Damit ist $x = 0$ und $1 = e^{iy} = \cos(y) + i \sin(y)$. Also ist $\sin(y) = 0$ und $\cos(y) = 1$. Also ist $y = m\pi$ mit $m \in \mathbb{Z}$. Wäre $m = 2k + 1$, so würde $\cos(y) = \cos(2k\pi + \pi) = \cos(\pi) = -1$ folgen. Also ist $m = 2k$.

- (d) Ist nun $\sin(z) = \frac{1}{2i}(e^{iz} - e^{-iz}) = 0$, so ist $e^{iz} = e^{-iz}$ und somit $e^{2iz} = 1$ und wegen (c) ist $z = k\pi$. Wegen $\cos(z) = \sin(z + \frac{\pi}{2})$ folgt die Aussage auch für den Cosinus.

□

Bemerkungen.

- (a) $\cos : [0, \pi] \rightarrow [-1, 1]$ ist streng monoton fallend, denn für $0 \leq x < y \leq \pi$ ist wegen 2.6.3

$$\cos(y) - \cos(x) = -2 \sin\left(\frac{y+x}{2}\right) \sin\left(\frac{y-x}{2}\right) < 0,$$

denn $0 < \frac{y-x}{2} < \frac{y+x}{2} < \pi$ und $\sin(t) > 0$ für $t \in (0, \pi)$. Da $\cos(0) = 1$ und $\cos(\pi) = -1$, wird aus dem Zwischenwertsatz die Bijektivität von $\cos : [0, \pi] \rightarrow [-1, 1]$ folgen. Dann existiert auch die Umkehrfunktion $\arccos : [-1, 1] \rightarrow [0, \pi]$. Diese Funktion heißt *Arcus Cosinus*.

- (b) Da $\cos(x) = \cos(-x)$, ist $\cos : [-\pi, 0] \rightarrow [-1, 1]$ wegen (a) monoton steigend. Aus $\sin(x) = \cos(x - \frac{\pi}{2})$ folgt dann, dass $\sin : [-\frac{\pi}{2}, \frac{\pi}{2}] \rightarrow [-1, 1]$ ebenfalls streng monoton steigend ist. Wiederum folgt aus dem Zwischenwertsatz, dass $\sin : [-\frac{\pi}{2}, \frac{\pi}{2}] \rightarrow [-1, 1]$ bijektiv ist. Die Umkehrfunktion $\arcsin : [-1, 1] \rightarrow [-\frac{\pi}{2}, \frac{\pi}{2}]$ heißt *Arcus Sinus*.

Zum Schluss wollen wir noch *Tangens* und *Cotangens* definieren.

Definition 2.6.14. Für alle $z \in \mathbb{C} \setminus \{\frac{\pi}{2} + k\pi \mid k \in \mathbb{Z}\}$ bzw. $z \in \mathbb{C} \setminus \{k\pi \mid k \in \mathbb{Z}\}$ definiere

$$\tan(z) := \frac{\sin(z)}{\cos(z)}, \quad \text{bzw.} \quad \cot(z) = \frac{\cos(z)}{\sin(z)}$$

2.7 Potenzreihen

Die Exponentialfunktion, die Cosinusfunktion, Sinusfunktion sowie die geometrische Reihe sind Beispiele von Potenzreihen. Allgemein versteht man darunter das Folgende.

Definition 2.7.1. Sei $(a_k)_{k \in \mathbb{N}_0}$ eine Folge komplexer Zahlen. Dann heißt die Reihe

$$\sum_{k=0}^{\infty} a_k (z - z_0)^k$$

eine *Potenzreihe* mit *Entwicklungspunkt* z_0 .

Beispiel. Die Exponentialfunktion $\exp z = \sum_{k=0}^{\infty} \frac{1}{k!} z^k$ ist eine Potenzreihe mit Entwicklungspunkt 0.

Satz 2.7.2. Es sei $\sqrt[k]{|a_k|}$ beschränkt und $\overline{\lim}_{k \rightarrow \infty} \sqrt[k]{|a_k|} = r$. Dann ist die Potenzreihe

$$\sum_{k=0}^{\infty} a_k (z - z_0)^k$$

im Falle $r = 0$ absolut konvergent für alle $z \in \mathbb{C}$. Ist $r > 0$, so ist die Potenzreihe absolut konvergent für alle $z \in \mathbb{C}$ mit

$$|z - z_0| < \frac{1}{r}$$

und divergent für alle $z \in \mathbb{C}$ mit

$$|z - z_0| > \frac{1}{r}.$$

Ist $\sqrt[k]{|a_k|}$ unbeschränkt, so konvergiert sie nur am Entwicklungspunkt $z = z_0$.

Beweis. Es gilt: $\sqrt[k]{|a_k|} |z - z_0|^k = |z - z_0|^k \sqrt[k]{|a_k|}$. Aus dem Wurzelkriterium folgt somit unmittelbar: Für jedes $z \in \mathbb{C}$ mit $|z - z_0| r < 1$ ist die Potenzreihe absolut konvergent. Ist $r = 0$, so liegt Konvergenz für jedes $z \in \mathbb{C}$ vor.

Für jedes $z \in \mathbb{C}$ mit $|z - z_0| r > 1$ ist die Potenzreihe divergent. Ist $\sqrt[k]{|a_k|}$ unbeschränkt, so ist auch $|z - z_0| \sqrt[k]{|a_k|}$ für alle $z \neq z_0$ unbeschränkt, d.h. die Reihe divergiert für alle $z \neq z_0$. \square

Bemerkung.

- (a) Die Zahl $R = \frac{1}{r}$ heißt *Konvergenzradius*. Ist $r = 0$, so setzen wir $R := \infty$.
- (b) Ist $R > 0$, so konvergiert die Reihe $\sum_{k=0}^{\infty} a_k(z - z_0)^k$ für jede Zahl in der Kreisscheibe $B(z_0, R) := \{z \mid |z - z_0| < R\}$ und definiert daher eine Abbildung $f : B(z_0, R) \rightarrow \mathbb{C}$ mit $f(z) = \sum_{k=0}^{\infty} a_k(z - z_0)^k$. Ist $R = \infty$, so ist die Abbildung auf ganz \mathbb{C} definiert.
- (c) Ist $(a_k)_{k \in \mathbb{N}_0}$ eine Folge komplexer Zahlen mit $a_m = 0$ für $m > n$, so gilt

$$\sum_{k=0}^{\infty} a_k z^k = a_0 + a_1 z + \dots + a_n z^n,$$

d.h. die Potenzreihe ist ein *Polynom*. Ist $a_n \neq 0$, so heißt n der *Grad* des Polynoms.

Beispiele.

- (a) Betrachte die geometrische Reihe $\sum_{k=0}^{\infty} z^k$. Der Konvergenzradius ist 1, denn $\sqrt[k]{1} = 1$. Daher konvergiert die Reihe für $|z| < 1$. Dies haben wir natürlich auch schon im Abschnitt 2.5 gesehen. Dort haben wir auch gezeigt: $\sum_{k=0}^{\infty} z^k = \frac{1}{1-z}$. Für $z \in \mathbb{C}$ mit $|z| = 1$ divergiert die Reihe, denn ihre Glieder bilden keine Nullfolge.
- (b) Ist $|a_k|$ durch eine Konstante $\rho > 0$ beschränkt, so ist $\sum a_k z^k$ konvergent für $|z| < 1$, denn $\sum \rho |z|^k$ ist für jedes z mit $|z| < 1$ eine konvergente Majorante von $\sum_{k=0}^{\infty} |a_k| |z|^k$.

Der folgende Satz über Addition und Multiplikation von Potenzreihen folgt aus den entsprechenden Sätzen für absolut konvergente Reihen. Wir beschränken uns auf den Fall mit Entwicklungspunkt $z_0 = 0$. Für beliebiges z_0 muß z durch $z - z_0$ ersetzt werden.

Satz 2.7.3. Sei $\sum_{n=0}^{\infty} a_n z^n$ Potenzreihe mit Konvergenzradius R_1 und $\sum_{n=0}^{\infty} b_n z^n$ Potenzreihe mit Konvergenzradius R_2 . Ist $R = \min\{R_1, R_2\}$, so gilt für alle $z \in B(0, R)$:

$$\sum_{k=0}^{\infty} a_k z^k + \sum_{k=0}^{\infty} b_k z^k = \sum_{k=0}^{\infty} (a_k + b_k) z^k$$

$$\left(\sum_{k=0}^{\infty} a_k z^k \right) \left(\sum_{j=0}^{\infty} b_j z^j \right) = \sum_{n=0}^{\infty} \left(\sum_{k=0}^n a_k b_{n-k} \right) z^n.$$

Beweis. Der erste Teil folgt aus den Rechenregeln für konvergente Reihen, da beide Reihen für $|z| < R$ konvergieren. Der zweite Teil folgt aus dem Cauchyprodukt für absolut konvergente Reihen (Satz 2.6.4), denn

$$\left(\sum_{k=0}^{\infty} a_k z^k \right) \left(\sum_{j=0}^{\infty} b_j z^j \right) = \sum_{n=0}^{\infty} \sum_{k+j=n} a_k b_j z^{k+j} = \sum_{n=0}^{\infty} \left(\sum_{k=0}^n a_k b_{n-k} \right) z^n.$$

□

Eine Potenzreihe $\sum_{k=0}^{\infty} a_k z^k$ ist im allgemeinen nicht auf dem ganzen Konvergenzgebiet beschränkt, z.B. $\sum_{k=0}^{\infty} z^k$. Es gilt aber:

Satz 2.7.4. Sei $f(z) = \sum_{k=0}^{\infty} a_k z^k$ Potenzreihe mit Konvergenzradius R . Sei $r < R$, so ist f auf $\bar{B}(0, r) := \{z \mid |z| \leq r\} \subset B(0, R)$ beschränkt.

Beweis. Sei $x \in B(0, r)$. Dann gilt:

$$\left| \sum_{k=0}^{\infty} a_k z^k \right| \leq \sum_{k=0}^{\infty} |a_k| |z|^k \leq \sum_{k=0}^{\infty} |a_k| r^k.$$

Da $r \in B(0, R)$, ist diese Reihe konvergent. □

Von besonderem Interesse ist das Verhalten des Restterms $R_n(z) := \sum_{k=n}^{\infty} a_k z^k$ einer konvergenten Potenzreihe.

Satz 2.7.5. Sei $\sum_{k=0}^{\infty} a_k z^k$ eine konvergente Potenzreihe mit Konvergenzradius $R > 0$. Sei $0 < r < R$, so existiert eine Konstante $c > 0$ mit

$$|R_n(z)| \leq c|z|^n$$

für $|z| \leq r$.

Beweis. Sei $|z| \leq r$. Dann gilt:

$$\begin{aligned} |R_n(z)| &\leq \sum_{k=n}^{\infty} |a_k| |z|^k = |z|^n (|a_n| + |a_{n+1}| |z| + \dots) \\ &\leq |z|^n \sum_{j=0}^{\infty} |a_{n+j}| r^j \leq c|z|^n \end{aligned}$$

mit $c = \frac{1}{r^n} \sum_{j=0}^{\infty} |a_{n+j}| r^{j+n}$. □

Bemerkung. Ist $|a_n|$ durch eine Konstante ρ beschränkt, so ist $R \geq 1$, und es gilt für $|z| < 1$:

$$|R_n(z)| \leq |z|^n \rho (1 + |z| + \dots + |z|^n + \dots) = |z|^n \frac{\rho}{1 - |z|}.$$

Aus obigem Satz folgt, dass sich die Nullstellen in der Umgebung des Entwicklungspunktes nicht häufen können.

Satz 2.7.6. Sei $f(z) = \sum_{n=0}^{\infty} a_n z^n$ eine nicht triviale konvergente Potenzreihe (d.h. $a_k \neq 0$ für wenigstens ein $k \in \mathbb{N}_0$). Dann gibt es ein $r > 0$, so dass f auf $\bar{B}(0, r) \setminus \{0\}$ keine Nullstellen besitzt. Ist $a_0 = f(0) \neq 0$, so existieren also auch keine Nullstellen in $B(0, r)$.

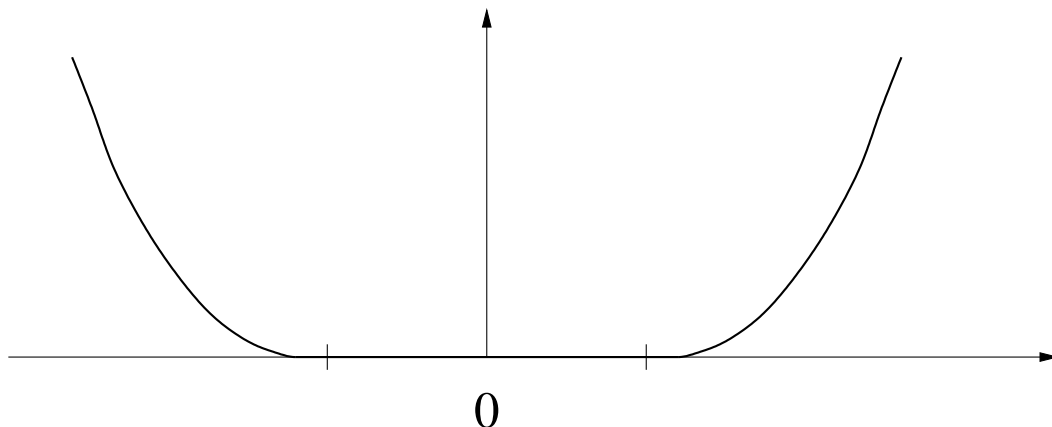


Abbildung 2.2: Beispiel einer nicht in eine Potenzreihe entwickelbaren Funktion

Beweis. Sei n der erste Index mit $a_n \neq 0$. Sei $0 < r_0 < R$, wobei R den Konvergenzradius von f bezeichnet. Dann folgt aus Satz 2.7.5:

$$|f(z) - a_n z^n| = |R_{n+1}(z)| < c|z|^{n+1}$$

für $|z| < r_0$ und ein $c > 0$. Da $|a + b| \geq |a| - |b|$, folgt

$$|f(z)| = |a_n z^n + f(z) - a_n z^n| \geq |a_n| |z|^n - |f(z) - a_n z^n| \geq |a_n| |z|^n (1 - c|z|).$$

Ist $z \neq 0$, so folgt: $f(z) \neq 0$ falls $|z| < r_0$ und $|z| < \frac{1}{c}$. Setze also $r = \min\{\frac{1}{c}, r_0\}$ □

Satz 2.7.6 impliziert den folgenden Identitätssatz.

Satz 2.7.7. (Identitätssatz)

Seien $f(z) = \sum_{k=0}^{\infty} a_k z^k$ und $g(z) = \sum_{k=0}^{\infty} b_k z^k$ zwei konvergente Potenzreihen. Ist (z_n) eine Nullfolge mit $z_n \neq 0$ und ist $f(z_n) = g(z_n)$, so gilt $a_k = b_k$ für alle $k \in \mathbb{N}_0$.

Beweis. Ist $h(z) = f(z) - g(z)$, so folgt $h(z_n) = 0$ für eine Nullfolge (z_n) mit $z_n \neq 0$. Dann sind aber wegen Satz 2.7.6 alle Koeffizienten $a_k - b_k = 0$, denn sonst gäbe es ein $r > 0$, so dass h auf $B(0, r) \setminus \{0\}$ keine Nullstellen hätte. □

Bemerkung. Funktionen, die sich in Potenzreihen entwickeln lassen, haben also alle obige Eigenschaft. Zum Beispiel können sich nicht triviale Funktionen, die in einer kleinen Umgebung der Null identisch Null sind, nicht in eine Potenzreihe entwickeln lassen (siehe Abbildung 2.2).

Eine wichtige Klasse von Funktionen stellen die sogenannten *Lipschitz-stetigen* (*dehnungsbeschränkten*) Funktionen dar.

Definition 2.7.8. Sei $D \subset \mathbb{C}$ und $f: D \rightarrow \mathbb{R}$ eine Funktion. f heißt *Lipschitz-stetig*, falls es eine Konstante $L > 0$ gibt, mit $|f(z) - f(w)| \leq L|z - w|$ für alle $z, w \in D$.

Satz 2.7.9. Sei $f(z) = \sum_{n=0}^{\infty} a_n z^n$ eine Potenzreihe mit Konvergenzradius $R > 0$. Ist $0 < r < R$, so ist f Lipschitz-stetig auf $\bar{B}(0, r)$.

Beweis. Da

$$(z - w)(z^{n-1} + z^{n-2}w + z^{n-3}w^2 + \dots + w^{n-1}) = z^n - w^n,$$

folgt:

$$f(z) - f(w) = \sum_{n=1}^{\infty} a_n(z^n - w^n) = (z - w) \cdot \sum_{n=1}^{\infty} a_n(z^{n-1} + z^{n-2}w + z^{n-3}w^2 + \dots + w^{n-1}).$$

Sei nun $|z|, |w| \leq r$, so folgt

$$|f(z) - f(w)| \leq |z - w| \sum_{n=1}^{\infty} |a_n|n \cdot r^{n-1}.$$

Die Reihe $\sum_{n=1}^{\infty} |a_n|n \cdot r^{n-1}$ konvergiert, denn

$$\overline{\lim}_{n \rightarrow \infty} \sqrt[n]{|a_n|n} = \overline{\lim}_{n \rightarrow \infty} \sqrt[n]{|a_n|} \lim_{n \rightarrow \infty} \sqrt[n]{n} = \overline{\lim}_{n \rightarrow \infty} \sqrt[n]{|a_n|},$$

d.h. ihr Konvergenzradius stimmt mit dem der Reihe $\sum_{n=0}^{\infty} a_n z^n$ überein. Da $r < R$ ist, folgt die Konvergenz. Insbesondere ist f Lipschitz-stetig auf $\bar{B}(0, r)$ mit $L = \sum_{n=1}^{\infty} |a_n|n \cdot r^{n-1}$. \square

Nun wollen wir uns mit dem wichtigen Begriff der Stetigkeit beschäftigen.

Kapitel 3

Stetigkeit

3.1 Definition und grundlegende Eigenschaften stetiger Funktionen

In diesem Abschnitt betrachten wir Funktionen $f: D \rightarrow \mathbb{C}$ mit Definitionsbereich $D \subset \mathbb{R}$ oder $D \subset \mathbb{C}$. Außerdem ist $f(D) \subset \mathbb{R}$ nicht ausgeschlossen. Die Stetigkeit einer solchen Funktion in einem Punkt $a \in D$ soll bedeuten, dass sich die Funktion in der Umgebung von a nicht zu stark ändert. Dies lässt sich wie folgt präzisieren.

Definition 3.1.1. Eine Funktion $f: D \rightarrow \mathbb{C}$ heißt *stetig* in $a \in D$, falls zu jeder ϵ -Umgebung $B(f(a), \epsilon)$ von $f(a)$ eine δ -Umgebung $B(a, \delta)$ von a existiert, so dass gilt:

$$f(B(a, \delta) \cap D) \subset B(f(a), \epsilon).$$

f heißt *stetig auf D* , falls f in jedem Punkt $a \in D$ stetig ist.

Bemerkungen.

- (a) Die Stetigkeit in einem Punkte $a \in D$ lässt sich in der Sprache der Aussagenlogik auch wie folgt schreiben:

$$\forall \epsilon > 0 \quad \exists \delta > 0 \quad \forall x \in D: |x - a| < \delta \Rightarrow |f(x) - f(a)| < \epsilon.$$

- (b) In den Abbildungen 3.1 und 3.2 wird die Stetigkeit einer Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$ und einer Funktion $f: \mathbb{C} \rightarrow \mathbb{C}$ veranschaulicht. Im reellen Falle betrachten wir den Graph der Funktion.

Satz 3.1.2. Jede Lipschitz-stetige Funktion $f: D \rightarrow \mathbb{C}$ ist stetig auf D .

Beweis. Ist f Lipschitz-stetig, so existiert ein $L > 0$ mit $|f(x) - f(y)| \leq L|x - y|$. Ist $a \in D$ und $\epsilon > 0$, so wähle $\delta = \frac{\epsilon}{L}$. Ist $x \in D$ und $|x - a| < \delta = \frac{\epsilon}{L}$, so folgt: $|f(x) - f(a)| \leq L|x - a| < \epsilon$. \square

Bemerkung. In Abschnitt 2.7 haben wir gesehen, dass eine Potenzreihe auf jeder abgeschlossenen Kreisscheibe $\bar{B}(0, r)$ Lipschitz-stetig ist, falls r kleiner als ihr Konvergenzradius R gewählt wird. Da für jedes $a \in B(0, R)$, ein $r < R$ existiert mit $a \in \bar{B}(0, r)$, ist mit dem obigen Satz jede Potenzreihe stetig auf ganz $B(0, R)$. Allerdings ist die Potenzreihe i. allg. nicht auf ganz $B(0, R)$ Lipschitz-stetig, denn die Lipschitzkonstanten auf $\bar{B}(0, r)$ sind von r abhängig.

Ein nützliches Kriterium für die Stetigkeit ist der folgende Satz.

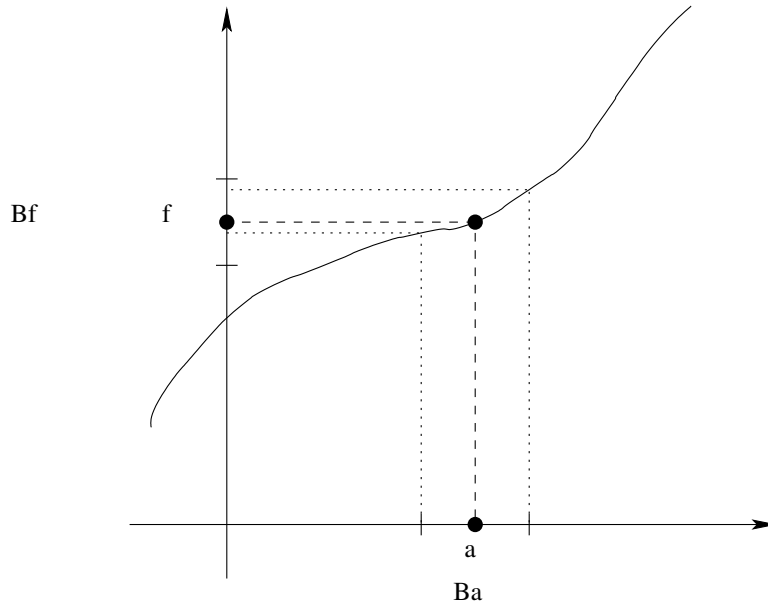


Abbildung 3.1: Stetigkeit einer Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$ im Punkt a .

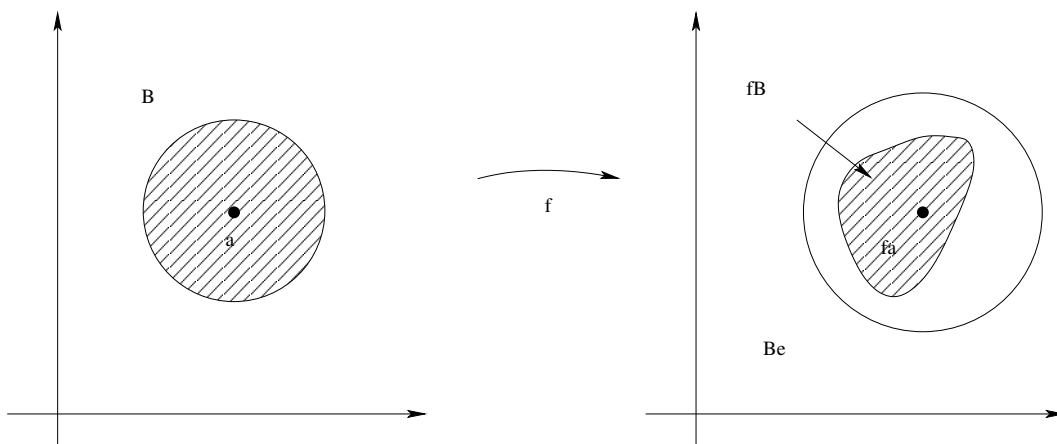


Abbildung 3.2: Stetigkeit einer Funktion $f: \mathbb{C} \rightarrow \mathbb{C}$ im Punkt a .

Satz 3.1.3. Sei $f: D \rightarrow \mathbb{C}$ eine Funktion. Dann gilt:

f ist stetig in $a \in D$ genau dann, falls für jede Folge (x_n) mit $x_n \in D$ und $\lim_{n \rightarrow \infty} x_n = a$ folgt:

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

Bemerkung. Stetige Funktionen bilden also konvergente Folgen auf konvergente Folgen ab.

Beweis. “ \Rightarrow ”: Seien f stetig in $a \in D$ und $x_n \in D$ eine Folge mit $\lim_{n \rightarrow \infty} x_n = a$. Wir zeigen:

$$\lim_{n \rightarrow \infty} f(x_n) = f(a), \text{ d.h. } \forall \epsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : |f(x_n) - f(a)| < \epsilon.$$

Sei $\epsilon > 0$. Da f stetig in $a \in D$ ist, existiert ein $\delta > 0$ mit $f(B(a, \delta) \cap D) \subset B(f(a), \epsilon)$. Wegen

$\lim_{n \rightarrow \infty} x_n = a$, gibt es ein n_0 , so dass $x_n \in B(a, \delta)$ für alle $n \geq n_0$. Dann folgt: $f(x_n) \in B(f(a), \epsilon)$

und somit $|f(x_n) - f(a)| < \epsilon$ für alle $n \geq n_0$.

“ \Leftarrow ”: Angenommen f wäre nicht stetig, d.h.

$$\exists \epsilon > 0 \quad \forall \delta > 0 \quad \exists x \in D : |x - a| < \delta \wedge |f(x) - f(a)| \geq \epsilon.$$

Dann existiert für jedes $n \in \mathbb{N}$ und $\delta = \frac{1}{n}$ ein $x_n \in D$ mit $|x_n - a| < \frac{1}{n}$ und $|f(x_n) - f(a)| \geq \epsilon$. Damit ist die Folge (x_n) konvergent mit $\lim_{n \rightarrow \infty} x_n = a$, aber die Folge $f(x_n)$ konvergiert nicht gegen $f(a)$ im Widerspruch zur Voraussetzung. \square

Funktionen $f, g: D \rightarrow \mathbb{C}$ kann man addieren und multiplizieren. Dies führt, wie der folgende Satz zeigt, nicht aus der Klasse der stetigen Funktionen heraus.

Satz 3.1.4. Sei $x_0 \in D$ und $f, g: D \rightarrow \mathbb{R}$ stetig in x_0 . Dann sind $(f + g): D \rightarrow \mathbb{C}$ mit

$$(f + g)(x) := f(x) + g(x)$$

sowie $(f \cdot g): D \rightarrow \mathbb{C}$ mit

$$(f \cdot g)(x) = f(x) \cdot g(x)$$

wieder stetig in x_0 . Ist $f(x_0) \neq 0$, so existiert $\delta > 0$, so dass $f(x) \neq 0$ für alle $x \in D' := B(x_0, \delta) \cap D$. Außerdem ist $\frac{1}{f}: D' \rightarrow \mathbb{R}$ mit

$$\frac{1}{f}(x) := \frac{1}{f(x)}$$

stetig in $x_0 \in D'$.

Beweis. Der erste Teil folgt aus den Rechenregeln für konvergente Folgen. Denn ist (x_n) Folge in D mit $\lim_{n \rightarrow \infty} x_n = x_0$, so ist $\lim_{n \rightarrow \infty} f(x_n) = f(x_0)$ und $\lim_{n \rightarrow \infty} g(x_n) = g(x_0)$. Damit erhalten wir

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

und

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

Ist $f(x_0) \neq 0$, so existiert ein $\epsilon > 0$ mit $0 \notin B(f(x_0), \epsilon)$. Wegen der Stetigkeit von f gibt es ein $\delta > 0$ mit $f(B(x_0, \delta) \cap D) \subset B(f(x_0), \epsilon)$. Insbesondere gilt: $f(x) \neq 0$ auf D' und die Stetigkeit in x_0 folgt wieder aus den Rechenregeln für konvergente Folgen. \square

Die Komposition stetiger Funktionen ist natürlich auch wieder stetig.

Satz 3.1.5. *Es seien $D_1, D_2 \subset \mathbb{C}$ und $f: D_1 \rightarrow D_2$ sowie $g: D_2 \rightarrow \mathbb{C}$ Funktionen. Ist f stetig in x_0 und g stetig in $f(x_0)$, so ist $(g \circ f): D_1 \rightarrow \mathbb{C}$ stetig in x_0 .*

Beweis. Sei (x_n) Folge in D_1 mit $\lim_{n \rightarrow \infty} x_n = x_0$. Da f stetig in x_0 ist, gilt: $\lim_{n \rightarrow \infty} f(x_n) = f(x_0)$. Aus der Stetigkeit von g in $f(x_0)$ folgt:

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

□

Bemerkung. Da, wie am Anfang des Abschnittes bemerkt, alle Potenzreihen auf ihrem Konvergenzgebiet stetig sind, sind sowohl die Exponentialfunktion $\exp: \mathbb{C} \rightarrow \mathbb{C}$ als auch die trigonometrischen Funktionen $\cos: \mathbb{C} \rightarrow \mathbb{C}$ und $\sin: \mathbb{C} \rightarrow \mathbb{C}$ Beispiele für stetige Funktionen. Wegen Satz 3.1.4 sind damit auch die in 2.6.14 definierten Funktionen $\tan: \mathbb{C} \setminus \{\frac{\pi}{2} + k\pi \mid k \in \mathbb{Z}\} \rightarrow \mathbb{C}$ bzw. $\cot: \mathbb{C} \setminus \{k\pi \mid k \in \mathbb{Z}\} \rightarrow \mathbb{C}$ mit

$$\tan(z) := \frac{\sin(z)}{\cos(z)}, \quad \text{bzw.} \quad \cot(z) = \frac{\cos(z)}{\sin(z)}$$

stetig.

3.2 Stetige reelle Funktionen auf Intervallen und der Zwischenwertsatz

Nun wollen wir uns mit reellen Funktionen beschäftigen, d.h mit Funktionen deren Definitionsbereich und Wertebereich Teilmengen der reellen Zahlen sind. Insbesondere wollen wir den Fall, dass der Definitionsbereich ein Intervall ist, untersuchen. Dabei sind Intervalle Teilmengen von \mathbb{R} der folgenden Form. Ist $a < b$, so definiere

$$[a, b] = \{x \in \mathbb{R} \mid a \leq x \leq b\} \quad (\text{kompaktes Intervall})$$

$$(a, b) = \{x \in \mathbb{R} \mid a < x < b\} \quad (\text{offenes Intervall})$$

$$(a, b] = \{x \in \mathbb{R} \mid a < x \leq b\} \\ (\text{halboffene Intervalle})$$

$$[a, b) = \{x \in \mathbb{R} \mid a \leq x < b\}$$

Außerdem betrachten wir *nicht beschränkte (uneigentliche) Intervalle* der Form:

$$(-\infty, a] = \{x \in \mathbb{R} \mid x \leq a\}, \quad (-\infty, a) = \{x \in \mathbb{R} \mid x < a\}, \quad (-\infty, +\infty) = \mathbb{R}$$

sowie analog $[a, +\infty)$ und $(a, +\infty)$. Die Intervalle $[a, b]$, $(-\infty, a]$, $[a, +\infty)$ sowie $(-\infty, +\infty)$ heißen auch *abgeschlossen*. Die einzigen beschränkten und abgeschlossenen Intervalle sind also von der Form $[a, b]$. In der Definition von $[a, b]$ lassen wir auch $a = b$ zu.

Die folgende Charakterisierung von Intervallen ist nützlich.

Satz 3.2.1. (Charakterisierung von Intervallen) *Sei $D \subset \mathbb{R}$. Dann ist D genau dann ein Intervall, wenn für alle $x_1, x_2 \in D$ mit $x_1 \leq x_2$ auch das Intervall $[x_1, x_2]$ in D enthalten ist.*

Beweis. Intervalle haben offensichtlich diese Eigenschaft. Sei nun umgekehrt $D \subset \mathbb{R}$, so dass für alle $x_1, x_2 \in D$ mit $x_1 \leq x_2$ auch das Intervall $[x_1, x_2]$ in D enthalten ist. Sei zunächst D beschränkt. Wir können annehmen, dass D wenigstens 2 Elemente enthält, denn sonst ist D von der Form $[a, a]$. Wir zeigen:

$$(\inf D, \sup D) \subset D \subset [\inf D, \sup D].$$

Da jede Menge mit dieser Eigenschaft ein Intervall ist, folgt die Behauptung für beschränkte Teilmengen von \mathbb{R} . Zunächst ist $D \subset [\inf D, \sup D]$ immer erfüllt, denn $\inf D$ bzw. $\sup D$ sind eine untere bzw. eine obere Schranke von D . Sei also $x \in (\inf D, \sup D)$ (nach Annahme ist $\inf D \neq \sup D$). Dann ist y weder untere Schranke noch obere Schranke von D und es existieren daher $x_1, x_2 \in D$ mit $x_1 < x < x_2$. Nach Voraussetzung an D ist daher $x \in [x_1, x_2] \subset D$.

Sei D nicht nach oben, jedoch nach unten beschränkt. Dann folgt mit analogen Argumenten:

$$(\inf D, \infty) \subset D \subset [\inf D, \infty).$$

und D ist ein unbeschränktes Intervall. Ähnliches gilt, falls D nicht nach unten, jedoch nach oben beschränkt ist. Ist D weder nach unten noch nach oben beschränkt, so ist $D = (-\infty, \infty) = \mathbb{R}$. \square

Einer der wichtigsten Sätze über stetige reellen Funktionen ist der *Zwischenwertsatz*. Der Zwischenwertsatz besagt: Eine stetige Funktion $f: [a, b] \rightarrow \mathbb{R}$ nimmt alle Werte zwischen $f(a)$ und $f(b)$ an.

Satz 3.2.2. (Zwischenwertsatz) Sei $f: [a, b] \rightarrow \mathbb{R}$ eine stetige Funktion. Dann existiert zu jeder Zahl y zwischen $f(a)$ und $f(b)$ ein $c \in [a, b]$ mit $f(c) = y$.

Beweis. Wir nehmen $f(a) < f(b)$ an (sonst betrachte $-f$).

Wir zeigen: Für alle $y \in [f(a), f(b)]$ existiert ein $c \in [a, b]$ mit $f(c) = y$. Sei $f(a) < y < f(b)$. Betrachte die Menge

$$M = \{x \in [a, b] \mid f(x) \leq y\}.$$

Da M beschränkt und nicht leer ist ($a \in M$), gibt es ein $c \in \mathbb{R}$ mit $c = \sup M$.

Wir zeigen: $c \in [a, b]$ und $f(c) = y$.

Trivialerweise ist a eine untere Schranke und b eine obere Schranke von M , also gilt $c \in [a, b]$. Da c kleinste obere Schranke ist, existiert für jedes $n \in \mathbb{N}$ ein $a_n \in M$ mit $c - \frac{1}{n} \leq a_n \leq c$. Damit definiert (a_n) eine konvergente Folge mit $\lim_{n \rightarrow \infty} a_n = c$. Außerdem ist $f(a_n) \leq y$ und wegen $\lim_{n \rightarrow \infty} f(a_n) = f(c)$ folgt $f(c) \leq y$.

Nach Annahme ist $y < f(b)$ und damit $c \neq b$. Betrachte eine konvergente Folge $b_n \in (c, b]$ mit $\lim_{n \rightarrow \infty} b_n = c$. Da $b_n > c$ ist $b_n \notin M$, also folgt $f(b_n) > y$. Somit erhalten wir: $y \leq \lim_{n \rightarrow \infty} f(b_n) = f(c)$, also $f(c) = y$. \square

Bemerkung.

Dieser Satz ist für nicht stetige Funktionen falsch. Betrachte z.B. die Funktion $f: [-1, +1] \rightarrow \mathbb{R}$ mit

$$f(x) = \begin{cases} 0, & \text{für } -1 \leq x < 0 \\ 1, & \text{für } 0 \leq x \leq 1 \end{cases}$$

Diese Funktion ist im Nullpunkt nicht stetig. Der Zwischenwertsatz ist nicht erfüllt, denn der Zwischenwert $\frac{1}{2}$ mit $f(-1) < \frac{1}{2} < f(1)$ wird nicht angenommen.

Wir wollen einige Folgerungen ableiten.

Korollar 3.2.3. *Sei I ein beliebiges Intervall und $f: I \rightarrow \mathbb{R}$ stetig. Dann ist $f(I)$ ebenfalls ein Intervall.*

Beweis. Seien $y_1, y_2 \in f(I)$ mit $y_1 \leq y_2$ und $y \in [y_1, y_2]$. Seien $x_1, x_2 \in I$ mit $f(x_1) = y_1$ sowie $f(x_2) = y_2$. Dann existiert wegen des Zwischenwertsatzes ein $x \in [x_1, x_2]$ (oder $[x_2, x_1]$ falls $x_2 < x_1$) mit $f(x) = y$. Da I ein Intervall ist, ist $x \in I$ und daher $y \in f(I)$. Wegen Satz 3.2.1 ist daher $f(I)$ ebenfalls ein Intervall. \square

Bemerkung. Offene Intervalle (a, b) müssen nicht unbedingt auf offene Intervalle abgebildet werden, wie z.B. die Abbildung $f: (-1, +1) \rightarrow \mathbb{R}$ mit $f(x) = x^2$ zeigt, denn $f((-1, +1)) = [0, 1)$. Hingegen ist das Bild kompakter Intervalle, wie der folgende Satz zeigt, unter stetigen Funktionen wieder kompakt. Insbesondere werden Infimum und Supremum angenommen. Genauer gilt:

Satz 3.2.4. *Sei $f: [a, b] \rightarrow \mathbb{R}$ stetig. Dann ist das Intervall $f([a, b]) \subset \mathbb{R}$ kompakt. Insbesondere existieren $x_M, x_m \in [a, b]$ mit*

$$f(x_M) = \sup f([a, b]) \text{ und } f(x_m) = \inf f([a, b]).$$

und somit ist

$$f([a, b]) = [f(x_m), f(x_M)].$$

Beweis. Wir zeigen zunächst: das Intervall $f([a, b])$ ist beschränkt.

Ist f unbeschränkt, so existiert zu jedem $n \in \mathbb{N}$ ein $x_n \in [a, b]$ mit $|f(x_n)| \geq n$. Da die Folge (x_n) beschränkt ist, existiert wegen des Satzes von Bolzano-Weierstraß 2.3.5 eine konvergente Teilfolge $x_{k(n)}$ mit $x_0 := \lim_{n \rightarrow \infty} x_{k(n)} \in [a, b]$.

Da $|f(x_{k(n)})| \geq |k(n)| \geq n$ unbeschränkt ist, konvergiert $f(x_{k(n)})$ nicht und somit ist f nicht stetig in x_0 .

Nun sei $s = \sup f([a, b])$. Dann existiert zu jedem $n \in \mathbb{N}$ ein $x_n \in [a, b]$ mit $s - \frac{1}{n} < f(x_n) \leq s$. Sei $x_{k(n)}$ konvergente Teilfolge von (x_n) so gilt:

$$\lim_{n \rightarrow \infty} x_{k(n)} =: x_0 \in [a, b] \text{ und } f(x_0) = \lim_{n \rightarrow \infty} f(x_{k(n)}) = s.$$

Genauso folgt, dass das Infimum angenommen wird. \square

Bemerkung.

- (a) Das Supremum der Menge $f([a, b])$ ist damit ein Maximum, das Infimum von $f([a, b])$ ist ein Minimum. Man sagt auch: Stetige Funktionen nehmen auf kompakten Intervallen ihr Maximum und Minimum an.
- (b) Der Satz ist für unstetige Funktionen i.allg. falsch. Ein Beispiel hierfür ist die Funktion $f: [0, 1] \rightarrow \mathbb{R}$ mit

$$f(x) = \begin{cases} 0 & \text{für } x = 0 \\ \frac{1}{x} & \text{für } x > 0 \end{cases}$$

- (c) Der Satz wird auch für stetige Funktionen falsch, wenn wir kompakte durch offene oder beschränkte Intervalle ersetzen. Betrachte z.B. die Funktion $f: (0, 1) \rightarrow \mathbb{R}$ mit $f(x) = \frac{1}{x}$. Dann ist $f((0, 1))$ unbeschränkt.

Streng monotone Abbildungen sind injektiv. Für stetige Abbildungen auf Intervallen gilt auch die Umkehrung.

Satz 3.2.5. *Sei $I \subset \mathbb{R}$ ein Intervall und $f : I \rightarrow \mathbb{R}$ stetig und injektiv. Dann ist f streng monoton.*

Beweis. Sei f injektiv aber weder streng monoton wachsend noch streng monoton fallend. Dann existieren $a, b, c, d \in I$ mit $a < b, c < d$ und $f(a) < f(b)$ sowie $f(c) > f(d)$. Betrachte die Funktion $g : [0, 1] \rightarrow \mathbb{R}$ mit

$$g(t) = f(a + t(c - a)) - f(b + t(d - b)).$$

Diese Funktion ist wohldefiniert und stetig, denn $a + t(c - a) \in I$ sowie $b + t(d - b) \in I$ sind stetige Funktionen auf $[0, 1]$.

Es gilt: $g(0) = f(a) - f(b) \leq 0$ und $g(1) = f(c) - f(d) \geq 0$. Wegen des Zwischenwertsatzes existiert ein $t_1 \in [0, 1]$ mit

$$0 = g(t_1) = f(a + t_1(c - a)) - f(b + t_1(d - b)).$$

Da f injektiv ist, folgt: $a + t_1(c - a) = b + t_1(d - b)$ und damit $a(1 - t_1) + t_1c = b(1 - t_1) + t_1d$. Auf der anderen Seite erhalten wir dann wegen $a < b$ und $c < d$ mit $a(1 - t_1) + t_1c < b(1 - t_1) + t_1d$ einen Widerspruch. \square

In vielen Fällen ist die Umkehrfunktion einer stetigen Funktion stetig.

Satz 3.2.6. *Sei I ein Intervall und $f : I \rightarrow \mathbb{R}$ eine stetige streng monotone Abbildung. Dann ist $J = f(I)$ ein Intervall und die Umkehrfunktion $f^{-1} : J \rightarrow I$ ist stetig und von gleichem Monotonietyp wie f .*

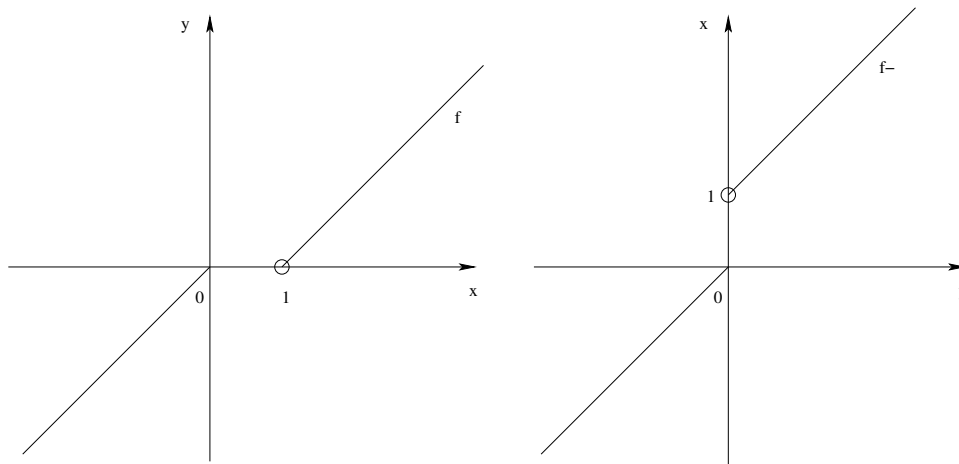
Beweis. Da f stetig ist, ist $J = f(I)$ wegen Korollar 3.2.3 ein Intervall. Sei f monoton steigend (sonst betrachte $-f$), so ist die Umkehrabbildung $f^{-1} : J \rightarrow I$ wegen der Bemerkung nach Definition 1.3.3 ebenfalls monoton steigend.

Sei $(y_n) \in J$ eine konvergente Folge mit $\lim_{n \rightarrow \infty} y_n = y_0 \in J$. Da $(y_n) \in J$ als konvergente Folge beschränkt ist, existieren $A = \inf\{y_n \mid n \in \mathbb{N}\}$ und $B = \sup\{y_n \mid n \in \mathbb{N}\}$ und es gilt: $y_n \in [A, B]$ für alle $n \in \mathbb{N}$. Außerdem sind $A, B \in J$, denn ist $A \notin \{y_n \mid n \in \mathbb{N}\} \subset J$, so ist A HP der Folge $(y_n)_{n \in \mathbb{N}}$, denn zu jedem $\epsilon > 0$ existiert wegen der Definition des Infimums ein $n \in \mathbb{N}$ mit $A < y_n < A + \epsilon$. Dies impliziert: $A = y_0 \in J$. Genauso folgt $B \in J$ und somit ist $[A, B] \subset J$:

Sei $f(a) = A, f(b) = B$. Da f streng monoton steigend ist, ist $a < b$ und mit dem Zwischenwertsatz folgt $f([a, b]) = [A, B]$. Also ist $f^{-1}(y_n) \in [a, b]$ eine beschränkte Folge und besitzt somit eine konvergente Teilfolge. Wir zeigen nun, dass alle ihre konvergenten Teilfolgen $f^{-1}(y_0)$ als Grenzwert besitzen. Als beschränkte Folge ist dann auch $(f^{-1}(y_n))$ konvergent. Sei also $a_n = f^{-1}(y_{k(n)})$ eine konvergente Teilfolge mit Grenzwert a , so ist wegen der Stetigkeit von f auch die Bildfolge $(f(a_n))$ konvergent und es gilt:

$$f(a) = \lim_{n \rightarrow \infty} f(a_n) = \lim_{n \rightarrow \infty} y_{k(n)} = y_0.$$

Da f injektiv ist, ist $a = f^{-1}(y_0)$. Damit ist auch $(f^{-1}(y_n))$ konvergent und $\lim_{n \rightarrow \infty} f^{-1}(y_n) = f^{-1}(y_0)$. \square



Bemerkung. Ist $D \subset \mathbb{R}$ eine beliebige Teilmenge und $f : D \rightarrow f(D)$ eine umkehrbare stetige Abbildung, so muß die Umkehrabbildung i. allg. nicht stetig sein. Die Abbildung $f : (-\infty, 0] \cup (1, \infty) \rightarrow \mathbb{R}$ mit

$$f(x) = \begin{cases} x & \text{für } x \leq 0 \\ x - 1 & \text{für } x > 1 \end{cases}$$

ist streng monoton steigend und stetig. Die Umkehrabbildung ist jedoch nicht stetig.

Aus den Resultaten dieses Abschnittes ergeben sich folgende Eigenschaften sowohl der reellen Exponentialfunktion als auch der reellen trigonometrischen Funktionen.

Korollar 3.2.7. (a) Die Exponentialfunktion $\exp : \mathbb{R} \rightarrow \mathbb{R}_+$ ist streng monoton steigend, bijektiv und stetig. Daher existiert auch ihre Umkehrfunktion

$$\log : \mathbb{R}_+ \rightarrow \mathbb{R}.$$

Sie ist ebenfalls streng monoton steigend, bijektiv und stetig.

(b) Die trigonometrischen Funktionen $\cos : [0, \pi] \rightarrow [-1, 1]$ bzw. $\sin : [-\frac{\pi}{2}, \frac{\pi}{2}] \rightarrow [-1, 1]$ sind streng monoton fallend bzw. streng monoton steigend. Beide Funktionen sind bijektiv und stetig. Daher sind auch ihre Umkehrfunktionen

$$\arccos : [-1, 1] \rightarrow [0, \pi] \quad \text{bzw.} \quad \arcsin : [-1, 1] \rightarrow \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$$

sowohl streng monoton fallend bzw. streng monoton steigend als auch bijektiv und stetig.

(c) Der Tangens $\tan : \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \rightarrow \mathbb{R}$ ist streng monoton steigend, bijektiv und stetig. Insbesondere hat seine Umkehrfunktion $\arctan : \mathbb{R} \rightarrow \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ (Arcus Tangens) die gleichen Eigenschaften.

Beweis. Wie oben bemerkt, sind \exp , \cos , \sin als auch \tan stetige Funktionen. Ihre Monotonieeigenschaften wurden schon im letzten Kapitel bis auf den Tangens $\tan : \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \rightarrow \mathbb{R}$ bewiesen. Die Monotonie des Tangens folgt aus den entsprechenden Eigenschaften von Sinus

und Cosinus. Die Surjektivität der betrachteten Funktionen folgt aus dem Zwischenwertsatz. Betrachtet man z.B. $\exp : \mathbb{R} \rightarrow \mathbb{R}_+$, so existiert zu $y \in \mathbb{R}$ mit $y > 0$ ein $n \in \mathbb{N}$ mit $y \in [e^{-n}, e^n]$. Also folgt aus dem Zwischenwertsatz die Existenz von $c \in [-n, n]$ mit $\exp c = y$. Die Aussagen folgen damit alle aus Satz 3.2.6. \square

Satz 3.2.8. (Polarkoordinaten) *Jede komplexe Zahl $z \in \mathbb{C}$ besitzt die Darstellung*

$$z = re^{i\varphi}$$

mit $r = |z|$ und $\varphi \in \mathbb{R}$. Ist $z \neq 0$ so ist φ bis auf ganzzahlige Vielfache von 2π eindeutig bestimmt.

Beweis. Ist $z = 0$, so gilt $z = 0e^{i\varphi}$ für beliebiges φ . Ist $z \neq 0$, so ist $z = rw$ mit $r = |z|$ und $|w| = 1$. Ist $w = x + iy$, so folgt $x^2 + y^2 = 1$. Da $x \in [-1, +1]$, existiert $\alpha = \arccos(x) \in [0, \pi]$. Also ist $e^{i\alpha} = \cos(\alpha) + i \sin(\alpha) = x + iy$. Ist also $\sin(\alpha) = y$, so setze $\varphi = \alpha$. Andernfalls setze $\varphi = -\alpha$.

Ist $z = re^{i\varphi_1} = re^{i\varphi_2}$, so ist $e^{i(\varphi_1 - \varphi_2)} = 1$ und wegen Korollar 2.6.13 ist $\varphi_2 - \varphi_1 = 2k\pi$ mit $k \in \mathbb{Z}$. \square

Korollar 3.2.9. *Die komplexe Exponentialfunktion*

$$\exp : \mathbb{C} \rightarrow \mathbb{C} \setminus \{0\}$$

ist surjektiv.

Beweis. Ist $z \neq 0$, so existiert ein $y \in \mathbb{R}$ mit $z = |z|e^{iy}$. Wegen der Bijektivität der reellen Exponentialfunktion existiert ein $x \in \mathbb{R}$ mit $e^x = |z|$. Also ist $e^{x+iy} = e^x e^{iy} = z$. \square

3.3 Kompakte Mengen und stetige Funktionen

Wir haben die Nützlichkeit der kompakten Intervalle schon gesehen. Wir wollen nun diesen Begriff auf allgemeine Teilmengen in \mathbb{R} oder \mathbb{C} erweitern. Zunächst benötigen wir den Begriff der Abgeschlossenheit.

Definition 3.3.1. Eine Teilmenge $A \subset \mathbb{R}$ oder $A \subset \mathbb{C}$ heißt *abgeschlossen*, falls jede konvergente Folge (a_n) mit $a_n \in A$ ihren Grenzwert in A hat.

Beispiele.

- (1) Unter den Intervallen sind nur die Intervalle $[a, b]$ und die unbeschränkten Intervalle $[a, \infty)$, $(-\infty, a]$ bzw. $(-\infty, \infty)$ abgeschlossen.
- (2) \mathbb{Z} ist abgeschlossen.
- (3) Die Menge \mathbb{Q} ist nicht abgeschlossen, denn jede reelle Zahl ist Grenzwert einer Folge rationaler Zahlen (dies sieht man z.B. mit Hilfe der Dezimalbruchentwicklung).
- (4) Ist $z_0 \in \mathbb{C}$ und $r > 0$, so ist

$$\bar{B}(z_0, r) = \{z \in \mathbb{C} \mid |z - z_0| \leq r\}$$

eine abgeschlossene Menge. Sie heißt auch *abgeschlossene Kreisscheibe*. Die Abgeschlossenheit folgt aus der Stetigkeit der Abbildung $f : \mathbb{C} \rightarrow \mathbb{R}$ mit $f(z) = |z - z_0|$. Allgemeiner gilt:

(5) Ist $f : \mathbb{C} \rightarrow \mathbb{R}$ eine beliebige stetige Abbildung, so ist für jedes $a \in \mathbb{R}$ die Menge

$$A := \{z \in \mathbb{C} \mid f(z) \leq a\}$$

abgeschlossen. Denn ist $z_n \in A$ eine konvergente Folge mit Grenzwert w , so ist auch $f(z_n)$ eine konvergente Folge mit Grenzwert $f(w)$. Insbesondere ist $f(w) = \lim_{n \rightarrow \infty} f(z_n) \leq a$ und somit ist $w \in A$.

Definition 3.3.2. Eine Teilmenge $K \subset \mathbb{R}$ oder $K \subset \mathbb{C}$ heißt *kompakt*, falls jede Folge in K eine konvergente Teilfolge mit Grenzwert in K besitzt.

Satz 3.3.3. Sei $K \subset \mathbb{C}$ ($K \subset \mathbb{R}$). Dann gilt: K ist genau dann kompakt, falls K abgeschlossen und beschränkt ist.

Beweis. Sei K abgeschlossen und beschränkt und (x_n) eine Folge in K . Da K beschränkt ist, existiert wegen des Satzes von Bolzano-Weierstraß 2.3.5 und 2.4.6 eine konvergente Teilfolge $x_{k(n)} \in K$. Da K abgeschlossen ist, ist ihr Grenzwert in K .

Sei nun K kompakt, so ist K beschränkt, denn sonst existiert zu jedem $n \in \mathbb{N}$ ein $x_n \in K$ mit $|x_n| \geq n$. Diese Folge besitzt aber dann keine konvergente Teilfolge im Widerspruch zur Annahme. Die Menge K ist auch abgeschlossen, denn ist (x_n) eine konvergente Folge mit $x_n \in K$ und Grenzwert a , so ist nach Definition der Kompaktheit $a \in K$. \square

Bemerkung. Das Supremum und Infimum einer kompakten Menge $K \subset \mathbb{R}$ gehören zu K . Denn es existieren Folgen (x_n) mit Grenzwerten $\sup K$ und $\inf K$. Zum Beispiel existiert nach Definition von $\sup K$ zu jedem $n \in \mathbb{N}$ ein $x_n \in K$ mit $\sup K - \frac{1}{n} < x_n \leq \sup K$. Also besitzen kompakte Teilmengen von \mathbb{R} ein Maximum und ein Minimum.

Satz 3.3.4. Sei $K \subset \mathbb{C}$ ($K \subset \mathbb{R}$) eine kompakte Menge und $f : K \rightarrow \mathbb{C}$ eine stetige Funktion. Dann ist $f(K)$ kompakt.

Beweis. Wir zeigen: jede Folge $y_n \in f(K)$ hat eine konvergente Teilfolge mit Grenzwert in $f(K)$.

Sei $y_n = f(x_n)$, wobei $x_n \in K$ eine Folge in K definiert. Wegen der Kompaktheit von K besitzt (x_n) eine konvergente Teilfolge $x_{k(n)}$ mit $\lim_{n \rightarrow \infty} x_{k(n)} = x_0 \in K$. Da f stetig, konvergiert auch die Folge $f(x_{k(n)}) = y_{k(n)}$ mit $\lim_{n \rightarrow \infty} f(x_{k(n)}) = f(x_0) \in f(K)$.

Korollar 3.3.5. Seien $K \subset \mathbb{C}$ ($K \subset \mathbb{R}$) eine kompakte Menge und $f : K \rightarrow \mathbb{R}$ eine stetige Funktion. Dann nimmt f sein Maximum und Minimum an, d.h. es gibt $x_m, x_M \in K$ mit $f(x_m) = \min f(K)$ und $f(x_M) = \max f(K)$.

Beweis. Da wegen Satz 3.3.4 die Menge $f(K) \subset \mathbb{R}$ kompakt ist, enthält sie wegen obiger Bemerkung ihr Minimum und Maximum. \square

Satz 3.3.6. Sei $K \subset \mathbb{C}$ ($K \subset \mathbb{R}$) eine kompakte Menge und $f : K \rightarrow \mathbb{C}$ eine stetige injektive Funktion. Dann besitzt $f : K \rightarrow f(K)$ eine stetige Umkehrfunktion $f^{-1} : f(K) \rightarrow K$.

Beweis. Der Beweis ist analog zu dem Beweis von Satz 3.2.6. Sei (y_n) eine konvergente Folge in $f(K)$ mit $\lim_{n \rightarrow \infty} y_n = y_0$. Wegen der Kompaktheit von K besitzt die Folge $(f^{-1}(y_n))$ eine konvergente Teilfolge. Wir zeigen, dass alle ihre konvergenten Teilfolgen $f^{-1}(y_0)$ einen

gemeinsamen Grenzwert besitzen. Dann ist die Folge $(f^{-1}(y_n))$ als beschränkte Folge konvergent. Ist nun $x_n = f^{-1}(y_{k(n)})$ eine konvergente Teilfolge, so folgt aus der Stetigkeit von f

$$f(\lim_{n \rightarrow \infty} x_n) = \lim_{n \rightarrow \infty} f(x_n) = \lim_{n \rightarrow \infty} y_{k(n)} = y_0 = f(f^{-1}(y_0)).$$

Da f injektiv ist, erhalten wir somit: $\lim_{n \rightarrow \infty} x_n = f^{-1}(y_0)$. \square

Als Anwendung von Korollar 3.3.5 beweisen wir nun den *Fundamentalsatz der Algebra*. Er besagt, dass jedes Polynom vom Grade ≥ 1 eine Nullstelle in \mathbb{C} besitzt. *Komplexe Polynome vom Grade $k \geq 0$* sind Abbildungen mit $p: \mathbb{C} \rightarrow \mathbb{C}$ mit

$$p(z) = a_k z^k + a_{k-1} z^{k-1} + \dots + a_0$$

mit $a_i \in \mathbb{C}$ und $a_k \neq 0$. Ist $a_i \in \mathbb{R}$, so heißt die Abbildung $p: \mathbb{R} \rightarrow \mathbb{R}$ *reelles Polynom*. Der Fundamentalsatz der Algebra lautet:

Satz 3.3.7. *Jedes komplexe Polynom vom Grade ≥ 1 hat eine Nullstelle.*

Bemerkung. Dieser Satz ist natürlich für reelle Polynome falsch. Betrachte zum Beispiel das reelle Polynom $p(x) = x^2 + 1$.

Zunächst zeigt man, dass der Betrag eines Polynoms sein Minimum annimmt.

Lemma 3.3.8. *Sei p ein Polynom. Dann existiert ein $z_0 \in \mathbb{C}$ mit*

$$|p(z_0)| = \inf\{|p(z)| \mid z \in \mathbb{C}\}.$$

Beweis. Zum Beweis dieser Aussage zeigt man die Existenz einer positiven Zahl r mit $|p(z)| > p(0)$ für alle $z \in \mathbb{C} \setminus \bar{B}(0, r)$. Dann ist

$$\inf\{|p(z)| \mid z \in \mathbb{C}\} = \inf\{|p(z)| \mid z \in \bar{B}(0, r)\}.$$

Da $z \mapsto |p(z)|$ eine auf \mathbb{C} stetige Abbildung definiert, nimmt diese Abbildung auf der kompakten Menge $\bar{B}(0, r)$ ihr Infimum an. \square

Beweis des Fundamentalsatzes der Algebra.

Zum Beweis des Fundamentalsatzes der Algebra zeigen wir nun: ist $|p(z_0)| = \inf\{|p(z)| \mid z \in \mathbb{C}\}$, so folgt $p(z_0) = 0$. Angenommen $p(z_0) \neq 0$. Betrachte dann das Polynom

$$q(w) = \frac{1}{p(z_0)} p(z_0 + w).$$

Da $q(0) = 1$, aber q nicht konstant ist, ist q von der Form

$$q(w) = 1 + b_k w^k + b_{k+1} w^{k+1} + \dots + b_{k+l} w^{k+l},$$

wobei $k \in \mathbb{N}$, $l \in \mathbb{N} \cup \{0\}$ und $b_k \neq 0$. Wähle nun $b \in \mathbb{C}$ mit $b^k = -\frac{1}{b_k}$ (Dies ist möglich, denn jede komplexe Zahl ist von der Form $re^{i\varphi}$ mit $r \geq 0$. Hier ist die einzige Stelle, wo verwendet wird, dass das betrachtete Polynom komplex ist). Dann ist $\tilde{q}(w) = q(bw)$ von der Form

$$\tilde{q}(w) = 1 - w^k + w^{k+1} s(w)$$

für ein weiteres Polynom s . Es existiert eine Zahl $r > 0$ mit $|s(w)| \leq r$ für alle $w \in \mathbb{C}$ mit $|w| \leq 1$. Dann folgt für alle $|w| \leq 1$:

$$|\tilde{q}(w)| \leq |1 - w^k| + r|w^{k+1}|$$

Ist insbesondere $t \in (0, 1]$ so folgt:

$$|\tilde{q}(t)| \leq 1 - t^k + rt^{k+1} = 1 - t^k(1 - rt) < 1$$

falls $rt < 1$. Für ein solches t gilt daher:

$$|\tilde{q}(t)| = |q(bt)| = \left| \frac{p(z_0 + bt)}{p(z_0)} \right| < 1$$

und somit erhalten wir mit $|p(z_0 + bt)| < |p(z_0)|$ einen Widerspruch zur Definition von z_0 . Damit ist die Annahme $p(z_0) \neq 0$ falsch. \square

Als Anwendung erhält man, dass komplexe Polynome sich in Linearfaktoren zerlegen lassen.

Lemma 3.3.9. *Sei p ein Polynom vom Grade $n \geq 1$ mit*

$$p(z) = a_n z^n + a_{n-1} z^{n-1} + \dots + a_0$$

und $w \in \mathbb{C}$ eine Nullstelle von p . Dann existiert ein Polynom q vom Grade $n - 1$ mit

$$p(z) = (z - w)q(z).$$

Beweis. Schreibe wie im Beweis von Satz 2.7.9:

$$(z - w)(z^{k-1} + z^{k-2}w + z^{k-3}w^2 + \dots + w^{k-1}) = z^k - w^k,$$

es folgt:

$$p(z) - p(w) = \sum_{k=1}^n a_k (z^k - w^k) = (z - w) \cdot \sum_{k=1}^n a_k (z^{k-1} + z^{k-2}w + z^{k-3}w^2 + \dots + w^{k-1}).$$

Da w eine Nullstelle von p ist, folgt

$$p(z) = (z - w)q(z)$$

mit $q(z) := \sum_{k=1}^n a_k (z^{k-1} + z^{k-2}w + z^{k-3}w^2 + \dots + w^{k-1})$. \square

Zusammen mit dem Fundamentalsatz der Algebra erhalten wir.

Satz 3.3.10. (Zerlegung in Linearfaktoren)

Sei p ein komplexes Polynom vom Grade $n \geq 1$. Dann existieren $m \leq n$ verschiedene Nullstellen z_1, \dots, z_k mit

$$p(z) = c(z - z_1)^{k_1}(z - z_2)^{k_2} \dots (z - z_m)^{k_m}$$

mit $k_i \in \mathbb{N}$ und $k_1 + k_2 + \dots + k_m = n$ und $c \neq 0$. Dabei heißt k_i die Vielfachheit der Nullstelle z_i .

3.4 Stetige Fortsetzbarkeit und Grenzwerte von Funktionen

Sei $f : D \rightarrow \mathbb{C}$ eine stetige Funktion und $a \in \mathbb{C} \setminus D$. Existiert eine stetige Funktion $g : D \cup \{a\} \rightarrow \mathbb{R}$ mit $g(x) = f(x) \forall x \in D$?

Definition 3.4.1. Sei $A \subset \mathbb{C}$ und $a \in \mathbb{C}$. Dann heißt a *Berührungspunkt* von A , falls für jedes $\epsilon > 0$ gilt: $B(a, \epsilon) \cap A \neq \emptyset$.

Bemerkungen.

1. $a \in \mathbb{C}$ ist genau dann ein Berührungspunkt von A , falls eine Folge $a_n \in A$ existiert mit $\lim_{n \rightarrow \infty} a_n = a$. Denn ist $a \in \mathbb{C}$, so existiert zu jedem $\epsilon = \frac{1}{n}$ mit $n \in \mathbb{N}$ ein $a_n \in B(a, \frac{1}{n}) \cap A$. Die Folge (a_n) konvergiert dann gegen a . Ist umgekehrt eine Folge in A mit Grenzwert $a \in \mathbb{C}$ gegeben, so ist für jedes $\epsilon > 0$ die Menge $B(a, \epsilon) \cap A$ nicht leer.
2. Jeder Punkt von A ist Berührungspunkt. Berührungspunkte von A müssen nicht zu A gehören. Für $A = (0, 1)$ oder $A = \{\frac{1}{n} \mid n \in \mathbb{N}\}$ ist 0 Berührungspunkt, der nicht zu M gehört.
3. Ist $A \subset \mathbb{C}$, so heißt

$$\bar{A} := \{z \mid z \text{ ist Berührungspunkt von } A\}$$

abgeschlossene Hülle von A . Die Menge \bar{A} ist abgeschlossen und enthält A . Sie ist sogar die kleinste abgeschlossene Menge, die A enthält, d.h. jede abgeschlossene Menge, die A enthält, enthält auch \bar{A} . Insbesondere ist A genau dann abgeschlossen, falls A mit ihrer abgeschlossenen Hülle \bar{A} übereinstimmt, d.h. wenn A alle ihre Berührungspunkte enthält.

4. Sei $f : D \rightarrow \mathbb{C}$ eine stetige Funktion und $a \in \mathbb{C} \setminus D$. Ist a kein Berührungspunkt von D , d.h. existiert ein $\delta > 0$ mit $B(a, \delta) \cap D = \emptyset$. Dann ist jede *Fortsetzung* $g : D \cup \{a\} \rightarrow \mathbb{C}$ von f auf die Menge $D \cup \{a\}$ stetig, d.h. für jedes $b \in \mathbb{C}$ ist

$$g(x) = \begin{cases} f(x) & \text{für } x \in D \\ b & \text{für } x = a \end{cases}$$

stetig. Denn für jedes $\epsilon > 0$ folgt

$$g(B(a, \delta) \cap (D \cup \{a\})) = \{g(a)\} = \{b\} \subset B(b, \epsilon),$$

falls $B(a, \delta) \cap D = \emptyset$.

Ist dagegen a ein Berührungspunkt von D , so existiert i. allg. keine stetige Fortsetzung. Als Beispiel betrachte die Funktion

$$f : \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R} \quad \text{mit} \quad f(x) = \frac{1}{x}.$$

Diese Funktion ist nicht stetig auf \mathbb{R} fortsetzbar. Denn wäre $g : \mathbb{R} \rightarrow \mathbb{R}$ eine Fortsetzung, so wäre $g([0, 1])$ beschränkt.

Ist $a \in \mathbb{C} \setminus D$ Berührungspunkt von $f : D \rightarrow \mathbb{C}$, so existiert, wie der folgende Satz zeigt, höchstens eine stetige Fortsetzung von f .

Satz 3.4.2. *Es sei $f : D \rightarrow \mathbb{C}$ stetig und $a \in \mathbb{C} \setminus D$ Berührungspunkt von D . Existiert eine stetige Fortsetzung $g : D \cup \{a\} \rightarrow \mathbb{C}$, so ist sie eindeutig bestimmt.*

Beweis. Es seien $g, g' : D \cup \{a\} \rightarrow \mathbb{C}$ zwei stetige Fortsetzungen. Dann ist die Funktion $h := g - g' : D \cup \{a\} \rightarrow \mathbb{C}$ stetig mit

$$h(x) = \begin{cases} 0 & \text{für } x \in D \\ c & \text{für } x = a \end{cases}$$

und $c = g(a) - g'(a)$. Da h stetig in a ist, existiert für jedes $\epsilon > 0$ ein $\delta > 0$ mit

$$h(B(a, \delta) \cap D) \cup \{a\} \subset B(c, \epsilon).$$

Sei $x \in B(a, \delta) \cap D$, so ist $h(x) = 0 \in B(c, \epsilon)$. Dies impliziert für alle $\epsilon > 0 : |c| = |c - 0| < \epsilon$. Insbesondere ist $c = 0$. \square

In engem Zusammenhang mit der Stetigkeit und stetigen Fortsetzbarkeit steht der folgende Begriff.

Definition 3.4.3. Sei $f : D \rightarrow \mathbb{C}$ und a Berührungspunkt von D . Die Zahl $b \in \mathbb{C}$ heißt *Grenzwert von f in a* , falls für jedes $\epsilon > 0$ ein $\delta > 0$ existiert mit $f(B(a, \delta) \cap D) \subset B(b, \epsilon)$. Wir schreiben

$$\lim_{\substack{x \rightarrow a \\ x \in D}} f(x) = b,$$

falls b der Grenzwert von f in a ist.

Bemerkungen.

(a) Ist $a \in D$, so ist $b = f(a)$, denn $f(a) \in B(b, \epsilon)$ für alle $\epsilon > 0$. Dann ist $\lim_{\substack{x \rightarrow a \\ x \in D}} f(x) = f(a)$

äquivalent zur Stetigkeit von f in a .

Ist a Berührungspunkt von D der nicht zu D gehört, so ist $\lim_{\substack{x \rightarrow a \\ x \in D}} f(x) = b$ äquivalent dazu,

dass die Funktion $g : D \cup \{a\} \rightarrow \mathbb{C}$ mit

$$g(x) = \begin{cases} f(x) & \text{für } x \in D \\ b & \text{für } x = a \end{cases}$$

stetig in a ist. Insbesondere lässt sich f stetig in a fortsetzen.

(b) Wenn klar ist, auf welchen Definitionsbereich der Grenzwert einer Funktion sich bezieht, so schreiben wir auch $\lim_{x \rightarrow a} f(x)$ statt $\lim_{\substack{x \rightarrow a \\ x \in D}} f(x)$.

Grenzwerte von Funktionen lassen sich auch mit Hilfe von Folgen charakterisieren.

Satz 3.4.4. Sei $f : D \rightarrow \mathbb{C}$ eine Funktion und a Berührungspunkt von D . Dann gilt:

$$\lim_{\substack{x \rightarrow a \\ x \in D}} f(x) = b \Leftrightarrow \text{für alle Folgen } x_n \in D \text{ und } \lim_{n \rightarrow \infty} x_n = a \text{ folgt: } \lim_{n \rightarrow \infty} f(x_n) = b.$$

Beweis. Dies ist eine Konsequenz aus dem Folgenkriterium für stetige Funktionen (siehe Satz 3.1.4). \square

Beispiel. Sei $D = \mathbb{R}_+ \setminus \{1\}$ und $f(x) = \frac{\sqrt{x}-1}{x-1}$. Es gilt: $\lim_{x \rightarrow 1} \frac{\sqrt{x}-1}{x-1} = \frac{1}{2}$, denn

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

Insbesondere ist $g(x) = \frac{1}{\sqrt{x}+1}$ stetige Fortsetzung von f .

Aus Satz 3.1.4 für stetige Funktionen ergeben sich folgende Rechenregeln für Grenzwerte von Funktionen.

Satz 3.4.5. (Rechenregeln für Grenzwerte von Funktionen)

Seien $f : D \rightarrow \mathbb{C}$ und $g : D \rightarrow \mathbb{C}$ Funktionen. Sei a Berührungspunkt von D und nehme an, dass die Grenzwerte

$$\lim_{\substack{x \rightarrow a \\ x \in D}} f(x) \quad \text{und} \quad \lim_{\substack{x \rightarrow a \\ x \in D}} g(x)$$

existieren. Dann gilt:

(a) Die Summe $(f+g) : D \rightarrow \mathbb{C}$ hat einen Grenzwert in a mit

$$\lim_{\substack{x \rightarrow a \\ x \in D}} (f+g)(x) = \lim_{\substack{x \rightarrow a \\ x \in D}} f(x) + \lim_{\substack{x \rightarrow a \\ x \in D}} g(x).$$

(b) Das Produkt $(fg) : D \rightarrow \mathbb{C}$ hat einen Grenzwert in a mit

$$\lim_{\substack{x \rightarrow a \\ x \in D}} (fg)(x) = \left(\lim_{\substack{x \rightarrow a \\ x \in D}} f(x) \right) \left(\lim_{\substack{x \rightarrow a \\ x \in D}} g(x) \right)$$

(c) Ist $\lim_{\substack{x \rightarrow a \\ x \in D}} g(x) \neq 0$, so existiert ein $\delta > 0$ mit $g(x) \neq 0$ für alle $x \in B(a, \delta) \cap D = D'$.

Dann hat der Quotient $\left(\frac{f}{g}\right) : D' \rightarrow \mathbb{C}$ einen Grenzwert in a mit

$$\lim_{\substack{x \rightarrow a \\ x \in D'}} \left(\frac{f}{g}\right)(x) = \frac{\lim_{\substack{x \rightarrow a \\ x \in D'}} f(x)}{\lim_{\substack{x \rightarrow a \\ x \in D'}} g(x)}.$$

Wir wollen noch ein Kriterium für die Existenz des Grenzwertes einer Funktion angeben, das nicht die Kenntnis des Grenzwertes verlangt. Es beruht auf dem Cauchy-Kriterium für Folgen.

Satz 3.4.6. Sei $f : D \rightarrow \mathbb{C}$ eine Funktion und $a \in D$ ein Berührungspunkt. Dann sind folgende Aussagen äquivalent:

(1) f hat in a einen Grenzwert.

(2) $\forall \epsilon > 0 \quad \exists \delta > 0$ mit $|f(x) - f(y)| < \epsilon$ für alle $x, y \in B(a, \delta) \cap D$.

Beweis. Es gelte (1), d.h. f habe in a den Grenzwert b . Sei $\epsilon > 0$, so existiert ein $\delta > 0$ mit $f(B(a, \delta) \cap D) \subset B(b, \epsilon/2)$.

Sind $x, y \in B(a, \delta) \cap D$, so folgt:

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

Sei nun (2) erfüllt und (x_n) eine Folge mit $x_n \in D$ und $\lim_{n \rightarrow \infty} x_n = a$.

Dann definiert $f(x_n)$ eine Cauchyfolge, denn ist $\epsilon > 0$, so betrachte $\delta > 0$ wie in (2). Da (x_n) gegen a konvergiert, existiert ein $n_0 \in \mathbb{N}$ mit $x_n \in B(a, \delta) \cap D$ für $n \geq n_0$. Aus (2) folgt: $|f(x_n) - f(x_m)| < \epsilon$ für $n, m \geq n_0$. Da Cauchyfolgen konvergieren, existiert $\lim_{n \rightarrow \infty} f(x_n) = b$.

Der Grenzwert b hängt nicht von der Wahl der Folge x_n ab. Denn ist x'_n eine andere Folge mit $\lim_{n \rightarrow \infty} x'_n = a$, so ist $x_1, x'_1, x_2, x'_2, \dots$ eine Folge mit Grenzwert a . Insbesondere ist $f(x_1), f(x'_1), f(x_2), f(x'_2), \dots$ eine Cauchyfolge. Da diese Folge konvergiert, konvergieren auch die Teilfolgen $(f(x_n))$ und $(f(x'_n))$ und haben den gleichen Grenzwert b . Wegen obiger Bemerkung ist b der Grenzwert von f in a . \square

Nun können wir ein Kriterium für die stetige Fortsetzbarkeit einer stetigen Funktion $f : D \rightarrow \mathbb{R}$ auf die abgeschlossene Hülle \overline{D} angeben. Dazu brauchen wir noch den Begriff der gleichmäßigen Stetigkeit.

Definition 3.4.7. Eine Funktion $f : D \rightarrow \mathbb{R}$ heißt *gleichmäßig stetig*, falls zu jedem $\epsilon > 0$ ein $\delta > 0$ existiert, so dass für alle $x, y \in D$ gilt:

$$|x - y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon.$$

Bemerkung. Gleichmäßig stetige Funktionen sind stetig, denn hält man $y = a$ fest, so ergibt sich aus Definition 3.1.1 die Stetigkeit im Punkte a . Aus der Stetigkeit folgt hingegen nicht die gleichmäßige Stetigkeit. Dies liegt daran, dass in der Definition der Stetigkeit von f in $a \in D$, die positive Zahl δ i. allg. von a abhängt.

Beispiel.

- (1) Sei $D = (0, 1]$ und $f(x) = \frac{1}{x}$. Dann ist f stetig, aber nicht gleichmäßig stetig. Eine Funktion $f : D \rightarrow \mathbb{R}$ ist nicht gleichmäßig stetig, falls es ein $\epsilon > 0$ gibt, so dass für jedes $\delta > 0$ Zahlen $x, y \in D$ mit $|x - y| < \delta$ und $|f(x) - f(y)| \geq \epsilon$ existieren. Betrachte nun

$$|f(x) - f(y)| = \left| \frac{1}{x} - \frac{1}{y} \right| = \frac{|y - x|}{xy}.$$

Sei $\delta > 0$, so wähle $x = \frac{1}{n} < \delta$ und $y = 2x$. Dann folgt: $|x - y| = \frac{1}{n} < \delta$ und

$$\frac{|y - x|}{x \cdot y} = \frac{x}{2x^2} = \frac{1}{2x} = \frac{n}{2} \geq \frac{1}{2}.$$

Also ist f nicht gleichmäßig stetig.

- (2) Hingegen sind Lipschitz-stetige Funktionen $f : D \rightarrow \mathbb{C}$ gleichmäßig stetig, denn ist $|f(x) - f(y)| \leq L|x - y|$, und $\epsilon > 0$, so wähle $\delta = \frac{\epsilon}{L}$.
Ist $|x - y| < \delta = \frac{\epsilon}{L}$, so ist $|f(x) - f(y)| < \epsilon$.

Satz 3.4.8. Ist $f : D \rightarrow \mathbb{C}$ eine gleichmäßig stetige Funktion, so existiert eine stetige Fortsetzung auf die abgeschlossene Hülle \overline{D} .

Beweis. Sei $a \in \overline{D}$, d.h. es sei a Berührungspunkt von D . Wegen der gleichmäßigen Stetigkeit existiert für jedes $\epsilon > 0$ ein $\delta > 0$ mit $|f(x) - f(y)| < \epsilon$, falls $x, y \in D$ und $|x - y| < \delta$.

Aus Satz 3.4.6 folgt, dass f einen Grenzwert in a besitzt. Mit Satz 3.4.4 erhalten wir dann, dass sich f stetig nach a fortsetzen läßt. \square

Bemerkung. Es ist einfach zu zeigen, dass die stetige Fortsetzung ebenfalls gleichmäßig stetig ist.

Wir haben gesehen, dass stetige Abbildungen i. allg. nicht gleichmäßig stetig sind. Stetige Funktionen auf kompakten Mengen hingegen sind, wie der folgende Satz zeigt, immer gleichmäßig stetig.

Satz 3.4.9. Sei $K \subset \mathbb{C}$ kompakt und $f: K \rightarrow \mathbb{C}$ stetig. Dann ist f auch gleichmäßig stetig.

Beweis. Wir beweisen dies indirekt, d.h. wir zeigen: Ist f nicht gleichmäßig stetig, so ist f nicht stetig. Sei f nicht gleichmäßig stetig. Dann gibt es ein $\epsilon > 0$, so dass für jedes $n \in \mathbb{N}$ und $\delta = \frac{1}{n}$ Zahlen $x_n, y_n \in K$ mit $|x_n - y_n| < \frac{1}{n}$ und $|f(x_n) - f(y_n)| \geq \epsilon$ existieren. Da K kompakt ist, besitzt x_n eine konvergente Teilfolge $x_{k(n)}$ mit $\lim_{n \rightarrow \infty} x_{k(n)} = a \in K$. Dann konvergiert aber auch die Teilfolge $y_{k(n)}$ mit $\lim_{k \rightarrow \infty} y_{k(n)} = a$.

Da $|f(x_{k(n)}) - f(y_{k(n)})| \geq \epsilon$, können nicht die beiden Folgen $(f(x_{k(n)}))$ und $(f(y_{k(n)}))$ gegen $f(a)$ konvergieren. Dann ist f aber nicht stetig in $a \in K$. \square

Seien $f: D \rightarrow \mathbb{C}$ und a Berührungspunkt einer Teilmenge $D' \subset D$. Dann kann natürlich der Grenzwert

$$\lim_{\substack{x \rightarrow a \\ x \in D'}} f(x)$$

existieren ohne dass der Grenzwert

$$\lim_{\substack{x \rightarrow a \\ x \in D}} f(x)$$

existiert. Ein Beispiel hierfür ist die Funktion (*Heaviside-Funktion*) $f: \mathbb{R} \rightarrow \mathbb{R}$ mit

$$f(x) = \begin{cases} 0 & \text{für } x \leq 0 \\ 1 & \text{für } x > 0 \end{cases}.$$

0 ist in $\mathbb{R}_- := \{x \in \mathbb{R} \mid x \leq 0\}$. Der Grenzwert $\lim_{\substack{x \rightarrow 0 \\ x \in \mathbb{R}_-}} f(x)$ existiert und ist 0 während der

Grenzwert $\lim_{\substack{x \rightarrow 0 \\ x \in \mathbb{R}}} f(x)$ nicht existiert.

Definition 3.4.10. (*Einseitige Grenzwerte*)

Sei $D \subset \mathbb{R}$ und $f: D \rightarrow \mathbb{C}$ eine Funktion. Sei a ein Berührungspunkt von $D^- := D \cap (-\infty, a)$ bzw. von $D^+ := D \cap (a, \infty)$. Dann sagt man, dass der *linksseitige* bzw. *rechtsseitige Grenzwert* existiert, falls der Grenzwert

$$\lim_{\substack{x \rightarrow a \\ x \in D^-}} f(x) \quad \text{bzw.} \quad \lim_{\substack{x \rightarrow a \\ x \in D^+}} f(x)$$

existiert. Für den linksseitigen Grenzwert bzw. rechtsseitigen Grenzwert schreibt man auch

$$\lim_{x \nearrow a} f(x) = f(a_-) \quad \text{bzw.} \quad \lim_{x \searrow a} f(x) = f(a_+).$$

Gehört a zu D , so heißt f in a *linksseitig stetig* (*rechtsseitig stetig*), falls der linksseitige Grenzwert existiert und $f(a) = f(a_-)$ ($f(a) = f(a_+)$) gilt.

Beispiel. Die Heaviside-Funktion ist in 0 linksseitig aber nicht rechtsseitig stetig.

3.5 Uneigentliche Grenzwerte

Für gewissen Anwendungen ist es nützlich den Grenzwertbegriff zu verallgemeinern. Wir betrachten dabei nur Funktionen, deren Definitionsbereich Teilmenge der reellen Zahlen ist. Zunächst wollen wir den Grenzwert einer Funktion in ∞ erklären.

Definition 3.5.1. Sei $f : D \rightarrow \mathbb{C}$ eine Funktion mit nach oben unbeschränktem Definitionsbereich $D \subset \mathbb{C}$. Dann heißt $b \in \mathbb{C}$ der *Grenzwert von f in ∞* , falls für alle $\epsilon > 0$ eine Zahl $N \in \mathbb{R}$ existiert mit

$$|f(x) - b| < \epsilon \quad \text{für alle } x \in D \text{ und } x > N.$$

Existiert dieser Grenzwert, so schreibt man $b = \lim_{x \rightarrow \infty} f(x)$ oder auch $f(x) \rightarrow b$ für $x \rightarrow \infty$.

Ist D nicht nach unten beschränkt, so heißt $b \in \mathbb{C}$ der *Grenzwert von f in $-\infty$* , falls für alle $\epsilon > 0$ eine Zahl $N \in \mathbb{R}$ gibt mit

$$|f(x) - b| < \epsilon \quad \text{für alle } x \in D \text{ mit } x < N.$$

Falls der Grenzwert existiert, schreibt man $b = \lim_{x \rightarrow -\infty} f(x)$ oder auch $f(x) \rightarrow b$ für $x \rightarrow -\infty$.

Bemerkung. Diese Definition erweitert den Begriff des Grenzwertes einer Folge (siehe Definition 2.1.3), denn eine Folge $(a_n)_{n \in \mathbb{N}}$ ist nichts anderes als eine Funktion $a : \mathbb{N} \rightarrow \mathbb{C}$ mit $a(n) = a_n$. Die Folge $(a_n)_{n \in \mathbb{N}}$ konvergiert genau dann gegen b , wenn der Grenzwert der Funktion $a : \mathbb{N} \rightarrow \mathbb{C}$ in ∞ existiert und mit b übereinstimmt.

Beispiele.

1. Betrachte die Funktion $f : \mathbb{R}_+ \rightarrow \mathbb{R}$ mit $f(x) = \frac{1}{x}$. Dann existiert zu jedem $\epsilon > 0$ ein $N \in \mathbb{R}$ mit $|\frac{1}{N}| < \epsilon$ (siehe Satz 2.1.5). Also ist auch $|f(x)| < \epsilon$ für $x > N$ und somit folgt $0 = \lim_{x \rightarrow \infty} f(x)$.

2. Betrachte die Funktion $f : \mathbb{R}_+ \rightarrow \mathbb{R}$ mit $f(x) = \sqrt{x+1} - \sqrt{x}$. Dann gilt $0 = \lim_{x \rightarrow \infty} f(x)$, denn

$$|f(x)| = \sqrt{x+1} - \sqrt{x} = \frac{1}{\sqrt{x+1} + \sqrt{x}} < \frac{1}{2\sqrt{x}}.$$

Ist $\epsilon > 0$ gegeben, so definiere $N := \frac{1}{4\epsilon^2}$. Dann gilt für alle $x > N$:

$$|f(x)| < \frac{1}{2\sqrt{x}} < \frac{1}{2\sqrt{N}} = \epsilon.$$

Bemerkung. Die in Satz 3.4.5 formulierten Rechenregeln übertragen sich auf für Grenzwerte von Funktionen in ∞ .

Nun wollen wir noch definieren, was wir unter dem ‘Konvergieren’ einer Funktion gegen unendlich verstehen wollen. Um diese Begriffsbildung von dem Begriff der Konvergenz gegen reelle oder komplexe Zahlen abzugrenzen, werden wir von *uneigentlichen Grenzwerten* sprechen.

Definition 3.5.2. Seien $D \subset \mathbb{R}$ und a ein Berührungspunkt von \mathbb{R} . Sei $f : D \rightarrow \mathbb{R}$ eine Funktion, so sagen wir, dass f in a den *uneigentlichen Grenzwert* ∞ ($-\infty$) besitzt, falls zu jedem $E \in \mathbb{R}$ ein $\delta > 0$ existiert mit

$$f(x) > E \quad (f(x) < E)$$

für alle $x \in B(a, \delta) \cap D$. Wir schreiben im Falle der Existenz des uneigentlichen Grenzwertes:

$$\lim_{x \rightarrow a} f(x) = \infty \quad (\lim_{x \rightarrow a} f(x) = -\infty).$$

Ist $D \subset \mathbb{R}$ nicht nach oben beschränkt, so sagen wir, dass f in ∞ den uneigentlichen Grenzwert ∞ ($-\infty$) besitzt, falls zu jedem $E \in \mathbb{R}$ ein N existiert mit

$$f(x) > E \quad (f(x) < E)$$

für alle $x > N$. Im Falle der Existenz des Grenzwertes schreiben wir dann

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

Entsprechendes lässt sich formulieren, wenn $D \subset \mathbb{R}$ nicht nach unten beschränkt ist.

Beispiel. Ist $f : \mathbb{R} \rightarrow \mathbb{R}$ eine monoton steigende und nicht nach oben beschränkte Funktion, so besitzt f einen uneigentlichen Grenzwert in ∞ und es gilt:

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

Denn ist $E \in \mathbb{R}$, so existiert (da f unbeschränkt ist) ein $N \in \mathbb{R}$ mit $f(N) > E$. Da f monoton steigend ist, ist somit $f(x) > E$ für $x > N$. Ist f nicht nach unten beschränkt, so folgt genauso:

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

Kapitel 4

Differentialrechnung

4.1 Differenzierbarkeit von Funktionen

Im folgenden sei der Definitionsbereich der betrachteten Funktionen, falls nichts Anderes gesagt wird, immer eine Teilmenge der reellen Zahlen. Sei $f : D \rightarrow \mathbb{R}$ eine Funktion und $a \in D$ sei Berührungspunkt von $D \setminus \{a\}$ (solche Punkte werden auch *Häufungspunkte* von D genannt. Ist D ein Intervall oder eine offene Menge, so ist jedes $x \in D$ auch Häufungspunkt). Die Funktion $q : D \setminus \{a\} \rightarrow \mathbb{R}$ mit

$$q(x) = \frac{f(x) - f(a)}{x - a}$$

heißt *Differenzenquotient* von f in a . Der Differenzenquotient beschreibt für jedes $x \in D$ die Steigung der Geraden durch $f(a)$ und $f(x)$ (s. Abb. 4.1). Läßt sich nun der Differenzenquotient stetig nach a fortsetzen, so nennen wir f in a *differenzierbar*.

Definition 4.1.1. Sei $f : D \rightarrow \mathbb{R}$ eine Funktion und $a \in D$ Häufungspunkt von D . f heißt genau dann *in a differenzierbar*, falls

$$\lim_{\substack{x \rightarrow a \\ x \in D \setminus \{a\}}} \frac{f(x) - f(a)}{x - a} = m$$

existiert. Die Zahl m heißt die *Ableitung* von f in a . Man schreibt: $f'(a) := m$. Ist jeder Punkt $a \in D$ Häufungspunkt von D und ist f differenzierbar für alle $a \in D$, so heißt f *differenzierbar*.

Bemerkungen.

- (1) Die Funktion $f : D \rightarrow \mathbb{R}$ ist also genau dann differenzierbar in $a \in D$, falls sich der Differenzenquotient $q(x) = \frac{f(x) - f(a)}{x - a}$ stetig nach a fortsetzen lässt, d.h. es existiert ein $m \in \mathbb{R}$, so dass

$$Q(x) = \begin{cases} q(x) & \text{für } x \in D \setminus \{a\} \\ m & \text{für } x = a \end{cases}$$

stetig in a ist. Geometrisch beschreibt der Wert m die Steigung der Funktion f im Punkte a . Die Zahl m ist, falls sie existiert, wegen Satz 3.4.2 eindeutig definiert, denn a ist Berührungspunkt von $D \setminus \{a\}$.

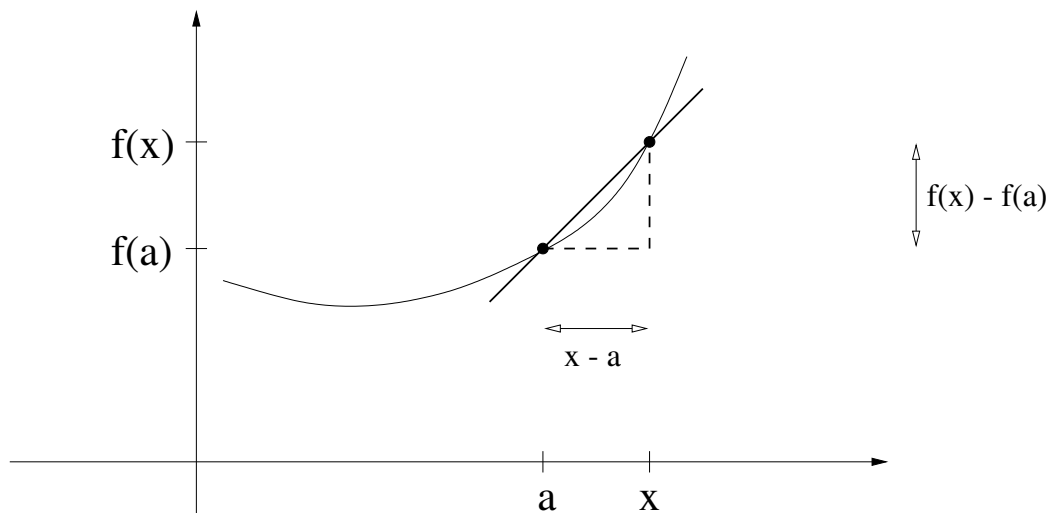


Abbildung 4.1: Graphische Darstellung des Differenzenquotienten

- (2) Die Funktion f ist genau dann in a differenzierbar, falls für jede konvergente Folge $x_n \in D \setminus \{a\}$ mit $\lim_{n \rightarrow \infty} x_n = a$ auch die Folge $q(x_n)$ konvergiert. Diese Bedingung impliziert, dass der Grenzwert der Folge $q(x_n)$ nicht von der Wahl der gegen a konvergierenden Folge $x_n \in D$ abhängt.

Beispiel.

- (1) Betrachte $f: \mathbb{R} \rightarrow \mathbb{R}$ mit $f(x) = x^n$ und $n \in \mathbb{N}$. Sei $a \in \mathbb{R}$ und $q: \mathbb{R} \setminus \{a\} \rightarrow \mathbb{R}$ mit

$$q(x) = \frac{f(x) - f(a)}{x - a}$$

der Differenzenquotient von f in a . Dann folgt:

$$q(x) = \frac{x^n - a^n}{x - a} = x^{n-1} + x^{n-2}a + x^{n-3}a^2 + \dots + a^{n-1}.$$

Insbesondere lässt sich die Funktion q stetig nach a fortsetzen mit $\lim_{x \rightarrow a} q(x) = n a^{n-1} = f'(a)$.

- (2) Betrachte die Exponentialfunktion

$$\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

- (a) Diese Funktion ist in 0 differenzierbar, denn für $x \neq 0$ gilt:

$$q(x) = \frac{\exp(x) - 1}{x} = \frac{x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots}{x} = 1 + \frac{x}{2!} + \frac{x^2}{3!} + \dots =: Q(x).$$

Insbesondere stellt $Q(x)$ eine Potenzreihe dar, die für jedes $x \in \mathbb{R}$ konvergiert. Da Potenzreihen stetig sind, folgt: $\lim_{x \rightarrow 0} q(x) = \lim_{x \rightarrow 0} Q(x) = Q(0) = 1$, d.h. $\exp'(0) = 1 = \exp(0)$.

(b) Ist $a \in \mathbb{R}$ beliebig, so betrachte für $x \neq a$

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

Für $x = a + h$ gilt:

$$\begin{aligned} q(a+h) &= \frac{\exp(a+h) - \exp(a)}{h} \\ &= \frac{\exp(a) \cdot \exp(h) - \exp(a)}{h} = \exp(a) \frac{\exp(h) - 1}{h}. \end{aligned}$$

Somit ergibt sich $\lim_{x \rightarrow a} q(x) = \lim_{h \rightarrow 0} q(a+h) = \exp(a)$, d.h. $\exp'(a) = \exp(a)$.

(3) Betrachte die Sinusfunktion $\sin : \mathbb{R} \rightarrow \mathbb{R}$, die definiert durch die Potenzreihe ist:

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} + \dots$$

Aus Satz 2.6.9 folgt

$$\sin(x) - \sin(a) = 2 \cos\left(\frac{x+a}{2}\right) \sin\left(\frac{x-a}{2}\right).$$

Damit folgt für $x \neq a$

$$q(x) = \frac{\sin(x) - \sin(a)}{x - a} = \cos\left(\frac{x+a}{2}\right) \frac{\sin\left(\frac{x-a}{2}\right)}{\left(\frac{x-a}{2}\right)}.$$

Außerdem gilt für $h \in (-4, 4)$ wegen Satz 2.6.9:

$$\sin(h) = h + R_3(h)$$

mit $|R_3(h)| \leq \frac{|h|^3}{3!}$. Also ist

$$\left| \frac{\sin(h)}{h} - 1 \right| = \frac{|R_3(h)|}{|h|} \leq \frac{h^2}{3!}$$

und somit

$$\lim_{x \rightarrow a} \frac{\sin\left(\frac{x-a}{2}\right)}{\left(\frac{x-a}{2}\right)} = 1.$$

Da der Cosinus stetig ist, folgt aus den Rechenregeln für die Grenzwerte von Funktionen:

$$\sin'(a) = \lim_{x \rightarrow a} \frac{\sin(x) - \sin(a)}{x - a} = \lim_{x \rightarrow a} \cos\left(\frac{x+a}{2}\right) = \cos(a)$$

Wegen

$$\cos(x) - \cos(a) = -2 \sin\left(\frac{x+a}{2}\right) \sin\left(\frac{x-a}{2}\right)$$

schließt man mit analogen Argumenten

$$\cos'(a) = \lim_{x \rightarrow a} \frac{\cos(x) - \cos(a)}{x - a} = -\sin(a).$$

Differenzierbare Funktionen lassen sich dadurch kennzeichnen, dass sie sich gut durch *affin lineare Funktionen* approximieren lassen.

Satz 4.1.2. Sei $f: D \rightarrow \mathbb{R}$ eine Funktion und $a \in D$ ein Häufungspunkt. Dann sind folgende Aussagen äquivalent:

(1) f ist differenzierbar in a .

(2) Es existieren eine Zahl $m \in \mathbb{R}$ und eine in a stetige Funktion $r: D \rightarrow \mathbb{R}$ mit $r(a) = 0$, so dass

$$f(x) = f(a) + m(x - a) + r(x)(x - a).$$

Beweis. Sei $q(x) = \frac{f(x)-f(a)}{x-a}$ der Differenzenquotient von f in a . Nach Definition 4.1.1 ist die Funktion $f: D \rightarrow \mathbb{R}$ genau dann in a differenzierbar, falls q sich in $x = a$ stetig fortsetzen lässt, d.h. falls ein $m \in \mathbb{R}$ existiert, so dass

$$Q(x) = \begin{cases} q(x) & \text{für } x \neq a \\ m & \text{für } x = a \end{cases}$$

stetig ist. Für $x \neq a$ betrachte:

$$\frac{f(x) - f(a) - m(x - a)}{x - a} = q(x) - m =: r(x).$$

Ist f differenzierbar in a , so ist somit r in a stetig fortsetzbar mit $r(a) = Q(a) - m = 0$. Damit ist (2) erfüllt.

Ist umgekehrt (2) erfüllt, so ist $r(x) = q(x) - m$ stetig in a fortsetzbar, und somit ist auch q stetig in a fortsetzbar. \square

Bemerkung. Sei $a \in D$ ein Häufungspunkt von D und $f, g: D \rightarrow \mathbb{R}$ Funktionen. Wir sagen " f und g stimmen in a in 1. Ordnung überein", falls

$$f(x) - g(x) = r(x)(x - a),$$

wobei $r: D \rightarrow \mathbb{R}$ eine in a stetige Funktion mit $r(a) = 0$ ist. Damit lässt sich Satz 4.1.2 auch so formulieren: f ist in $a \in D$ genau dann differenzierbar, falls eine *affin lineare Abbildung* $T: \mathbb{R} \rightarrow \mathbb{R}$ mit $T(x) = f(a) + m(x - a)$ existiert, die mit f in a in 1. Ordnung übereinstimmt. Die eindeutig bestimmte Abbildung T heißt auch *Tangente*.

Eine in a differenzierbare Funktion ist natürlich auch in a stetig.

Satz 4.1.3. Sei a ein Häufungspunkt von D und $f: D \rightarrow \mathbb{R}$ in a differenzierbar. Dann ist f in a stetig.

Beweis. Ist f in a differenzierbar, so existiert wegen Satz 4.1.2 ein $m \in \mathbb{R}$ und eine in a stetige Funktion $r: D \rightarrow \mathbb{R}$ mit $r(a) = 0$ und $f(x) = f(a) + m(x - a) + r(x)(x - a)$. Da die rechte Seite in a stetig ist, ist auch f in a stetig. \square

Bemerkung. Umgekehrt sind stetige Funktionen im allgemeinen nicht differenzierbar. Ein Beispiel hierfür ist die Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$ mit $f(x) = |x|$. Diese Funktion ist stetig, aber nicht differenzierbar in 0. Der Differenzenquotient

$$q(x) = \frac{|x|}{x} = \begin{cases} 1, & x > 0 \\ -1, & x < 0 \end{cases}$$

ist nicht in 0 stetig fortsetzbar.

4.2 Geschwindigkeit und Beschleunigung von Kurven im \mathbb{R}^n

Um die Bewegung eines Teilchens (eines Systems von Teilchen) beschreiben zu können, ist der Begriff der *Kurve* im n -dimensionalen Vektorraum

$$\mathbb{R}^n := \{x = (x_1, \dots, x_n) \mid x_i \in \mathbb{R}\}$$

von fundamentaler Bedeutung. Die Elemente des \mathbb{R}^n nennt man *Vektoren*. Auf \mathbb{R}^n ist eine *Addition* und eine *Multiplikation mit reellen Zahlen (skalare Multiplikation)* erklärt. Sind $x = (x_1, \dots, x_n), y = (y_1, \dots, y_n)$ zwei Vektoren und $\lambda \in \mathbb{R}$ ein Skalar, so definiere $x + y$ und λx durch

$$x + y = (x_1, \dots, x_n) + (y_1, \dots, y_n) := (x_1 + y_1, \dots, x_n + y_n)$$

und

$$\lambda x = \lambda(x_1, \dots, x_n) := (\lambda x_1, \dots, \lambda x_n).$$

Definition 4.2.1. Sei $c : D \rightarrow \mathbb{R}^n$ eine Abbildung, wobei D eine Teilmenge von \mathbb{R} ist. Ist

$$c(t) = (c_1(t), \dots, c_n(t))$$

so heißen für jedes $i \in \{1, \dots, n\}$ die Funktionen $c_i : D \rightarrow \mathbb{R}$ die *Komponentenfunktionen* von c . Sei $a \in \mathbb{R}$ ein Berührungspunkt von D , so sagt man, dass der Grenzwert von c in a existiert, falls für jedes $i \in \{1, \dots, n\}$ der Grenzwert von $c_i : D \rightarrow \mathbb{R}$ existiert. Wir schreiben dann:

$$\lim_{t \rightarrow a} c(t) = \left(\lim_{t \rightarrow a} c_1(t), \dots, \lim_{t \rightarrow a} c_n(t) \right).$$

Ist $a \in D$ und existiert $\lim_{t \rightarrow a} c(t) = c(a)$, so heißt c *stetig* in a .

Im Falle, dass $D = I$ ein Intervall ist, heißt $c : I \rightarrow \mathbb{R}^n$ eine *Kurve*. In diesem Falle ist jedes Element $a \in I$ ein Berührungspunkt.

Bemerkungen.

- (a) Eine Abbildung $c : D \rightarrow \mathbb{R}^n$ ist genau dann stetig in a , falls alle ihre Komponentenfunktionen in a stetig sind.
- (b) Da \mathbb{C} sich mit \mathbb{R}^2 identifizieren lässt, kann man eine Abbildung $c : I \rightarrow \mathbb{C}$ mit $I \subset \mathbb{R}$ auch als Kurve in \mathbb{R}^2 auffassen.
- (c) Beschreibt $c : I \rightarrow \mathbb{R}^n$ die Bewegung eines Teilchens (oder mehrerer Teilchen) im \mathbb{R}^n , so beschreibt $c(t)$ den Punkt, in dem sich das Teilchen zum Zeitpunkt $t \in D$ befindet.

Nun lässt sich der Begriff des Differenzenquotienten und der Ableitung unmittelbar auf Kurven übertragen.

Definition 4.2.2. Sei D eine Teilmenge von \mathbb{R} und $a \in \mathbb{R}$ ein Häufungspunkt von D . Ist $c : D \rightarrow \mathbb{R}^n$ eine Abbildung, so heißt

$$q(t) = \frac{c(t) - c(a)}{t - a}$$

Differenzenquotient von c in a . Wir sagen, dass die Abbildung $c : D \rightarrow \mathbb{R}^n$ in a *differenzierbar* ist, falls der Grenzwert von q in a existiert. Wir schreiben dann:

$$c'(a) := \lim_{t \rightarrow a} q(t).$$

Bemerkungen.

- (a) Nach Definition ist die Abbildung $c : D \rightarrow \mathbb{R}^n$ genau dann differenzierbar, falls ihre Komponentenfunktionen differenzierbar sind. Im Falle der Differenzierbarkeit gilt:

$$c'(a) = (c'_1(a), \dots, c'_n(a)).$$

- (b) Sei $c : I \rightarrow \mathbb{R}^n$ eine differenzierbare Kurve, die die Bewegung eines Teilchens beschreibt. Dann definiert der Vektor $c'(a)$ die *Geschwindigkeit* des Teilchens (*Geschwindigkeitsvektor*) zum Zeitpunkt a . Ist $c'(a) \neq 0$, so lässt sich $c'(a)$ als *Tangentialvektor an die Kurve im Punkte $c(a)$* interpretieren. Es ist manchmal üblich, $\dot{c}(a)$ statt $c'(a)$ zu schreiben.

- (c) Über die Änderung der Geschwindigkeit wird die *Beschleunigung* eines Teilchens definiert. Für ihre Definition benötigt man die Existenz der Ableitung der Kurve $c : I \rightarrow \mathbb{R}^n$ für jedes $a \in I$. Dann definiert die Ableitung $c'(t)$ wieder eine Kurve $c' : I \rightarrow \mathbb{R}^n$. Die Kurve $c' : I \rightarrow \mathbb{R}^n$ heißt auch *Geschwindigkeitsfeld* von c . Ist diese Kurve in a differenzierbar, so heißt

$$c''(a) := (c')'(a)$$

die *Beschleunigung* von c in a .

Da die Analysis der Kurven auf die Analysis ihrer Komponentenfunktionen zurückgeführt werden kann, wenden wir uns jetzt wieder reellwertigen Funktion zu.

4.3 Rechenregeln für differenzierbare Funktionen

Satz 4.3.1. *Seien $a \in D$ ein Häufungspunkt von D und $f, g : D \rightarrow \mathbb{R}$ in a differenzierbar. Dann sind auch $f + g$, $f \cdot g$ und $\frac{1}{f}$ (falls $f(a) \neq 0$) in a differenzierbar und es gilt:*

$$(1) (f + g)'(a) = f'(a) + g'(a).$$

$$(2) (f \cdot g)'(a) = f'(a) \cdot g(a) + g'(a) \cdot f(a).$$

$$(3) \left(\frac{1}{f}\right)'(a) = -\frac{f'(a)}{f^2(a)}.$$

Beweis. Da f und g in a differenzierbar sind, existieren in a stetige Funktionen $f_1, g_1 : D \rightarrow \mathbb{R}$ (nämlich die stetigen Fortsetzungen der Differenzenquotienten von f und g), mit

$$f(x) = f(a) + f_1(x)(x - a)$$

und

$$g(x) = g(a) + g_1(x)(x - a).$$

Insbesondere ist $f'(a) = f_1(a)$ und $g'(a) = g_1(a)$. Dann gilt:

(1)

$$f(x) + g(x) = f(a) + g(a) + (f_1(x) + g_1(x))(x - a).$$

Also besitzt der Differenzenquotient von $f + g$ in a die stetige Fortsetzung $Q(x) := f_1(x) + g_1(x)$ und somit ist $f + g$ in a differenzierbar mit

$$(f + g)'(a) = Q(a) = f_1(a) + g_1(a).$$

(2)

$$\begin{aligned}
f(x) \cdot g(x) &= f(a)g(a) + f(a)g_1(x)(x-a) + f_1(x)(x-a)g(a) \\
&\quad + f_1(x)g_1(x)(x-a)^2 \\
&= f(a)g(a) + (f(a)g_1(x) + f_1(x)g(a) + f_1(x)g_1(x)(x-a))(x-a).
\end{aligned} \tag{4.1}$$

Also besitzt der Differenzenquotient von $f \cdot g$ in a die stetige Fortsetzung $Q(x) := f(a)g_1(a) + f_1(a)g(a) + f_1(x)g_1(x)(x-a)$ und somit ist $f \cdot g$ in a differenzierbar mit

$$(f \cdot g)'(a) = Q(a) = f(a)g_1(a) + f_1(a)g(a) = f(a)g'(a) + f'(a)g(a).$$

(3) Da $f(a) \neq 0$ ist, existiert ein $\delta > 0$, so dass für $x \in B(a, \delta) \cap D$ auch $f(x)$ von Null verschieden ist. Für solche x gilt:

$$\frac{1}{f(x)} - \frac{1}{f(a)} = \frac{f(a) - f(x)}{f(x)f(a)} = \frac{-f_1(x)(x-a)}{f^2(a) + f_1(x)(x-a)f(a)}.$$

Also gilt

$$\frac{1}{f(x)} = \frac{1}{f(a)} + \frac{-f_1(x)}{f^2(a) + f_1(x)(x-a)f(a)}(x-a).$$

und der Differenzenquotient von $\frac{1}{f}$ in a hat die stetige Fortsetzung

$$Q(x) = \frac{-f_1(x)}{f^2(a) + f_1(x)(x-a)f(a)}.$$

Somit ist $\frac{1}{f}$ in a differenzierbar mit

$$\left(\frac{1}{f}\right)'(a) = Q(a) = \frac{-f_1(a)}{f^2(a)} = \frac{-f'(a)}{f^2(a)}.$$

□

Bemerkung. Aus der *Produktregel* (2) erhalten wir mit Hilfe von (3) die folgende *Quotientenregel*: Sind $f, g: D \rightarrow \mathbb{R}$ in a differenzierbar mit $g(a) \neq 0$, so gilt:

$$\left(\frac{f}{g}\right)'(a) = \left(f \cdot \frac{1}{g}\right)'(a) = f'(a) \cdot \frac{1}{g(a)} + f(a) \frac{-g'(a)}{g^2(a)} = \frac{f'(a)g(a) - f(a)g'(a)}{g^2(a)}.$$

Die nächste Regel behandelt die Verknüpfung differenzierbarer Abbildungen.

Satz 4.3.2. (Kettenregel)

Es seien $f: D \rightarrow E \subset \mathbb{R}$ und $g: E \rightarrow \mathbb{R}$ Funktionen. Ist a Häufungspunkt von D , $f(a) = b$ Häufungspunkt von E und sind f in a und g in $f(a)$ differenzierbar, so ist auch $g \circ f$ in a differenzierbar mit

$$(g \circ f)'(a) = g'(f(a)) \cdot f'(a).$$

Beweis. Nach Voraussetzung existieren Funktionen $f_1 : D \rightarrow \mathbb{R}$ und $g_1 : E \rightarrow \mathbb{R}$, die in $a \in D$ und $b \in E$ stetig sind, so dass gilt:

$$f(x) = f(a) + f_1(x)(x - a)$$

$$g(y) = g(b) + g_1(y)(y - b).$$

Dann folgt:

$$\begin{aligned} g \circ f(x) &= g(b) + g_1(f(x))(f(x) - f(a)) \\ &= g \circ f(a) + g_1(f(a) + f_1(x)(x - a))f_1(x)(x - a). \end{aligned}$$

Also besitzt der Differenzenquotient $g \circ f$ in a die stetige Fortsetzung

$$Q(x) = g_1(f(a) + f_1(x)(x - a)) \cdot f_1(x)$$

und es gilt:

$$(g \circ f)'(a) = Q(a) = g_1(f(a))f_1(a) = g'(f(a)) \cdot f'(a).$$

□

Bemerkung. Die Kettenregel erlaubt es, bei differenzierbaren, umkehrbaren Abbildungen $f : D \rightarrow f(D) = E$, die *Ableitung der Umkehrfunktion* $f^{-1} : E \rightarrow D$ zu berechnen. Dafür müssen wir allerdings wissen, dass f^{-1} differenzierbar ist. Sind f in a und f^{-1} in $f(a)$ differenzierbar, so gilt: $f^{-1} \circ f(x) = x$, und aus der Kettenregel folgt $(f^{-1})'(f(a)) \cdot f'(a) = 1$, d.h. $(f^{-1})'(f(a)) = \frac{1}{f'(a)}$ und somit $(f^{-1})'(b) = \frac{1}{f'(f^{-1}(b))}$ für $b = f(a)$.

Der folgende Satz zeigt, dass aus der Differenzierbarkeit einer umkehrbaren Funktion die Differenzierbarkeit ihrer Umkehrfunktion folgt, falls diese stetig ist.

Satz 4.3.3. *Es seien $D, E \subset \mathbb{R}$ und $f : D \rightarrow E$ eine bijektive Abbildung.*

Es seien $a \in D$ Häufungspunkt und f in a differenzierbar mit $f'(a) \neq 0$; die Umkehrabbildung $f^{-1} : E \rightarrow D$ sei in $b = f(a)$ stetig. Dann ist f^{-1} in b differenzierbar, und es gilt:

$$(f^{-1})'(b) = \frac{1}{f'(f^{-1}(b))}.$$

Bemerkung. $a \in D$ ist genau dann Häufungspunkt von D , falls eine Folge $x_n \in D \setminus \{a\}$ existiert mit $\lim_{n \rightarrow \infty} x_n = a$. Da f injektiv und stetig in a ist, ist $f(x_n) \neq f(a)$ und $\lim_{n \rightarrow \infty} f(x_n) = f(a)$, d.h. $f(a)$ ist ebenfalls Häufungspunkt von E .

Beweis. Nach Voraussetzung gilt für alle $x \in D$:

$$f(x) = f(a) + f_1(x)(x - a),$$

wobei $f_1 : D \rightarrow \mathbb{R}$ eine stetige Funktion ist mit $f_1(a) = f'(a) \neq 0$. Dann ist $f_1(x) \neq 0$ auf einer kleinen Umgebung von $a = f^{-1}(b)$. Ist $y \in E$ und $x = f^{-1}(y) \in D$, so erhalten wir:

$$y = b + f_1(f^{-1}(y))(f^{-1}(y) - f^{-1}(b)).$$

Da $f_1(f^{-1}(y))$ stetig ist und an der Stelle b den Wert $f'(a) \neq 0$ annimmt, existiert eine Umgebung von b mit $f_1(f^{-1}(y)) \neq 0$ für alle y in dieser Umgebung. Damit erhalten wir:

$$f^{-1}(y) - f^{-1}(b) = \frac{1}{f_1(f^{-1}(y))}(y - b).$$

Da $Q(y) = \frac{1}{f_1(f^{-1}(y))}$ stetig in b ist, ist f^{-1} differenzierbar in b , und es gilt:

$$(f^{-1})'(b) = Q(b) = \frac{1}{f_1(f^{-1}(b))} = \frac{1}{f'(f^{-1}(b))}.$$

□

Korollar 4.3.4. Sei I ein Intervall und $f : I \rightarrow J$ stetig und bijektiv. Ist f in $a \in I$ differenzierbar, so ist die Umkehrabbildung $f^{-1} : J \rightarrow I$ in $f(a)$ differenzierbar.

Beweis. Ist $f : I \rightarrow J$ bijektiv und stetig, so ist wegen Satz 3.2.6 auch J ein Intervall, und die Umkehrabbildung $f^{-1} : J \rightarrow I$ ist stetig. Da jeder Punkt $a \in I$ Häufungspunkt ist, folgt das Korollar aus obigem Satz. □

Beispiele.

- (1) Sei $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ die Funktion $f(x) = x^n$ mit $n \geq 2$ und $f^{-1} : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ die Umkehrabbildung mit $f^{-1}(x) = \sqrt[n]{x} =: x^{\frac{1}{n}}$. Da $f'(x) = nx^{n-1} \neq 0$ für alle $x \in \mathbb{R}_+$, ist f^{-1} differenzierbar, und es gilt:

$$(f^{-1})'(x) = \frac{1}{f'(f^{-1}(x))} = \frac{1}{n \left(x^{\frac{1}{n}}\right)^{n-1}} = \frac{1}{n} x^{\frac{1}{n}(1-n)} = \frac{1}{n} x^{\frac{1}{n}-1}.$$

Erweitern wir den Definitionsbereich von f auf $[0, \infty)$, so ist f^{-1} noch in 0 stetig, aber nicht mehr in 0 differenzierbar.

- (2) Sei $\exp : \mathbb{R} \rightarrow \mathbb{R}_+$ die Exponentialfunktion und $\log : \mathbb{R}_+ \rightarrow \mathbb{R}$ die Umkehrfunktion. Dann gilt:

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

Exponentialfunktionen mit beliebiger Basis $a > 0$ definiert man wie folgt:

$$a^x := \exp(x \log a).$$

4.4 Extrema und Mittelwertsätze

Definition 4.4.1. Sei $f : D \rightarrow \mathbb{R}$ eine Funktion und $a \in D$. Dann heißt a

- (1) ein *globales Maximum* bzw. *Minimum*, falls $f(a) \geq f(x)$ bzw. $f(a) \leq f(x)$ für alle $x \in D$ gilt.
- (2) ein *lokales Maximum* bzw. *Minimum*, falls ein $\epsilon > 0$ existiert mit $f(a) \geq f(x)$ bzw. $f(a) \leq f(x)$ für alle $x \in D \cap B(a, \epsilon)$.

Bemerkungen.

- (a) Ist $D \subset \mathbb{R}$ eine kompakte Menge, so besitzt jede stetige Funktion auf D ein globales Maximum und Minimum.
- (b) Man spricht auch von einem *globalen* bzw. *lokalen Extremum* von f , wenn man nicht zwischen Maximum oder Minimum zu unterscheiden braucht.

Definition 4.4.2. Sei $f : D \rightarrow \mathbb{R}$ eine in $a \in D$ differenzierbare Funktion. Dann nennen wir a *kritischer Punkt* von f , falls $f'(a) = 0$.

Für differenzierbare Funktionen gilt die folgende notwendige Bedingung für die Existenz eines lokalen Extremums.

Satz 4.4.3. Sei $f : D \rightarrow \mathbb{R}$ eine in $a \in D$ differenzierbare Funktion und a lokales Extremum. Sei $a \in D$ ein innerer Punkt, d.h. $B(a, \delta) \subset D$ für ein $\delta > 0$. Dann ist a ein kritischer Punkt.

Bemerkung. Die Umkehrung ist falsch, denn für $f : \mathbb{R} \rightarrow \mathbb{R}$ mit $f(x) = x^3$ gilt $f'(0) = 0$, aber 0 ist kein Extremum.

Beweis. Sei $f : D \rightarrow \mathbb{R}$ eine in $a \in D$ differenzierbare Funktion. Wir nehmen an, dass a ein lokales Maximum ist (sonst betrachte $-f$). Da $a \in D$ innerer Punkt ist, existiert ein $\epsilon > 0$ mit $f(a) \geq f(x)$ für alle $x \in B(a, \epsilon)$. Insbesondere folgt aus $x > a$ mit $x - a < \epsilon$:

$$q(x) = \frac{f(x) - f(a)}{x - a} \leq 0$$

und aus $x < a$ mit $a - x < \epsilon$:

$$q(x) = \frac{f(x) - f(a)}{x - a} \geq 0.$$

Da q stetig in 0 fortsetzbar ist, existieren die linksseitigen und rechtsseitigen Grenzwerte und stimmen überein. Aus den obigen Ungleichungen folgt somit: $\lim_{x \rightarrow 0} q(x) = 0$, d.h. $f'(0) = 0$. \square

Eine hinreichende Bedingung für die Existenz eines kritischen Punktes liefert der folgende *Satz von Rolle*.

Satz 4.4.4 (Satz von Rolle).

Sei $a < b$ und $f : [a, b] \rightarrow \mathbb{R}$ eine stetige Funktion mit $f(a) = f(b)$. Ist f auf dem offenen Intervall (a, b) differenzierbar, so existiert ein $\xi \in (a, b)$ mit $f'(\xi) = 0$.

Beweis. Ist f eine konstante Funktion, so ist $f'(x) = 0$ für alle $x \in (a, b)$. Sei nun f nicht konstant. Da $f : [a, b] \rightarrow \mathbb{R}$ stetig ist, besitzt f ein globales Maximum x_M und ein globales Minimum x_m . Da $f(x_m) \leq f(x) \leq f(x_M)$ für alle $x \in [a, b]$ und $f(a) = f(b)$, liegt x_m oder x_M in (a, b) , denn sonst ist f konstant. Insbesondere existiert ein Extremum $\xi \in (a, b)$ von f und aus Satz 4.4.3 folgt $f'(\xi) = 0$. \square

Aus dem Satz von Rolle läßt sich der wichtige *Mittelwertsatz der Differentialrechnung* ableiten.

Satz 4.4.5 (Mittelwertsatz der Differentialrechnung)

Es sei $f : [a, b] \rightarrow \mathbb{R}$ eine stetige Funktion, die auf (a, b) differenzierbar ist. Dann existiert ein $\xi \in (a, b)$ mit

$$f'(\xi) = \frac{f(b) - f(a)}{b - a}.$$

Bemerkung. Geometrisch bedeutet dieser Satz: Ist $g(a, b)$ die Gerade, die $(a, f(a))$ mit $(b, f(b))$ verbindet, so existiert ein $\xi \in (a, b)$, so dass die Tangente von f in ξ parallel zu $g(a, b)$ ist (s. Abb. 5.2).

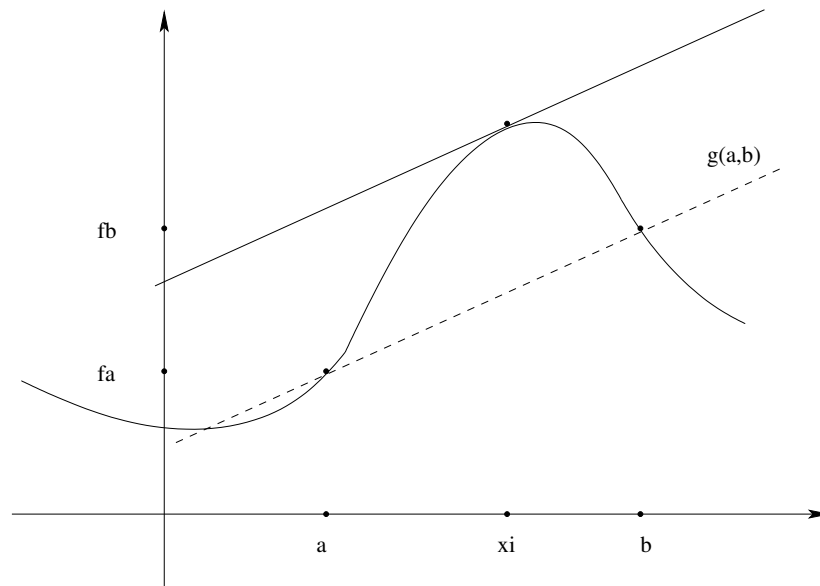


Abbildung 4.2: Der Mittelwertsatz

Beweis. Betrachte die Funktion: $g: [a, b] \rightarrow \mathbb{R}$ mit

$$g(x) = f(x) - (x - a) \frac{f(b) - f(a)}{b - a}.$$

Dann gilt: $g(a) = f(a)$ und $g(b) = f(b) - (f(b) - f(a)) = f(a)$. Da g ebenso wie f auf $[a, b]$ stetig und auf (a, b) differenzierbar ist, existiert wegen des Satzes von Rolle ein $\xi \in (a, b)$ mit $g'(\xi) = 0 = f'(\xi) - \frac{f(b) - f(a)}{b - a}$. \square

Aus dem Mittelwertsatz ergibt sich folgendes Korollar.

Korollar 4.4.6. Ist $f: [a, b] \rightarrow \mathbb{R}$ stetig und auf (a, b) differenzierbar, so gilt:

- (1) f ist monoton wachsend bzw. fallend genau dann, falls $f'(\xi) \geq 0$ bzw. $f'(\xi) \leq 0$ für alle $\xi \in (a, b)$.
- (2) Ist $f'(\xi) > 0$ bzw. $f'(\xi) < 0$ für alle $\xi \in (a, b)$, so ist f streng monoton wachsend bzw. fallend.
- (3) Ist $|f'(\xi)| \leq L$ für $\xi \in (a, b)$, so ist $f: [a, b] \rightarrow \mathbb{R}$ Lipschitz-stetig mit Lipschitz-Konstante L .
- (4) f ist genau dann konstant, falls $f'(\xi) = 0$ für alle $\xi \in (a, b)$.

Bemerkung. Die Umkehrung von (2) in Korollar 4.4.6 ist falsch. Alle Aussagen werden falsch, falls der Definitionsbereich von f kein Intervall ist.

Beweis.

(1) Ist f monoton wachsend und $\xi \in (a, b)$, so ist

$$q(x) = \frac{f(x) - f(\xi)}{x - \xi} \geq 0$$

für alle $x \in [a, b] \setminus \{\xi\}$. Damit ist auch $f'(\xi) = \lim_{x \rightarrow \xi} q(x) \geq 0$.

Ist $a \leq x_1 < x_2 \leq b$, so existiert wegen des Mittelwertsatzes ein $\xi \in (x_1, x_2)$ mit

$$f'(\xi) = \frac{f(x_2) - f(x_1)}{x_2 - x_1}. \quad (*)$$

Ist $f'(\xi) \geq 0$ für alle $\xi \in (a, b)$, so ist f wegen (*) monoton wachsend.

In Fall $f'(\xi) \leq 0$ behandelt man analog.

(2) Ist $f'(\xi) > 0$ für alle $\xi \in (a, b)$, so ist f wegen (*) streng monoton wachsend.

In Fall $f'(\xi) < 0$ behandelt man analog.

(3) Ist $|f'(\xi)| \leq L$ für alle $\xi \in (a, b)$, und ist $a \leq x_1 < x_2 \leq b$, so ergibt sich aus (*)

$$|f(x_2) - f(x_1)| \leq L|x_2 - x_1|,$$

d.h. L ist Lipschitz-stetig mit Lipschitz-Konstante L .

(4) Ist f konstant, so sind alle Differenzenquotienten von f identisch 0, also $f'(\xi) = 0$ für alle $\xi \in (a, b)$. Ist $f'(\xi) = 0$ für alle $\xi \in (a, b)$, so ist f wegen (3) Lipschitz-stetig mit Lipschitz-Konstante 0. Insbesondere ist f konstant. □

Der Mittelwertsatz besitzt folgende Verallgemeinerung.

Satz 4.4.7. (Verallgemeinerter Mittelwertsatz)

Es sei $a < b$, und $f, g : [a, b] \rightarrow \mathbb{R}$ seien auf $[a, b]$ stetig und auf (a, b) differenzierbar. Ist $g'(x) \neq 0$ für alle $x \in (a, b)$, so existiert ein $\xi \in (a, b)$ mit

$$\frac{f'(\xi)}{g'(\xi)} = \frac{f(b) - f(a)}{g(b) - g(a)}.$$

Bemerkung. Eine naive Anwendung von Satz 4.4.4 würde

$$\frac{f'(\xi)}{g'(\mu)} = \frac{f(b) - f(a)}{g(b) - g(a)}$$

liefern. Wichtig ist hier aber, dass auf der linken Seite der gleiche Wert ξ im Zähler und Nenner gewählt werden kann.

Beweis. Betrachte die Funktion $h : [a, b] \rightarrow \mathbb{R}$ mit

$$h(x) = (g(b) - g(a))f(x) - (f(b) - f(a))g(x).$$

Dann gilt: $h(b) = -g(a)f(b) + f(a)g(b)$ und $h(a) = g(b)f(a) - f(b)g(a) = h(b)$. Wegen des Satzes von Rolle gibt es ein $\xi \in (a, b)$ mit

$$0 = h'(\xi) = (g(b) - g(a))f'(\xi) - (f(b) - f(a))g'(\xi).$$

Da $g'(x) \neq 0$ für alle $x \in (a, b)$, ist wegen des Mittelwertsatzes $g(b) \neq g(a)$ und es gilt:

$$\frac{f'(\xi)}{g'(\xi)} = \frac{f(b) - f(a)}{g(b) - g(a)}$$

□

4.5 Höhere Ableitungen und die Taylorsche Formel

Sei $D \subset \mathbb{R}$ eine Teilmenge, so dass jedes $x \in D$ Häufungspunkt von D ist (z.B. ist dies erfüllt, falls D ein Intervall ist). Ist $f: D \rightarrow \mathbb{R}$ differenzierbar für jedes $x \in D$, so definiert die erste Ableitung $f': D \rightarrow \mathbb{R}$ wieder eine Funktion. Ist allgemein die k -te Ableitung $f^{(k)}: D \rightarrow \mathbb{R}$ erklärt und differenzierbar, so definieren wir:

$$f^{(k+1)} := (f^{(k)})'.$$

Mit

$$C^k(D) = \{f: D \rightarrow \mathbb{R} \mid f \text{ ist } k\text{-mal differenzierbar und } f^{(k)} \text{ ist stetig}\}$$

bezeichnen wir die *Menge der k -mal stetig differenzierbaren Funktionen*, mit $C^\infty(D)$ die *Menge der unendlich oft differenzierbaren Funktionen*. Offensichtlich gilt:

$$C^\infty(D) = \bigcap_{k \geq 1} C^k(D).$$

$C^0(D)$ ist dann die Menge der stetigen Funktion. Es gilt:

$$C^0(D) \supset C^1(D) \supset \dots \supset C^k(D) \supset \dots \supset C^\infty(D).$$

Wir haben gesehen, dass sich differenzierbare Funktionen in 1. Ordnung durch lineare Funktionen approximieren lassen; ist nämlich $f: D \rightarrow \mathbb{R}$ differenzierbar, so approximiert $T(x) = f(x_0) + f'(x_0)(x - x_0)$ die Funktion f in x_0 in 1. Ordnung. Es gilt $T(x_0) = f(x_0)$ und $T'(x_0) = f'(x_0)$.

Ist nun f n -mal differenzierbar, so ist es naheliegend, f durch ein Polynom vom Grade $\leq n$ zu approximieren.

Wir suchen ein Polynom

$$T_n(x) = \sum_{k=0}^n a_k (x - x_0)^k$$

mit

$$T_n(x_0) = f(x_0), \quad T_n'(x_0) = f'(x_0), \quad \dots, \quad T_n^{(n)}(x_0) = f^{(n)}(x_0).$$

Da $T_n^{(k)}(x_0) = k! \cdot a_k$, folgt $a_k = \frac{f^{(k)}(x_0)}{k!}$. Das Polynom

$$T_n(x) = T_n(x, x_0) := \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k$$

heißt das *n -te Taylorpolynom* von f in x_0 . Der Punkt x_0 heißt *Entwicklungspunkt* des Taylorpolynoms. Das folgende Lemma zeigt, dass die Ableitung dieser Funktion als Funktion vom Entwicklungspunkt, nur von der $(n + 1)$ -ten Ableitung von f abhängt.

Lemma 4.5.1. Sei $f: (a, b) \rightarrow \mathbb{R}$ eine $(n + 1)$ -mal differenzierbare Abbildung, $n \in \mathbb{N}$. Für ein festes $x \in (a, b)$ betrachte die Funktion $g: (a, b) \rightarrow \mathbb{R}$ mit

$$g(t) = T_n(x, t) = \sum_{k=0}^n \frac{f^{(k)}(t)}{k!} (x - t)^k.$$

Dann gilt:

$$g'(t) = \frac{f^{(n+1)}(t)}{n!} (x - t)^n.$$

Beweis.

$$\begin{aligned} g'(t) &= \left(\sum_{k=0}^n \frac{f^{(k)}(t)}{k!} (x-t)^k \right)' = \sum_{k=0}^n \frac{f^{(k+1)}(t)}{k!} (x-t)^k + \sum_{k=0}^n \frac{f^{(k)}(t)}{k!} k(x-t)^{k-1} (-1) \\ &= \sum_{k=0}^n \frac{f^{(k+1)}(t)}{k!} (x-t)^k - \sum_{k=1}^n \frac{f^{(k)}(t)}{(k-1)!} (x-t)^{k-1} = \frac{f^{(n+1)}(t)}{n!} (x-t)^n \end{aligned}$$

□

Bemerkung. Ist also $f^{(n+1)}(t) = 0$ für alle $t \in (a, b)$, so ist $g'(t) = 0$ und somit g eine konstante Funktion. Dann folgt:

$$f(x) = g(x) = g(x_0) = \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x-x_0)^k,$$

d.h. f stimmt mit seinem n -ten Taylorpolynom überein.

Seien $x_1, x_2 \in \mathbb{R}$, so bezeichnen wir mit

$$I_{x_1, x_2} = \begin{cases} [x_1, x_2], & \text{falls } x_1 < x_2 \\ [x_2, x_1], & \text{falls } x_2 \leq x_1 \end{cases}$$

das abgeschlossene Intervall, welches x_1 und x_2 als Randpunkte besitzt, und mit $\overset{\circ}{I}_{x_1, x_2} = I_{x_1, x_2} \setminus \{x_1, x_2\}$ das zugehörige offene Intervall.

Satz 4.5.2. Sei $f : [a, b] \rightarrow \mathbb{R}$ eine n -mal stetig differenzierbare Abbildung, die auf (a, b) sogar $(n+1)$ -mal differenzierbar ist. Für $x_0, x \in [a, b]$ mit $x_0 \neq x$ sei $h : I_{x_0, x} \rightarrow \mathbb{R}$ eine stetige Funktion, die auf $\overset{\circ}{I}_{x_0, x}$ differenzierbar ist, mit $h'(t) \neq 0$ für alle $t \in \overset{\circ}{I}_{x_0, x}$. Dann existiert ein $\xi \in \overset{\circ}{I}_{x_0, x}$ mit

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x-x_0)^k + \frac{f^{(n+1)}(\xi)}{n!} (x-\xi)^n \frac{h(x) - h(x_0)}{h'(\xi)}.$$

Beweis. Die Funktionen $g(t) = \sum_{k=0}^n \frac{f^{(k)}(t)}{k!} (x-t)^k$ und $h(t)$ sind nach Voraussetzung auf $I_{x_0, x}$ stetig und auf $\overset{\circ}{I}_{x_0, x}$ differenzierbar mit $h'(t) \neq 0$ für $t \in \overset{\circ}{I}_{x_0, x}$. Aus dem verallgemeinerten Mittelwertsatz folgt:

$$\frac{g(x) - g(x_0)}{h(x) - h(x_0)} = \frac{g'(\xi)}{h'(\xi)}$$

für ein $\xi \in \overset{\circ}{I}_{x_0, x}$. Da $g(x) = f(x)$, $g(x_0) = \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x-x_0)^k$ und $g'(\xi) = \frac{f^{(n+1)}(\xi)}{n!} (x-\xi)^n$, folgt die Behauptung. □

Bemerkungen.

- (a) Die Differenz $R_{n+1}(x, x_0) := f(x) - \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x-x_0)^k$ nennen wir das $(n+1)$ -te Restglied. Durch die Wahl einer Funktion h mit den im Satz 4.5.2 angegebenen Eigenschaften erhalten wir eine Darstellung des Restglieds.

- (b) Setzen wir speziell $h(t) = (x - t)^p$ mit $p \in \mathbb{N}$, so gilt $h'(t) = -p(x - t)^{p-1}$, und wir erhalten: (*Restglied von Schlömilch*)

$$\begin{aligned} R_{n+1}(x, x_0) &= \frac{f^{(n+1)}(\xi)}{n!} (x - \xi)^n \frac{(x - x_0)^p}{p(x - \xi)^{p-1}} \\ &= \frac{f^{(n+1)}(\xi)}{p n!} (x - \xi)^{n+1-p} (x - x_0)^p \end{aligned}$$

für ein $\xi \in \overset{\circ}{I}_{x_0, x}$. Daraus ergeben sich zwei wichtige Spezialfälle:

- (i) $p = n + 1$: (*Restglied von Lagrange*)

$$R_{n+1}(x, x_0) = \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - x_0)^{n+1}$$

für ein $\xi \in \overset{\circ}{I}_{x_0, x}$.

- (ii) $p = 1$: (*Restglied von Cauchy*)

$$R_{n+1}(x, x_0) = \frac{f^{(n+1)}(\xi)}{n!} (x - \xi)^n (x - x_0)$$

für ein $\xi \in \overset{\circ}{I}_{x_0, x}$.

- (c) Die Identität

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k + \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - x_0)^{n+1}$$

wird auch *Taylor'sche Formel* oder *Satz von Taylor* genannt.

Beispiel. Betrachte die Funktion $f(x) = \cos(x)$. Wir wollen für $x_0 = 0$ die zu f gehörige Taylorformel berechnen. Aus $\cos^{(2k)}(x) = (-1)^k \cos(x)$ und $\cos^{(2k+1)}(x) = (-1)^{k+1} \sin(x)$ erhalten wir

$$\begin{aligned} f(x) = \cos(x) &= \sum_{k=0}^{2n+1} \frac{f^{(k)}(0)}{k!} x^k + \frac{f^{(2n+2)}(\xi)}{(2n+2)!} x^{2n+2} \\ &= \sum_{k=0}^n (-1)^k \frac{x^{2k}}{k!} + (-1)^{n+1} \cos(\xi) \frac{x^{2n+2}}{(2n+2)!} \end{aligned}$$

für ein $\xi \in \overset{\circ}{I}_{x_0, x}$. Insbesondere erhalten wir für alle $x \in \mathbb{R}$ die Abschätzung:

$$\left| \cos(x) - \sum_{k=0}^n (-1)^k \frac{x^{2k}}{k!} \right| \leq \frac{|x|^{2n+2}}{(2n+2)!}.$$

Diese Abschätzung hatten wir in Satz 2.6.9 direkt aus der Reihendarstellung des Cosinus hergeleitet. Allerdings konnten wir sie mit den dort zur Verfügung stehenden Hilfsmittel nur für $|x| \leq 2n + 3$ beweisen.

Definition 4.5.3. Es seien $f, g : D \rightarrow \mathbb{R}$ Funktionen und $x_0 \in D$. Wir sagen, dass f, g in n -ter Ordnung in x_0 übereinstimmen, falls eine in x_0 stetige Funktion $r : D \rightarrow \mathbb{R}$ existiert mit $r(x_0) = 0$ und $f(x) - g(x) = r(x)(x - x_0)^n$.

Es gilt nun der Satz:

Satz 4.5.4. Sei I ein Intervall und $f : I \rightarrow \mathbb{R}$ eine n -mal stetig differenzierbare Abbildung. Ist $x_0 \in I$, so stimmt das n -te Taylorpolynom $T_n = \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k$ mit f in x_0 in n -ter Ordnung überein.

Beweis. Aus der Taylorformel folgt:

$$\begin{aligned} f(x) - \sum_{k=0}^{n-1} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k &= \frac{f^{(n)}(\xi_x)}{n!} (x - x_0)^n \\ &= \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n + \frac{f^{(n)}(\xi_x) - f^{(n)}(x_0)}{n!} (x - x_0)^n \end{aligned}$$

mit $\xi_x \in \overset{\circ}{I}_{x_0, x} \subset I$. Also folgt

$$f(x) - \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k = \frac{f^{(n)}(\xi_x) - f^{(n)}(x_0)}{n!} (x - x_0)^n.$$

Da $f^{(n)}$ stetig ist, ist wegen $\xi_x \in \overset{\circ}{I}_{x_0, x}$ auch

$$r(x) := \frac{f^{(n)}(\xi_x) - f^{(n)}(x_0)}{n!} (x - x_0)^n$$

in x_0 stetig und es gilt $r(x_0) = 0$. □

Aus der Taylorsche Formel lassen sich leicht hinreichende Kriterien für lokale Extrema herleiten.

Satz 4.5.5. Es sei $f : (a, b) \rightarrow \mathbb{R}$ n -mal stetig differenzierbar und $x_0 \in (a, b)$, so dass $f'(x_0) = f''(x_0) = \dots = f^{(n-1)}(x_0) = 0$ und $f^{(n)}(x_0) \neq 0$. Dann gilt:

- (a) Ist n ungerade, so ist x_0 kein Extremum.
- (b) Ist n gerade, so ist x_0 ein lokales Minimum, falls $f^{(n)}(x_0) > 0$, und ein lokales Maximum, falls $f^{(n)}(x_0) < 0$.

Beweis. Ist $f'(x_0) = \dots = f^{(n-1)}(x_0) = 0$, so folgt aus der Taylorformel:

$$f(x) = f(x_0) + \frac{f^{(n)}(\xi)}{n!} (x - x_0)^n$$

mit $\xi \in \overset{\circ}{I}_{x_0, x}$. Da $f^{(n)}(x)$ stetig ist, existiert ein $\delta > 0$ mit $f^{(n)}(x) \neq 0$ für alle $x \in B(x_0, \delta)$. Ist n ungerade, so nimmt $(x - x_0)^n$ und somit auch $f(x) - f(x_0)$ auf der Umgebung $B(\delta, x_0)$ beide Vorzeichen an.

Sei n gerade. Ist $f^{(n)}(x_0) > 0$, so folgt $f^{(n)}(x) > 0$ für alle $x \in B(x_0, \delta)$ und somit $f(x) - f(x_0) > 0$ für alle $x \in B(x_0, \delta) \setminus \{x_0\}$. Ist $f^{(n)}(x_0) < 0$, so folgt $f(x) - f(x_0) < 0$ für alle $x \in B(x_0, \delta) \setminus \{x_0\}$. □

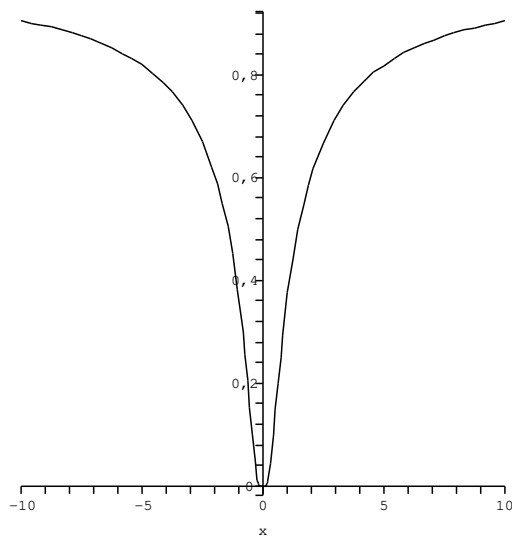


Abbildung 4.3: Eine Funktion, deren Taylorreihe in 0 überall verschwindet

Bemerkung. Der Beweis zeigt, dass es sich im Falle (b) sogar um ein striktes lokales Extremum handelt.

Ist I ein Intervall, so können wir jeder unendlich oft differenzierbaren Funktion eine Taylorreihe zuordnen.

Definition 4.5.6. Seien I ein Intervall und $f \in C^\infty(I)$. Ist $x_0 \in I$, so heißt

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k$$

die *Taylorreihe* von f mit *Entwicklungspunkt* x_0 .

Bemerkungen.

- (1) Taylorreihen sind Potenzreihen, d.h. Reihen der Form $\sum_{k=0}^{\infty} a_k (x - x_0)^k$. Wie in Abschnitt 2.7 gezeigt wurde, konvergieren sie absolut für alle $x \in B(x_0, R)$, falls $R = \frac{1}{\lim_{k \rightarrow \infty} \sqrt[k]{|a_k|}}$ der Konvergenzradius der Reihe ist.
- (2) Ist die zu einer Funktion gehörige Taylorreihe konvergent, so muss sie mit dieser Funktion nicht übereinstimmen. Zum Beispiel verschwinden alle Ableitungen der Funktion

$$f(x) = \begin{cases} \exp\left(-\frac{1}{|x|}\right) & \text{für } x \neq 0 \\ 0 & \text{für } x = 0 \end{cases}$$

im Nullpunkt (s. Abb. 4.3). Insbesondere beschreibt die Taylorreihe die Nullfunktion (d.h. die Funktion, die jeder reellen Zahl den Wert Null zuordnet), die offensichtlich nur für $x = 0$ mit f übereinstimmt.

(3) Offensichtlich gilt für $x, x_0 \in I$:

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k = f(x) \text{ genau dann, falls } \lim_{n \rightarrow \infty} R_{n+1}(x, x_0) = 0.$$

Unter Verwendung der Taylorformel erhalten wir ein einfaches hinreichendes Kriterium für die Darstellung einer Funktion durch ihre Taylorreihe.

Satz 4.5.7. *Seien $I \subset \mathbb{R}$ ein Intervall, $x_0, x \in I$ und $f \in C^\infty(I)$. Existiert eine Konstante $L > 0$ mit $|f^{(n)}(x)| \leq L$ für alle $n \in \mathbb{N}$ und alle $x \in I$, so gilt:*

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k$$

für alle $x \in I$.

Beweis. Sei $x \in I$, so erhalten wir aus der Taylorformel:

$$\left| R_{n+1}(x, x_0) \right| = \left| f(x) - \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k \right| = \left| \frac{f^{(n+1)}(\xi)}{(n+1)!} \right| |x - x_0|^{n+1}$$

für ein $\xi \in I_{x_0, x} \subset I$.

Da $|f^{(n)}(x)| \leq L$ für alle $n \in \mathbb{N}$ und $x \in I$, und da $\frac{|x-x_0|^{n+1}}{(n+1)!} \rightarrow 0$ für $n \rightarrow \infty$ gilt (siehe Beispiel (e) nach dem Satz 2.1.5), folgt:

$$\lim_{n \rightarrow \infty} R_{n+1}(x, x_0) = 0$$

□

Es sei $f(x) = \sum_{k=0}^{\infty} a_k (x - x_0)^k$ eine Potenzreihe mit Konvergenzradius $R > 0$.

Wie wir in Abschnitt 3.1 bemerkt haben, ist f stetig auf $B(x_0, R)$. Wir wollen nun zeigen, dass f beliebig oft differenzierbar ist, und dass die Taylorreihe von f in x_0 mit der Potenzreihe übereinstimmt. Dazu benötigen wir folgendes Lemma. Dieses Lemma gilt natürlich auch für komplexe Potenzreihen und wir werden es in diesem allgemeinen Fall formulieren und beweisen.

Lemma 4.5.8. *Es sei $(a_n)_{n \in \mathbb{N}_0}$ eine Folge komplexer Zahlen und $\sum_{n=0}^{\infty} a_n (z - z_0)^n$ eine Potenzreihe mit Entwicklungspunkt z_0 und Konvergenzradius R . Ist $z_1 \in B(x_0, R)$, so existiert eine Potenzreihe $\sum_{k=0}^{\infty} b_k (z - z_1)^k$ mit Entwicklungspunkt z_1 , so dass*

$$\sum_{n=0}^{\infty} a_n (z - z_0)^n = \sum_{k=0}^{\infty} b_k (z - z_1)^k$$

für alle z mit $|z - z_1| < R - |z_1 - z_0|$ gilt (s. Abb. 4.4). Die Koeffizienten b_k sind gegeben durch

$$b_k = \sum_{n=k}^{\infty} a_n \binom{n}{k} (z_1 - z_0)^{n-k}.$$

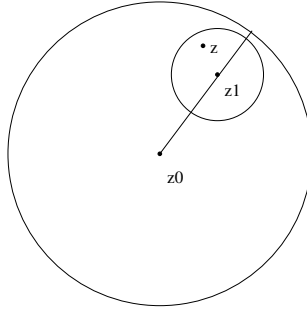


Abbildung 4.4:

Beweis. Es folgt aus der binomischen Formel:

$$\begin{aligned} \sum_{n=0}^{\infty} a_n (z - z_0)^n &= \sum_{n=0}^{\infty} a_n ((z - z_1) + (z_1 - z_0))^n \\ &= \sum_{n=0}^{\infty} \sum_{k=0}^n a_n \binom{n}{k} (z - z_1)^k (z_1 - z_0)^{n-k}. \end{aligned}$$

Da $\binom{n}{k} = \frac{n(n-1)\dots(n-k+1)}{k!} = 0$ für $k > n$, ist die rechte Seite gleich

$$\sum_{n=0}^{\infty} \sum_{k=0}^{\infty} a_n \binom{n}{k} (z - z_1)^k (z_1 - z_0)^{n-k}.$$

Das Lemma ist bewiesen, falls wir die Summationsreihenfolge vertauschen dürfen. Denn dann ist

$$\begin{aligned} \sum_{n=0}^{\infty} a_n (z - z_0)^n &= \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} a_n \binom{n}{k} (z_1 - z_0)^{n-k} (z - z_1)^k \\ &= \sum_{k=0}^{\infty} \left(\sum_{n=k}^{\infty} a_n \binom{n}{k} (z_1 - z_0)^{n-k} \right) (z - z_1)^k \end{aligned}$$

Nach Satz 2.6.3 dürfen wir diese Vertauschung durchführen, falls eine Konstante $L > 0$ existiert, mit

$$c_{ml} := \sum_{n=0}^m \sum_{k=0}^l \left| a_n \binom{n}{k} (z - z_1)^k (z_1 - z_0)^{n-k} \right| \leq L$$

für alle $l, m \in \mathbb{N}$. Wegen $\binom{n}{k} = 0$ für $k > n$ folgt aus der binomischen Formel:

$$\begin{aligned} c_{ml} &\leq \sum_{n=0}^m \sum_{k=0}^n |a_n| \binom{n}{k} |z - z_1|^k |z_1 - z_0|^{n-k} \\ &= \sum_{n=0}^m |a_n| (|z - z_1| + |z_1 - z_0|)^n. \end{aligned}$$

Da $|z - z_1| < R - |z_1 - z_0|$, existiert ein $\epsilon > 0$ mit $|z - z_1| + |z_1 - z_0| < R - \epsilon$. Somit ist

$$\sum_{n=0}^{\infty} |a_n|(R - \epsilon)^n < \infty,$$

denn der Konvergenzradius ist $R > 0$ und die Reihe konvergiert absolut im Innern des Konvergenzbereiches (Satz 2.7.2). Also ist mit $L := \sum_{n=0}^{\infty} |a_n|(R - \epsilon)^n$ die verlangte Konstante gefunden. \square

Für den nächsten Satz werden wir uns wieder auf reelle Potenzreihen beschränken, da wir Differenzierbarkeit nur für reelle Funktionen definiert haben.

Satz 4.5.9. Sei $\sum_{n=0}^{\infty} a_n(x - x_0)^n$ eine reelle Potenzreihe mit Konvergenzradius $R > 0$. Dann definiert $f(x) := \sum_{n=0}^{\infty} a_n(x - x_0)^n$ auf $B(x_0, R)$ eine differenzierbare Funktion und es gilt:

$$f'(x) = \sum_{n=1}^{\infty} n a_n(x - x_0)^{n-1}.$$

Außerdem hat $\sum_{n=1}^{\infty} n a_n(x - x_0)^{n-1}$ denselben Konvergenzradius wie $\sum_{n=0}^{\infty} a_n(x - x_0)^n$.

Beweis. Wir zeigen zunächst, dass f in x_0 differenzierbar ist. Da

$$f(x) = a_0 + \sum_{n=1}^{\infty} a_n(x - x_0)^n = f(x_0) + \left(\sum_{n=1}^{\infty} a_n(x - x_0)^{n-1} \right)(x - x_0),$$

und da Potenzreihen stetig sind, hat der Differenzenquotient von f in x_0 die stetige Fortsetzung

$$Q(x) = \sum_{n=1}^{\infty} a_n(x - x_0)^{n-1}$$

und es gilt $f'(x_0) = Q(x_0) = a_1$. Ist $x_1 \in B(x_0, R) \setminus \{x_0\}$, so folgt aus Lemma 4.5.8 für x mit $|x - x_1| < R - |x_1 - x_0|$ die Identität $f(x) = \sum_{k=0}^{\infty} b_k(x - x_1)^k$. Somit erhalten wir

$$f'(x_1) = b_1 = \sum_{n=1}^{\infty} a_n \binom{n}{1} (x_1 - x_0)^{n-1} = \sum_{n=1}^{\infty} a_n n (x_1 - x_0)^{n-1}.$$

Die Potenzreihe $\sum_{n=1}^{\infty} n a_n(x - x_0)^{n-1}$ hat den Konvergenzradius

$$R' := \frac{1}{\lim_{n \rightarrow \infty} \sqrt[n]{n} \sqrt[n]{|a_n|}} = \frac{1}{\lim_{n \rightarrow \infty} \sqrt[n]{|a_n|}} = R.$$

\square

Bemerkung. Durch iteratives Anwenden der Differentiationsformel erhalten wir:

$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)\dots(n-k+1)a_n(x-x_0)^{n-k}$$

In x_0 sind, außer für $n = k$, alle Summanden gleich 0, also gilt

$$f^{(k)}(x_0) = k! a_k.$$

Insbesondere stimmt die Potenzreihe mit ihrer Taylorreihe überein.

Definition 4.5.10. Sei $D \subset \mathbb{R}$ eine offene Menge. Dann heißt $f : D \rightarrow \mathbb{R}$ *analytisch*, falls sich f in jedem Punkt $x_0 \in D$ in eine Potenzreihe entwickeln lässt, d.h. falls es eine Potenzreihe $\sum_{k=0}^{\infty} a_k(x-x_0)^k$ mit Konvergenzradius $r > 0$ gibt, so dass $f(x) = \sum_{k=0}^{\infty} a_k(x-x_0)^k$ für alle $x \in B(x_0, r)$.

Bemerkung. Wegen Lemma 4.5.8 sind Potenzreihen analytisch.

4.6 Binomialreihen

Definition 4.6.1. Für $\alpha \in \mathbb{R}$ heißt die Funktion $f : \mathbb{R}_+ \rightarrow \mathbb{R}$ mit

$$f(x) = x^\alpha := \exp(\alpha \log x)$$

die *allgemeine Potenzfunktion*.

Bemerkungen.

(a) Es gelten folgende Rechenregeln. Sind $\alpha, \beta \in \mathbb{R}$ und $x > 0, y > 0$, so folgt

$$x^\alpha y^\alpha = (xy)^\alpha$$

sowie

$$x^{\alpha+\beta} = x^\alpha x^\beta, \quad x^{\alpha\beta} = (x^\alpha)^\beta.$$

(b) Die verallgemeinerte Potenzfunktion ist unendlich oft differenzierbar und es gilt:

$$\begin{aligned} f'(x) &= (x^\alpha)' = \exp(\alpha \log x) \cdot \frac{\alpha}{x} = \alpha \exp(\alpha \log x) \exp(-\log x) \\ &= \alpha \exp(\alpha \log x - \log x) = \alpha \exp((\alpha - 1) \log x) = \alpha x^{\alpha-1}. \end{aligned}$$

Durch Induktion erhält man:

$$f^{(k)}(x) = \alpha(\alpha-1)\dots(\alpha-k+1)x^{\alpha-k}.$$

(c) Betrachten wir die auf $(-1, \infty)$ definierte Funktion $g(x) = (1+x)^\alpha = f(1+x)$, so folgt aus der Kettenregel:

$$g^{(k)}(x) = f^{(k)}(1+x) = \alpha(\alpha-1)\dots(\alpha-k+1)(1+x)^{\alpha-k}.$$

Definieren wir für $\alpha \in \mathbb{R}, k \in \mathbb{N}_0$

$$\binom{\alpha}{0} := 1, \quad \binom{\alpha}{k} := \frac{\alpha(\alpha-1)\dots(\alpha-k+1)}{k!},$$

so ist die Taylorreihe von $g(x) = (1+x)^\alpha$ in 0 durch

$$\sum_{k=0}^{\infty} \binom{\alpha}{k} x^k$$

gegeben. Diese Reihe heißt auch *Binomialreihe* und die Koeffizienten $\binom{\alpha}{k}$ heißen *verallgemeinerte Binomialkoeffizienten*. Für $\alpha \in \mathbb{N}$ stimmen diese Koeffizienten mit denen in Definition 1.4.9 definierten Binomialkoeffizienten überein. Für $\alpha \in \mathbb{N}$ folgt aus der Binomialformel:

$$(1+x)^\alpha = \sum_{k=0}^{\alpha} \binom{\alpha}{k} x^k = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k,$$

denn wie schon bemerkt wurde, ist dann $\binom{\alpha}{k} = 0$ für $k > \alpha$. Wir wollen nun eine entsprechende Formel für beliebiges α herleiten.

Satz 4.6.2. Sei $\alpha \in \mathbb{R}$, so gilt für alle $x \in \mathbb{R}$ mit $|x| < 1$:

$$(1+x)^\alpha = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k.$$

Beweis. Wie oben bemerkt wurde, hat die Taylorreihe von $(1+x)^\alpha$ mit Entwicklungspunkt 0 die Gestalt

$$\sum_{k=0}^{\infty} \binom{\alpha}{k} x^k.$$

Wir wollen zeigen, dass diese Reihe für alle $x \in (-1, 1)$ gegen $g(x) = (1+x)^\alpha$ konvergiert. Dazu betrachten wir die Restglieddarstellung von Cauchy

$$(1+x)^\alpha - \sum_{k=0}^n \binom{\alpha}{k} x^k = R_{n+1}(x, 0) = \frac{g^{(n+1)}(\xi)}{n!} (x-\xi)^n \cdot x,$$

wobei $\xi \in \overset{\circ}{I}_{0,x}$. Da

$$\frac{g^{(n+1)}(\xi)}{n!} = (n+1) \frac{g^{(n+1)}(\xi)}{(n+1)!} = (n+1) \binom{\alpha}{n+1} (1+\xi)^{\alpha-(n+1)},$$

erhalten wir für $\delta \in (0, 1)$ mit $\xi = \delta x$:

$$\begin{aligned} R_{n+1}(x, 0) &= (n+1) \binom{\alpha}{n+1} (1+\delta x)^{\alpha-(n+1)} \cdot (x-\delta x)^n \cdot x \\ &= (n+1) \binom{\alpha}{n+1} (1+\delta x)^{\alpha-1} \frac{(1-\delta)^n}{(1+\delta x)^n} x^{n+1}. \end{aligned}$$

Da $|x| < 1$ und $\delta \in (0, 1)$, ist $0 < 1 - \delta \leq 1 + \delta x$ und somit $0 < \left(\frac{1-\delta}{1+\delta x}\right)^n \leq 1$. Außerdem ist die Funktion $\delta \rightarrow (1 + \delta x)^{\alpha-1}$ auf dem Definitionsbereich $(0, 1)$ monoton. Daher liegt $(1 + \delta x)^{\alpha-1}$ zwischen 1 und $(1 + x)^{\alpha-1}$. Insgesamt erhalten wir die Abschätzung

$$\begin{aligned} |R_{n+1}(x, 0)| &\leq (n+1) \left| \binom{\alpha}{n+1} \right| (1 + \delta x)^{\alpha-1} |x|^{n+1} \\ &\leq (n+1) \left| \binom{\alpha}{n+1} \right| A |x|^{n+1} =: b_{n+1}, \end{aligned}$$

wobei $A = \max\{1, (1+x)^{\alpha-1}\}$. Wir zeigen nun, dass b_n für alle $|x| < 1$ eine Nullfolge ist. Dazu betrachte

$$\begin{aligned} \frac{b_{n+1}}{b_n} &= \frac{n+1}{n} \left| \frac{\binom{\alpha}{n+1}}{\binom{\alpha}{n}} \right| |x| \\ &= \frac{n+1}{n} \frac{|\alpha(\alpha-1)\dots(\alpha-n)|}{(n+1)!} \frac{n!|x|}{|\alpha(\alpha-1)\dots(\alpha-n+1)|} \\ &= \frac{1}{n} |\alpha-n| |x| = \left| \frac{\alpha}{n} - 1 \right| |x| \end{aligned}$$

Daraus folgt: $\lim_{n \rightarrow \infty} \frac{b_{n+1}}{b_n} = |x| < 1$. Somit existiert ein $n_0 \in \mathbb{N}$, so dass $\left| \frac{b_{n+1}}{b_n} \right| \leq q < 1$ für alle $n \geq n_0$. Insbesondere gilt $0 < b_{n_0+k} \leq q^k \cdot b_{n_0}$, und b_n ist eine Nullfolge. Wegen

$$|R_{n+1}(x, 0)| \leq b_{n+1}$$

ist auch das Restglied eine Nullfolge und es gilt die Behauptung. \square

Beispiele.

(a) Ist $\alpha = n \in \mathbb{N}_0$, so ist $\binom{n}{k} = 0$ für $k > n$, und wir erhalten

$$(1+x)^n = \sum_{k=0}^{\infty} \binom{n}{k} x^k = \sum_{k=0}^n \binom{n}{k} x^k$$

(b) Ist $\alpha = -1$, so ist

$$\binom{-1}{k} = \frac{(-1) \cdot \dots \cdot (-k)}{1 \cdot \dots \cdot k} = (-1)^k,$$

und wir erhalten für $|x| < 1$ die geometrische Reihe

$$\frac{1}{1+x} = \sum_{k=0}^{\infty} (-1)^k x^k.$$

(c) Für $\alpha = \frac{1}{2}$ erhalten wir $\binom{\frac{1}{2}}{0} = 1$ und

$$\begin{aligned} \binom{\frac{1}{2}}{k} &= \frac{\frac{1}{2} \left(\frac{1}{2} - 1\right) \cdot \dots \cdot \left(\frac{1}{2} - k + 1\right)}{1 \cdot \dots \cdot k} = \frac{1(-1)(-3) \cdot \dots \cdot (3-2k)}{2 \cdot 4 \cdot \dots \cdot 2k} \\ &= (-1)^{k-1} \frac{1 \cdot 1 \cdot 3 \cdot \dots \cdot |(2k-3)|}{2 \cdot 4 \cdot \dots \cdot 2k}. \end{aligned}$$

Daher gilt:

$$\begin{aligned}\sqrt{1+x} &= 1 + \sum_{k=1}^{\infty} (-1)^{k-1} \frac{1 \cdot 1 \cdot 3 \cdot \dots \cdot |(2k-3)|}{2 \cdot 4 \cdot 6 \cdot \dots \cdot (2k)} x^k \\ &= 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \frac{1}{16}x^3 - \frac{5}{128}x^4 + R_5(x, 0).\end{aligned}$$

Benutzen wir die Restglieddarstellung von Lagrange, so erhalten wir:

$$R_5(x, 0) = \binom{\frac{1}{2}}{5} (1 + \xi)^{-\frac{9}{2}} x^5 = \frac{7}{256} (1 + \xi)^{-\frac{9}{2}} x^5,$$

wobei ξ eine Zahl zwischen 0 und x darstellt. Diese Beziehung kann man zum Beispiel verwenden um $\sqrt{2}$ näherungsweise zu berechnen.

$$\begin{aligned}\sqrt{2} &= \sqrt{\frac{9}{4} \cdot \frac{8}{9}} = \frac{3}{2} \sqrt{1 - \frac{1}{9}} \\ &= \frac{3}{2} \left(1 - \frac{1}{2} \cdot \frac{1}{9} - \frac{1}{8} \cdot \frac{1}{9^2} - \frac{1}{16} \cdot \frac{1}{9^3} - \frac{5}{128} \cdot \frac{1}{9^4} \right) + \frac{3}{2} R_5 \left(-\frac{1}{9}, 0 \right) \\ &= \frac{791779}{559872} + \frac{3}{2} R_5 \left(-\frac{1}{9}, 0 \right) = 1,4142143\dots + \frac{3}{2} R_5 \left(-\frac{1}{9}, 0 \right)\end{aligned}$$

wobei

$$\left| \frac{3}{2} R_5 \left(-\frac{1}{9}, 0 \right) \right| \leq \frac{3}{2} \frac{7}{256} \left(\frac{9}{8} \right)^{\frac{9}{2}} \frac{1}{9^5} < \frac{21}{2 \cdot 256 \cdot 8^5} < 1.26 \cdot 10^{-6}.$$

Also gilt:

$$|\sqrt{2} - 1.4142143| \leq \left| \sqrt{2} - \frac{791779}{559872} \right| + \left| \frac{791779}{559872} - 1.4142143 \right| \leq 1.26 \cdot 10^{-6} + 10^{-7} = 1.36 \cdot 10^{-6}.$$

Die Überprüfung mit Hilfe des Taschenrechners liefert: $\sqrt{2} = 1.4142135\dots$. Dies bestätigt die obige Abschätzung.

(d) Für $\alpha = -\frac{1}{2}$ erhalten wir $\binom{-\frac{1}{2}}{0} = 1$ und

$$\begin{aligned}\binom{-\frac{1}{2}}{k} &= \frac{-\frac{1}{2}(-\frac{1}{2}-1) \cdot \dots \cdot (-\frac{1}{2}-k+1)}{1 \cdot \dots \cdot k} \\ &= (-1)^k \cdot \frac{1 \cdot 3 \cdot \dots \cdot (2k-1)}{2 \cdot 4 \cdot \dots \cdot 2k}.\end{aligned}$$

Daher gilt:

$$\begin{aligned}\frac{1}{\sqrt{1+x}} &= 1 + \sum_{k=1}^{\infty} (-1)^k \frac{1 \cdot 3 \cdot \dots \cdot (2k-1)}{2 \cdot 4 \cdot \dots \cdot 2k} x^k \\ &= 1 - \frac{1}{2}x + \frac{3}{8}x^2 - \frac{5}{16}x^3 + \dots\end{aligned}$$

Weitere Anwendungen der Binomialreihe sind die Potenzreihenentwicklungen der Arcusfunktionen

$$\arcsin : (-1, 1) \rightarrow \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$$

und

$$\arctan : \mathbb{R} \rightarrow \left(-\frac{\pi}{2}, \frac{\pi}{2}\right).$$

Aus $(f^{-1})'(y) = \frac{1}{f'(f^{-1}(y))}$ folgt:

$$\arcsin'(y) = \frac{1}{\cos(\arcsin y)} = \frac{1}{\sqrt{1 - \sin^2(\arcsin y)}} = \frac{1}{\sqrt{1 - y^2}}.$$

Genauso berechnet man:

$$\arctan'(y) = \frac{1}{1 + y^2}.$$

Satz 4.6.3. Für alle $y \in \mathbb{R}$ mit $|y| < 1$ gelten folgende Potenzreihendarstellungen:

$$\begin{aligned} \arcsin y &= \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} \binom{-\frac{1}{2}}{k} y^{2k+1} = y + \frac{1}{2} \frac{y^3}{3} + \frac{1 \cdot 3}{2 \cdot 4} \frac{y^5}{5} + \frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6} \frac{y^7}{7} + \dots, \\ \arctan y &= \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} y^{2k+1} = y - \frac{y^3}{3} + \frac{y^5}{5} - \frac{y^7}{7} + \dots \end{aligned}$$

Beweis. Es gilt:

$$\arcsin' y = \frac{1}{\sqrt{1 - y^2}} = (1 - y^2)^{-1/2} = \sum_{k=0}^{\infty} \binom{-\frac{1}{2}}{k} (-1)^k y^{2k}$$

für y mit $|y| < 1$. Die Reihe

$$\sum_{k=0}^{\infty} \binom{-\frac{1}{2}}{k} \frac{1}{2k+1} (-1)^k y^{2k+1} =: F(y)$$

hat den gleichen Konvergenzradius wie die obige Reihe. Da man wegen Satz 4.5.9 Potenzreihen gliedweise differenzieren darf, erhalten wir: $F'(y) - \arcsin' y = 0$. Dies impliziert $F(y) = \arcsin y + a$. Die Konstante a ist wegen $F(0) = 0$ und $\arcsin(0) = 0$ gleich 0.

Genauso folgt:

$$\arctan'(y) = \frac{1}{1 + y^2} = \sum_{k=0}^{\infty} (-1)^k y^{2k}$$

und somit

$$\arctan y = \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} y^{2k+1} + a.$$

Da $\arctan(0) = 0$, ist auch $a = 0$. □

Außerdem lässt sich auf ähnliche Weise eine Potenzreihenentwicklung des Logarithmus angeben. Ist $f(x) = \log(1 + x)$, so ist $f'(x) = \frac{1}{x+1}$. Also folgt wie oben, unter Benutzung des Beispiels (b):

Satz 4.6.4. Für alle $x \in \mathbb{R}$ mit $|x| < 1$ gilt:

$$\log(1+x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} x^{k+1}.$$

Der folgende Satz zeigt, dass die Reihendarstellungen des Logarithmus als auch des Arcustangens für $x = 1$ gelten.

Satz 4.6.5 (Abelscher Grenzwertsatz).

Sei $\sum_{k=0}^{\infty} a_k x^k$ eine reelle Potenzreihe mit Konvergenzradius $R > 0$. Konvergiert die Reihe auch für $x = R$, so ist die Funktion $f : (-R, R] \rightarrow \mathbb{R}$ mit

$$f(x) = \sum_{k=0}^{\infty} a_k x^k$$

auch im Punkte $x = R$ stetig.

Bemerkungen.

- (a) Wir hatten schon gesehen, dass eine Potenzreihe innerhalb ihres Konvergenzkreises stetig ist. Der Abelsche Grenzwertsatz zeigt, dass die Stetigkeit auch noch in den Randpunkten gilt, falls die Reihe dort konvergiert.
- (b) Die Logarithmusreihe $\sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} x^{k+1}$ als auch die Arcustangensreihe $\sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} y^{2k+1}$ konvergieren wegen des Leibnizkriteriums 2.5.5 für $x = 1$. Unter Verwendung des Abelschen Grenzwertsatzes erhält man somit

$$\log(2) = \lim_{x \rightarrow 1} \log(1+x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1}$$

und

$$\arctan(1) = \lim_{x \rightarrow 1} \arctan(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1}.$$

Wegen $\tan(\frac{\pi}{4}) = 1$ erhalten wir $\arctan(1) = \frac{\pi}{4}$ und somit folgt:

$$\frac{\pi}{4} = \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1}.$$

Kapitel 5

Integralrechnung

Wir wollen uns nun der Integralrechnung zuwenden. In der Integralrechnung geht es darum, Flächeninhalte oder allgemeine Volumina in höherdimensionalen Räumen zu berechnen. Die einfachsten Funktionen, die wir integrieren können, sind Treppenfunktionen.

5.1 Treppenfunktionen

Definition 5.1.1. Sei $a, b \in \mathbb{R}$ und $a < b$. Eine Funktion $\varphi : [a, b] \rightarrow \mathbb{R}$ heißt *Treppenfunktion*, falls eine Unterteilung

$$a = x_0 < x_1 < \dots < x_n = b$$

existiert, so dass φ auf jedem offenen Teilintervall (x_{k-1}, x_k) konstant ist. Mit $T([a, b])$ bezeichnen wir die Menge der Treppenfunktionen (s. Abbildung 5.1).

Bemerkung.

- (a) Die Funktionswerte auf den Unterteilungspunkten $\{x_0, \dots, x_k\}$ unterliegen keiner Einschränkung.
- (b) Verschiedene Unterteilungen können durchaus gleiche Treppenfunktionen beschreiben.
Wir können z.B. Unterteilungspunkte hinzufügen, ohne, daß die Funktion sich ändert.

Satz 5.1.2. Sind $\varphi, \psi \in T([a, b])$ und $c \in \mathbb{R}$, so sind auch $\varphi + \psi$ als auch $\varphi \cdot \psi$ Treppenfunktionen.

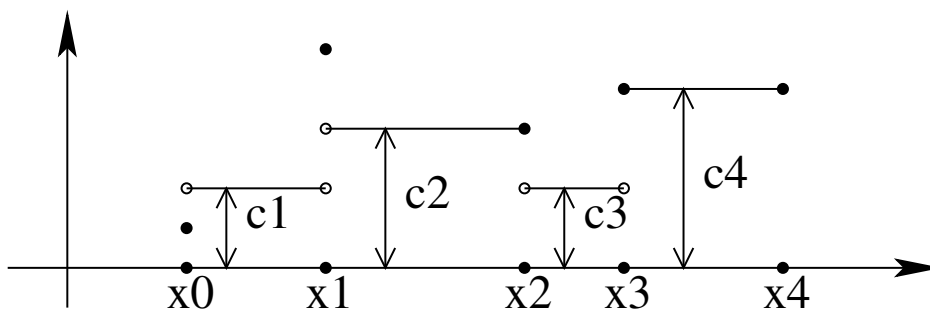


Abbildung 5.1:

Bemerkung. Insbesondere ist für jedes $c \in \mathbb{R}$ und jedes $\varphi \in T([a, b])$ auch $c \cdot \varphi \in T([a, b])$.

Beweis. Wähle die Unterteilung $a = z_0 < \dots < z_m = b$, so dass die Menge $\{z_0, \dots, z_m\}$ aus der Vereinigung der zu φ und ψ gehörigen Unterteilungspunkte besteht. Also sind $\varphi + \psi$ und $\varphi \cdot \psi$ auf jedem der Teilintervalle (z_i, z_{i+1}) konstant und somit wieder Treppenfunktionen. \square

Nun definieren wir das Integral einer Treppenfunktion.

Definition 5.1.3. Sei $\varphi \in T([a, b])$ und $a = x_0 < x_1 < \dots < x_n = b$ eine zu φ gehörige Unterteilung von $[a, b]$ mit $c_k = \varphi(x)$ für $x \in (x_{k-1}, x_k)$. Dann heißt

$$\int \varphi := \sum_{k=1}^n c_k (x_k - x_{k-1})$$

das *Integral* von φ .

Bemerkungen. Ist $\varphi(x) \geq 0$ für $x \in [a, b]$, so beschreibt $\int \varphi$ den Flächeninhalt unterhalb des Graphen von φ .

Die Definition hängt nicht von der Wahl der Unterteilung ab, denn das Hinzufügen von Unterteilungspunkten ändert den Wert der Integralsumme nicht. Hat man 2 Unterteilungen, so betrachte die Unterteilung, die aus der Vereinigung der Punkte beider Unterteilungen besteht. Die zugehörige Integralsumme stimmt dann mit den "Integralsummen" beider Unterteilungen überein.

Satz 5.1.4. Seien $\varphi, \psi \in T([a, b])$ Treppenfunktionen, so gilt:

$$(a) \int \lambda_1 \varphi + \lambda_2 \psi = \lambda_1 \int \varphi + \lambda_2 \int \psi, \text{ für alle } \lambda_1, \lambda_2 \in \mathbb{R} \quad (\text{Linearität})$$

$$(b) \text{ Ist } \varphi(x) \leq \psi(x) \text{ für alle } x \in [a, b], \text{ so folgt } \int \varphi \leq \int \psi \quad (\text{Monotonie})$$

$$(c) \left| \int \varphi \right| \leq \int |\varphi| \leq (b - a) \sup\{|\varphi(x)| \mid x \in [a, b]\} \quad (\text{Beschränktheit})$$

Beweis. Seien $\varphi, \psi \in T([a, b])$ Treppenfunktionen mit gemeinsamer Unterteilung $a = x_0 < x_1 < \dots < x_n = b$ und $\varphi(x) = c_k, \psi(x) = d_k$ für $x \in (x_{k-1}, x_k)$. Dann ist

$$\begin{aligned} \int \lambda_1 \varphi + \lambda_2 \psi &= \sum_{k=1}^n (\lambda_1 c_k + \lambda_2 d_k) (x_k - x_{k-1}) = \lambda_1 \sum_{k=1}^n c_k (x_k - x_{k-1}) + \lambda_2 \sum_{k=1}^n d_k (x_k - x_{k-1}) \\ &= \lambda_1 \int \varphi + \lambda_2 \int \psi. \end{aligned}$$

Gilt darüberhinaus für das obige Paar φ, ψ von Treppenfunktionen zusätzlich $\varphi(x) \leq \psi(x)$ für alle $x \in [a, b]$, so ist $c_k \leq d_k$ für alle $k \in \{1, \dots, n\}$ und somit $\int \varphi \leq \int \psi$.

Teil (c) folgt aus

$$\begin{aligned} \left| \int \varphi \right| &= \left| \sum_{k=1}^n c_k (x_k - x_{k-1}) \right| \leq \sum_{k=1}^n |c_k| (x_k - x_{k-1}) = \int |\varphi| \\ &\leq \sum_{k=1}^n \max\{|c_k| \mid k \in \{1, \dots, n\}\} (x_k - x_{k-1}) \leq (b - a) \sup\{|\varphi(x)| \mid x \in [a, b]\}. \end{aligned}$$

\square

Nun wollen wir den Integralbegriff auf eine größere Funktionenklasse ausdehnen. Dazu führen wir die sogenannte Supremumsnorm auf der Menge der beschränkten Funktionen ein.

Definition 5.1.5. Sei $D \subset \mathbb{R}$ eine Menge und $B(D) = \{f \mid f : D \rightarrow \mathbb{R} \text{ beschränkt}\}$ die Menge der auf D beschränkten Funktionen. Dann heißt

$$\|f\| = \sup\{|f(x)| \mid x \in D\}$$

die *Supremumsnorm* von f .

Satz 5.1.6. Die Supremumsnorm hat folgende Eigenschaften. Für alle $f, g \in B(D)$ und $\lambda \in \mathbb{R}$ gilt:

- (1) $\|f\| = 0$ genau dann, falls $f = 0$,
- (2) $\|\lambda f\| = |\lambda| \|f\|$,
- (3) $\|f + g\| \leq \|f\| + \|g\|$ (Dreiecksungleichung).

Beweis. Der Beweis folgt aus den entsprechenden Eigenschaften des Betrages reeller Zahlen. \square

Bemerkungen.

- (a) Sind $f, g \in B(D)$, so misst $\|f - g\|$ den Abstand der Funktionen f, g .
- (b) Die Abschätzung (c) in Satz 5.1.4 läßt sich auch wie folgt schreiben. Ist $\varphi \in T([a, b])$ eine Treppenfunktion, so gilt:

$$\left| \int \varphi \right| \leq \int |\varphi| \leq (b - a) \|\varphi\|.$$

Wir wollen nun das Integral auf die Menge der Funktionen ausdehnen, die sich bezüglich der Supremumsnorm beliebig gut durch Treppenfunktionen approximieren lassen. Diese Funktionen werden auch Regelfunktionen genannt.

Definition 5.1.7. Eine Funktion $f : [a, b] \rightarrow \mathbb{R}$ heißt *Regelfunktion*, wenn für jedes $\epsilon > 0$ eine Treppenfunktion $\varphi \in T([a, b])$ existiert mit

$$\|f - \varphi\| \leq \epsilon.$$

Die Menge der Regelfunktionen auf dem Intervall $[a, b]$ wird auch mit $R([a, b])$ bezeichnet. Man sagt auch, dass $R([a, b])$ aus allen Funktionen besteht, die sich gleichmäßig durch Treppenfunktionen approximieren lassen.

Bemerkungen.

- (a) Eine Funktion $f : [a, b] \rightarrow \mathbb{R}$ ist also genau dann eine Regelfunktion, wenn eine Folge φ_n , $n \in \mathbb{N}$ von Treppenfunktionen in $T([a, b])$ existiert mit

$$\lim_{n \rightarrow \infty} \|f - \varphi_n\| = 0.$$

- (b) Sind f, g Regelfunktionen, so sind auch $f + g$ und $f \cdot g$ wieder Regelfunktionen. Mit anderen Worten ist $R([a, b])$ unter Addition und Multiplikation abgeschlossen.

Wir zeigen nun, dass $R([a, b])$ den Raum der auf $[a, b]$ stetigen Funktionen umfasst.

Satz 5.1.8. Die Menge der auf $[a, b]$ stetigen Funktionen $C^0([a, b])$ ist in $R([a, b])$ enthalten, d.h. jede stetige Funktion ist auch eine Regelfunktion.

Beweis. Sei $f \in C^0([a, b])$. Wir müssen zeigen: für jedes $\epsilon > 0$ existiert ein $\varphi \in T([a, b])$ mit $\|f - \varphi\| = \sup\{|f(x) - \varphi(x)| \mid x \in [a, b]\} < \epsilon$.

Da f auf dem kompakten Intervall $[a, b]$ stetig ist, ist f auch gleichmäßig stetig, d.h. für jedes $\epsilon > 0$ existiert ein $\delta > 0$ mit $|f(x) - f(y)| < \epsilon$, falls $|x - y| < \delta$.

Wähle eine Unterteilung $a = x_0 < x_1 < \dots < x_n = b$ mit $|x_k - x_{k-1}| < \delta$ für alle $k \in \{1, \dots, n\}$. Definiere $\varphi \in T([a, b])$ durch $\varphi(x) = f(x_k)$, falls $x \in (x_{k-1}, x_k]$. Dann gilt für $k \in \{1, \dots, n\}$ und $x \in (x_{k-1}, x_k]$:

$$|f(x) - \varphi(x)| = |f(x) - f(x_k)| < \epsilon,$$

d.h. $\|f - \varphi\| < \epsilon$. □

Bemerkung. Man kann zeigen: $R([a, b])$ besteht genau aus den Funktionen, die für jedes $x \in [a, b]$ einen rechtsseitigen und linksseitigen Grenzwert besitzen (siehe Definition 3.4.10). Für den Beweis siehe Königsberger Band 1, Seiten 193-195.

Nun können wir das Integral auf $R([a, b])$ wie folgt erklären:

Definition 5.1.9. Sei $f \in R([a, b])$. Dann definiere:

$$\int f := \lim_{n \rightarrow \infty} \int \varphi_n,$$

wobei $\varphi_n \in T([a, b])$ eine Folge von Treppenfunktionen ist, mit $\lim_{n \rightarrow \infty} \|\varphi_n - f\| = 0$.

Bemerkungen.

- (a) Der Limes $\lim_{n \rightarrow \infty} \int \varphi_n$ existiert, denn

$$\left| \int \varphi_n - \int \varphi_m \right| = \left| \int (\varphi_n - \varphi_m) \right| \leq (b-a) \|\varphi_n - \varphi_m\| \leq (b-a) (\|\varphi_n - f\| + \|f - \varphi_m\|).$$

Damit ist $\int \varphi_n$ eine Cauchyfolge reeller Zahlen und somit konvergent.

- (b) Die Definition hängt nicht von der Wahl der Folge φ_n ab, denn sind $\varphi_n, \psi_n \in T(I)$ Treppenfunktionen mit $\|\varphi_n - f\|, \|\psi_n - f\| \rightarrow 0$ für $n \rightarrow \infty$, so gilt:

$$\left| \int \varphi_n - \int \psi_n \right| \leq (b-a) \|\varphi_n - \psi_n\| \leq (b-a) (\|\varphi_n - f\| + \|f - \psi_n\|) \rightarrow 0.$$

- (c) Der Satz 5.1.4 überträgt sich unmittelbar auf Regelfunktionen (siehe Königsberger, Band 1, Seite 197).

Wie kann man $\int f$ berechnen? Die wichtigste Methode ist mittels des Hauptsatzes der Differential- und Integralrechnung, den wir jetzt besprechen werden.

5.2 Hauptsatz der Differential- und Integralrechnung

Ist $a < b$ und $f \in R([a, b])$, so schreiben wir auch $\int_a^b f(t)dt$ statt $\int f$. Ist $c \in (a, b)$, so folgt aus der Definition des Integrals:

$$\int_a^b f(t)dt = \int_a^c f(t)dt + \int_c^b f(t)dt. \quad (*)$$

Diese Formel bleibt für eine beliebige Anordnung von Zahlen a, b, c richtig, falls wir folgendes definieren.

$$\int_a^a f(t)dt = 0 \quad \text{und} \quad \int_a^b f(t)dt := - \int_b^a f(t)dt, \quad \text{falls } b < a.$$

Hierbei ist f eine Regelfunktion auf einem kompakten Intervall, welches a, b und c enthält.

Satz 5.2.1 (*Hauptsatz der Differential- und Integralrechnung*).

Sei $f : [a, b] \rightarrow \mathbb{R}$ eine stetige Funktion. Sei $F(x) := \int_a^x f(t)dt$ für $x \in [a, b]$.

Dann ist $F : [a, b] \rightarrow \mathbb{R}$ eine differenzierbare Funktion und es gilt: $F'(x) = f(x)$ für alle $x \in [a, b]$.

Beweis. Wir müssen zeigen: ist $x_0 \in [a, b]$, so ist

$$Q(x) := \begin{cases} \frac{F(x) - F(x_0)}{x - x_0} & \text{für } x \neq x_0 \\ f(x_0) & \text{für } x = x_0 \end{cases}$$

stetig in x_0 , d.h.

$$\forall \epsilon > 0 \quad \exists \delta > 0 \quad \forall x \in [a, b] : |x - x_0| < \delta \Rightarrow |Q(x) - f(x_0)| < \epsilon.$$

Ist $x \neq x_0$, so folgt unter Benutzung von (*):

$$\begin{aligned} \left| \frac{F(x) - F(x_0)}{x - x_0} - f(x_0) \right| &= \left| \frac{\int_a^x f(t)dt - \int_a^{x_0} f(t)dt}{x - x_0} - f(x_0) \frac{\int_{x_0}^x 1dt}{x - x_0} \right| \\ &= \frac{\left| \int_{x_0}^x f(t)dt - \int_{x_0}^x f(x_0)dt \right|}{|x - x_0|} = \frac{\left| \int_{x_0}^x f(t) - f(x_0)dt \right|}{|x - x_0|} \\ &\leq |x - x_0| \frac{\sup\{|f(t) - f(x_0)| \mid t \in I_{x_0, x}\}}{|x - x_0|} \\ &= \sup\{|f(t) - f(x_0)| \mid t \in I_{x_0, x}\} \end{aligned}$$

wobei

$$I_{x_0, x} = \begin{cases} [x_0, x] & \text{falls } x \geq x_0, \\ [x, x_0] & \text{falls } x < x_0. \end{cases}$$

Sei $\epsilon > 0$. Wegen der Stetigkeit von f existiert ein $\delta > 0$ mit $|f(t) - f(x_0)| < \epsilon$, falls $|t - x_0| < \delta$. Ist $|x - x_0| < \delta$, so folgt $\sup\{|f(t) - f(x_0)| \mid t \in I_{x_0, x}\} < \epsilon$. Also folgt für alle $x \in [a, b]$ mit $|x - x_0| < \delta$: $|Q(x) - f(x_0)| < \epsilon$, d.h. Q ist stetig in x_0 . \square

Definition 5.2.2. Sei $f : D \rightarrow \mathbb{R}$, $D \subset \mathbb{R}$ eine Funktion. Eine Funktion $F : D \rightarrow \mathbb{R}$ heißt *Stammfunktion* von f , falls $F'(x) = f(x)$, für alle $x \in D$.

Satz 5.2.3. Ist I ein beliebiges Intervall und $F, G : I \rightarrow \mathbb{R}$ Stammfunktionen von $f : I \rightarrow \mathbb{R}$. Dann existiert eine Konstante $c \in \mathbb{R}$ mit $F = G + c$.

Beweis. Da F, G Stammfunktionen von f sind, folgt:

$$(F - G)' = F' - G' = f - f = 0.$$

Da I ein Intervall ist, folgt mit Hilfe des Mittelwertsatzes 4.4.5, dass $F - G$ eine konstante Funktion ist. \square

Bemerkung. Ist $f : [a, b] \rightarrow \mathbb{R}$ stetig, so existiert wegen des Hauptsatzes eine Stammfunktion, nämlich $F(x) = \int_a^x f(t) dt$. Aus dem Hauptsatz ergibt sich zusammen mit Satz 5.2.3 folgender wichtiger Satz:

Satz 5.2.4. Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig und F eine beliebige Stammfunktion. Dann gilt:

$$\int_a^b f(t) dt = F(b) - F(a) =: F(x) \Big|_a^b$$

Beweis. Wegen des Hauptsatzes ist $\int_a^x f(t) dt$ eine Stammfunktion von f . Wegen Satz 5.2.3 existiert daher eine Konstante $c \in \mathbb{R}$ mit

$$\int_a^x f(t) dt = F(x) + c.$$

Setzen wir $x = a$, so erhalten wir: $c = -F(a)$. Daher ist

$$\int_a^b f(t) dt = F(b) - F(a).$$

\square

Beispiel.

f	F (Stammfunktion von f)
$x^\alpha, \alpha \neq -1$	$\frac{1}{\alpha+1} x^{\alpha+1}$
$\frac{1}{x}$	$\log x$
e^{ax}	$\frac{1}{a} e^{ax}$
$\cos x$	$\sin x$
$\sin x$	$-\cos x$
$\frac{1}{\cos^2 x}$	$\tan x$
$\frac{1}{\sin^2 x}$	$-\cot x$
$\frac{1}{\sqrt{1-x^2}}$	$\arcsin x$
$\sum_{k=0}^{\infty} a_k x^k$	$\sum_{k=0}^{\infty} \frac{a_k}{k+1} x^{k+1}$

5.3 Integrationsregeln

Satz 5.3.1 (*Partielle Integration*).

Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig mit einer Stammfunktion F und $g : [a, b] \rightarrow \mathbb{R}$ stetig differenzierbar.

Dann gilt:

$$\int_a^b f(x) \cdot g(x) dx = F \cdot g \Big|_a^b - \int_a^b F(x) g'(x) dx.$$

Beweis. Wegen der Produktregel 4.3.1 gilt:

$$(F \cdot g)' = F' \cdot g + F \cdot g' = f \cdot g + F \cdot g'.$$

Daraus folgt:

$$(F \cdot g)(b) - (F \cdot g)(a) = \int_a^b f(x)g(x)dx + \int_a^b F(x)g'(x)dx.$$

□

Beispiel. Sei $n \geq 2$ und $\sin^n(x) := (\sin x)^n$. Dann erhalten wir durch partielle Integration:

$$\begin{aligned} \int_0^\pi \sin^n(x) dx &= \int_0^\pi \sin(x) \cdot \sin^{n-1}(x) dx \\ &= \underbrace{-\cos(x) \cdot \sin^{n-1}(x)}_{=0} \Big|_0^\pi + (n-1) \int_0^\pi \cos(x) \sin^{n-2}(x) \cdot \cos(x) dx \\ &= (n-1) \int_0^\pi \cos^2(x) \sin^{n-2}(x) dx = (n-1) \int_0^\pi (1 - \sin^2(x)) \sin^{n-2}(x) dx \\ &= (n-1) \int_0^\pi \sin^{n-2}(x) dx - (n-1) \int_0^\pi \sin^n(x) dx. \end{aligned}$$

Daraus folgt:

$$n \int_0^\pi \sin^n(x) dx = (n-1) \int_0^\pi \sin^{n-2}(x) dx :$$

Insbesondere erhalten wir für $n = 2$: $\int_0^\pi \sin^2(x) dx = \frac{1}{2}\pi$.

Eine weitere wichtige Regel ist die *Substitutionsregel*.

Satz 5.3.2. Sei $f : [c, d] \rightarrow \mathbb{R}$ stetig und $g : [a, b] \rightarrow [c, d]$ stetig differenzierbar. Dann gilt:

$$\int_{g(a)}^{g(b)} f(x) dx = \int_a^b f(g(t)) \cdot g'(t) dt.$$

Beweis. Sei F Stammfunktion von f , so folgt aus der Kettenregel:

$$(F \circ g)'(t) = F'(g(t)) \cdot g'(t).$$

Daraus folgt:

$$\begin{aligned} \int_a^b f(g(t))g'(t)dt &= \int_a^b F'(g(t)) \cdot g'(t)dt = \int_a^b (F \circ g)'(t)dt \\ &= F(g(b)) - F(g(a)) = \int_{g(a)}^{g(b)} f(x)dx. \end{aligned}$$

□

5.4 Uneigentliche Integrale

Wir hatten bisher nur Integrale für Regelfunktionen auf kompakten Intervallen erklärt. Der Grund dafür besteht darin, dass auf nicht kompakten Intervallen stetige oder differenzierbare Funktionen im Allgemeinen keine gleichmäßige Approximation durch Treppenfunktionen besitzen (ein Beispiel ist eine auf einem offenen Intervall stetige aber unbeschränkte Funktion). In diesem Falle definieren wir das Integral durch “Ausschöpfung” über kompakte Teilintervalle.

Definition 5.4.1. Sei $-\infty < a < b \leq \infty$ und $f : [a, b) \rightarrow \mathbb{R}$ mit $f \in R([a, c])$ für alle $c < b$. Betrachte für jedes $c < b$ das Integral

$$F(c) := \int_a^c f(x)dx.$$

Existiert der Grenzwert $\lim_{c \rightarrow b} F(c) \in \mathbb{R}$, so heißt dieser Grenzwert das *uneigentliche Integral* von f auf $[a, b]$. Man schreibt

$$\int_a^b f(x)dx := \lim_{c \rightarrow b} \int_a^c f(x)dx.$$

Analog definiert man das Integral im Falle $-\infty \leq a < b < \infty$ und $f : (a, b] \rightarrow \mathbb{R}$ mit $f \in R([c, b])$ für alle $c > a$ als

$$\int_a^b f(x)dx := \lim_{c \rightarrow a} \int_c^b f(x)dx,$$

falls der Grenzwert existiert und endlich ist.

Sei $-\infty \leq a < b \leq \infty$ und $f : (a, b) \rightarrow \mathbb{R}$ mit $f \in R([c, d])$ für alle $a < c < d < b$. Existieren für ein $c \in (a, b)$ die uneigentlichen Integrale $\int_a^c f(x)dx$ und $\int_c^b f(x)dx$, so definiere

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx.$$

Dabei hängt der Wert des Integrals nicht von der Wahl von $c \in (a, b)$ ab.

Beispiele.

1. Sei $\alpha \in \mathbb{R}$. Das Integral $\int_1^\infty \frac{1}{x^\alpha} dx$ existiert genau dann, falls $\alpha > 1$. Denn für $c > 1$ gilt:

$$\begin{aligned} \int_1^c \frac{1}{x^\alpha} dx &= \begin{cases} \left. \frac{1}{1-\alpha} x^{1-\alpha} \right|_1^c & \text{falls } \alpha \neq 1, \\ \log x \Big|_1^c & \text{falls } \alpha = 1, \end{cases} \\ &= \begin{cases} \frac{1}{1-\alpha} (c^{1-\alpha} - 1) & \text{falls } \alpha \neq 1, \\ \log c & \text{falls } \alpha = 1. \end{cases} \end{aligned}$$

Ist $\alpha \leq 1$, so existiert das Integral nicht. Für $\alpha > 1$ ergibt sich:

$$\lim_{c \rightarrow \infty} \int_1^c \frac{1}{x^\alpha} dx = \frac{1}{\alpha - 1}.$$

2. Das Integral $\int_0^1 \frac{1}{x^\alpha}$ existiert genau dann, falls $\alpha < 1$. Denn

$$\int_c^1 \frac{1}{x^\alpha} dx = \begin{cases} \frac{1}{1-\alpha} (1 - c^{1-\alpha}) & \alpha \neq 1, \\ -\log c & \alpha = 1. \end{cases}$$

Also existiert das Integral nicht, falls $\alpha \geq 1$. Für $\alpha < 1$ folgt:

$$\int_0^1 \frac{1}{x^\alpha} dx = \lim_{c \rightarrow 0} \int_c^1 \frac{1}{x^\alpha} dx = \frac{1}{1-\alpha}.$$

3. $\int_0^\infty e^{-\alpha x} dx$ existiert genau dann, falls $\alpha > 0$, denn

$$\int_0^c e^{-\alpha x} dx = -\frac{1}{\alpha} e^{-\alpha x} \Big|_0^c = -\frac{1}{\alpha} (e^{-\alpha c} - 1).$$

Insbesondere gilt, falls $\alpha > 0$:

$$\int_0^\infty e^{-\alpha x} dx = \lim_{c \rightarrow \infty} \int_0^c e^{-\alpha x} dx = \frac{1}{\alpha}.$$

4. $\int_{-\infty}^{+\infty} \frac{dx}{1+x^2} = \pi$, denn

$$\int_0^c \frac{dx}{1+x^2} = \arctan x \Big|_0^c = \arctan c \rightarrow \frac{\pi}{2} \quad \text{für } c \rightarrow +\infty$$

und analog

$$\int_b^0 \frac{dx}{1+x^2} = -\arctan b \rightarrow \frac{\pi}{2} \quad \text{für } b \rightarrow -\infty.$$

Für die Existenz uneigentlicher Integrale ist das folgende Majorantenkriterium nützlich.

Satz 5.4.2. *Es sei $a < b \leq \infty$ und $f : [a, b) \rightarrow \mathbb{R}$ eine Funktion mit $f \in R([a, c])$ für alle $c \in [a, b)$. Es sei $\varphi : [a, b) \rightarrow [0, \infty)$ eine nicht-negative Funktion, für die das Integral $\int_a^b \varphi(x) dx$ existiert. Ist $|f(x)| \leq \varphi(x)$ für alle $x \in [a, b)$, so existiert auch $\int_a^b f(x) dx$.*

Beweis. Es seien $s, t \in [a, b)$, so gilt

$$\begin{aligned} \left| \int_a^t f(x) dx - \int_a^s f(x) dx \right| &= \left| \int_s^t f(x) dx \right| \leq \left| \int_s^t |f(x)| dx \right| \\ &\leq \left| \int_s^t \varphi(x) dx \right| = \left| \int_a^t \varphi(x) dx - \int_a^s \varphi(x) dx \right|. \end{aligned}$$

Ist nun $t_n \in [a, b)$ eine Folge mit $\lim_{n \rightarrow \infty} t_n = b$, so ist $\int_a^{t_n} \varphi(x) dx$ eine Cauchyfolge und somit ist auch $\int_a^{t_n} f(x) dx$ eine Cauchyfolge.

Ist $s_n \in [a, b)$ eine weitere Folge mit $\lim_{n \rightarrow \infty} s_n = b$, so existiert auch der Grenzwert $\int_a^{s_n} f(x) dx$ und wegen

$$\left| \int_a^{t_n} f(x) dx - \int_a^{s_n} f(x) dx \right| \leq \left| \int_a^{t_n} \varphi(x) dx - \int_a^{s_n} \varphi(x) dx \right| \rightarrow 0 \quad \text{für } n \rightarrow \infty$$

stimmen die beiden Grenzwerte überein. □

Beispiel. Betrachte für $\alpha > 0$ das uneigentliche Integral (*Gammafunktion*)

$$\Gamma(\alpha) = \int_0^{\infty} t^{\alpha-1} e^{-t} dt.$$

Dieses Integral besteht aus den uneigentlichen Integralen $\int_0^1 t^{\alpha-1} e^{-t} dt$ und $\int_1^{\infty} t^{\alpha-1} e^{-t} dt$. Beide Integrale existieren, denn $|t^{\alpha-1} e^{-t}| \leq t^{\alpha-1}$ für $t \in (0, 1]$ und wegen des obigen Beispiels 2 existiert $\int_0^1 t^{\alpha-1} dt$, denn $1 - \alpha < 1$. Außerdem existiert eine Konstante $c > 0$ mit $t^{\alpha-1} e^{-t} \leq ce^{-t/2}$ für $t \geq 1$, denn $\lim_{t \rightarrow \infty} t^{\alpha-1} e^{-t/2} = 0$. Also existiert wegen des Beispiels 3 und des Majorantenkriteriums auch das Integral $\int_1^{\infty} t^{\alpha-1} e^{-t} dt$.

Die Gammafunktion $\Gamma : \mathbb{R}_+ \rightarrow \mathbb{R}$ hat folgende Eigenschaften.

Satz 5.4.3.

(a) $\Gamma(\alpha + 1) = \alpha\Gamma(\alpha)$,

(b) $\Gamma(n) = (n - 1)!$ für $n \in \mathbb{N}$.

Beweis.

(a) Sei $0 < \epsilon < R < \infty$, so folgt mittels partieller Integration

$$\int_{\epsilon}^R t^{\alpha} e^{-t} dt = -t^{\alpha} e^{-t} \Big|_{\epsilon}^R + \alpha \int_{\epsilon}^R t^{\alpha-1} e^{-t} dt.$$

Bilden wir die Grenzübergänge $\epsilon \rightarrow 0$ und $R \rightarrow \infty$, so erhalten wir die Behauptung.

(b) Der Beweis von (b) folgt durch Induktion, denn $\Gamma(1) = \int_0^{\infty} e^{-t} dt = 1$.

Ist nun $\Gamma(n) = (n - 1)!$, so folgt mit (a): $\Gamma(n + 1) = n\Gamma(n) = n!$.

□

Bemerkung. Die Gammafunktion ist wegen obiger Eigenschaft eine Interpolation der Funktion $n \rightarrow n!$, die nur für natürliche Zahlen definiert ist.

Kapitel 6

Funktionenfolgen

Über Folgen von reellen Zahlen hinaus sind Folgen von Funktionen von großer Bedeutung. Diese haben wir schon in den letzten Kapiteln betrachtet. Insbesondere kann man Potenzreihen als Grenzfunktionen ihrer Partialsummenfolgen interpretieren. Wir wollen nun allgemeine *Funktionenfolgen* und ihr Konvergenzverhalten untersuchen.

6.1 Punktweise und gleichmäßige Konvergenz

Definition 6.1.1. Sei M eine Menge, und für jedes $n \in \mathbb{N}$ sei eine Funktion $f_n : M \rightarrow \mathbb{C}$ gegeben.

- (a) Die Folge von Funktionen $f_n : M \rightarrow \mathbb{C}$ heißt *punktweise konvergent*, falls für jedes $x \in M$ die Zahlenfolge $(f_n(x))_{n \in \mathbb{N}}$ konvergiert. Die Grenzfunktion $f : M \rightarrow \mathbb{C}$ mit $f(x) := \lim_{n \rightarrow \infty} f_n(x)$ heißt *punktweiser Grenzwert* der Folge (f_n) .
- (b) Die Folge von Funktionen $f_n : M \rightarrow \mathbb{C}$ heißt *gleichmäßig konvergent* gegen $f : M \rightarrow \mathbb{C}$, falls zu jedem $\epsilon > 0$ ein $n_0 \in \mathbb{N}$ existiert mit $|f_n(x) - f(x)| < \epsilon$ für alle $n \geq n_0$ und für alle $x \in M$.

Bemerkung. Im Gegensatz zur punktweisen Konvergenz, muß es bei der gleichmäßigen Konvergenz zu vorgegebenem $\epsilon > 0$ ein $n_0 \in \mathbb{N}$ geben, unabhängig von x , mit $|f_n(x) - f(x)| < \epsilon$ für alle $n \geq n_0$. Anschaulich bedeutet dies, dass für jedes $\epsilon > 0$ die Graphen der Funktionen f_n ab einem n_0 in einem ϵ -Schlauch um die Grenzfunktion f liegen, siehe Abb. 6.1. Folgendes einfache Kriterium ist nützlich für die Charakterisierung der gleichmäßigen Konvergenz.

Satz 6.1.2. Sei $f_n : M \rightarrow \mathbb{C}$ eine Funktionenfolge und $f : M \rightarrow \mathbb{C}$ eine Funktion. Dann sind folgende Aussagen äquivalent

- (1) f_n konvergiert gleichmäßig gegen f .
- (2) $\|f - f_n\|$ ist eine Nullfolge, wobei

$$\|f - f_n\| = \sup\{|f(x) - f_n(x)| \mid x \in M\}$$

die Supremumsnorm von $f - f_n$ auf M bezeichnet.

Beispiele.

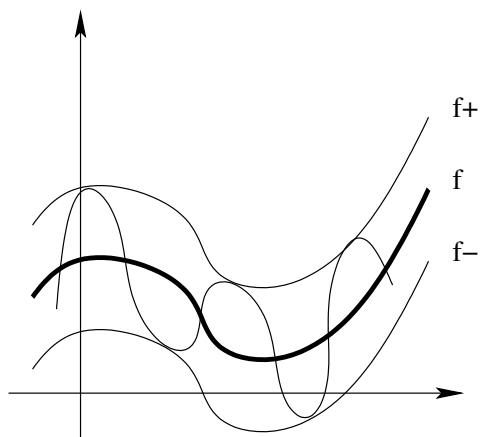


Abbildung 6.1:

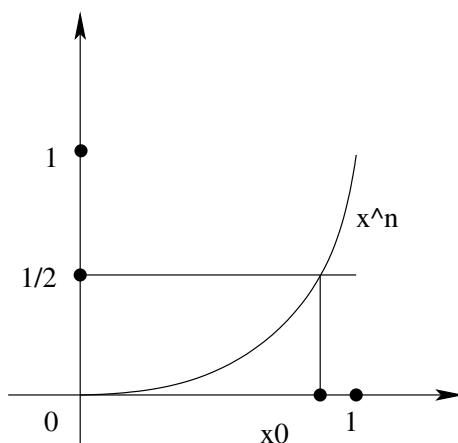


Abbildung 6.2:

- (a) Sei $f_n : [0, 1] \rightarrow \mathbb{R}$ eine Funktionenfolge mit $f_n(x) = x^n$. Für jedes $x \in [0, 1]$ ist die Folge $(f_n(x))$ konvergent und es gilt:

$$\lim_{n \rightarrow \infty} f_n(x) = \begin{cases} 0 & \text{für } x \in [0, 1), \\ 1 & \text{für } x = 1. \end{cases}$$

Diese Folge ist nicht gleichmäßig konvergent, d.h. es existiert ein $\epsilon > 0$, für das kein $n_0 \in \mathbb{N}$ existiert, ab dem die Funktionen f_n in einem ϵ -Schlauch um f liegen. Es existiert also ein $\epsilon > 0$, so dass für alle $n_0 \in \mathbb{N}$ ein $n \geq n_0$ und $x \in [0, 1]$ existiert mit $|f_n(x) - f(x)| \geq \epsilon$.

Zu $n_0 \in \mathbb{N}$ betrachte $x = \sqrt[n_0]{\frac{1}{2}} > 0$. Dann gilt für $n = n_0$: $f_{n_0}(x) = \frac{1}{2}$ und $f(x) = 0$. Insbesondere ist $|f_{n_0}(x) - f(x)| \geq \frac{1}{2}$.

- (b) Auf $[0, \frac{1}{2}]$ konvergiert hingegen die Funktionenfolge $f_n(x) = x^n$ gleichmäßig. Denn da $x \rightarrow x^n$ auf $[0, \frac{1}{2}]$ monoton ist, gilt: $|x^n - 0| \leq (\frac{1}{2})^n$. Da $(\frac{1}{2})^n$ eine Nullfolge ist, folgt die Behauptung.

(c) Sei $f_n = \mathbb{R} \rightarrow \mathbb{R}$ definiert durch $f_n(x) = \frac{n}{n^2+x^2}$.

Da $|f_n(x)| \leq \frac{n}{n^2} = \frac{1}{n}$ für alle $x \in \mathbb{R}$, konvergiert f_n gleichmäßig gegen $f = 0$.

(d) Sei

$$\sum_{k=0}^{\infty} a_k (z - z_0)^k$$

eine Potenzreihe mit Konvergenzradius $R > 0$ (siehe Abschnitt 2.7). Dann konvergiert für jedes $0 < r < R$ die Folge der Partialsummen

$$p_n(z) = \sum_{k=0}^n a_k (z - z_0)^k$$

gleichmäßig auf der abgeschlossenen Kreisscheibe $B(z_0, r) := \{z \mid |z - z_0| \leq r\}$ gegen

$$p(z) := \sum_{k=0}^{\infty} a_k (z - z_0)^k,$$

denn

$$|p(z) - p_n(z)| \leq \sum_{k=n+1}^{\infty} |a_k| |z - z_0|^k \leq \sum_{k=n+1}^{\infty} |a_k| r^k.$$

Da wegen Satz 2.7.2 Potenzreihen innerhalb des Konvergenzkreises absolut konvergieren, konvergiert auch die Reihe $\sum_{k=0}^{\infty} |a_k| r^k$. Insbesondere ist $b_n := \sum_{k=n+1}^{\infty} |a_k| r^k$ eine Nullfolge.

Wir haben gesehen, dass Cauchyfolgen komplexer Zahlen konvergieren. Dieses Cauchy Kriterium liefert ein Kriterium für die gleichmäßige Konvergenz von Funktionenfolgen. Wie bei den komplexen Zahlen, besteht der Vorteil dieses Kriteriums darin, dass man die Grenzfunktion nicht kennen muss.

Satz 6.1.3. *Sei $f_n : M \rightarrow \mathbb{C}$ eine Funktionenfolge, so sind folgende Aussagen äquivalent:*

- (1) *Es existiert eine Funktion $f : M \rightarrow \mathbb{C}$, so dass die Funktionenfolge $f_n : M \rightarrow \mathbb{C}$ gleichmäßig gegen $f : M \rightarrow \mathbb{C}$ konvergiert.*
- (2) *Die Funktionenfolge $f_n : M \rightarrow \mathbb{C}$ ist eine Cauchyfolge bezüglich der Supremumsnorm, d.h. zu jedem $\epsilon > 0$ existiert ein n_0 mit*

$$\|f_n - f_m\| < \epsilon$$

für alle $n, m \geq n_0$.

Beweis. Sei zunächst (1) erfüllt, so existiert ein $n_0 \in \mathbb{N}$ mit

$$\|f_n - f\| < \frac{\epsilon}{2}$$

für alle $n \geq n_0$. Also gilt für $n, m > n_0$:

$$\|f_n - f_m\| \leq \|f_n - f\| + \|f_m - f\| < \epsilon.$$

Sei nun (2) erfüllt. Dann ist für jedes $x \in M$ die Folge $f_n(x)$ eine Cauchyfolge komplexer Zahlen, denn für alle $n, m \in \mathbb{N}$ ist $|f_n(x) - f_m(x)| \leq \|f_n - f_m\|$. Insbesondere existiert für jedes $x \in M$ der Grenzwert dieser Zahlenfolge. Definiere $f : M \rightarrow \mathbb{C}$ durch

$$f(x) := \lim_{n \rightarrow \infty} f_n(x).$$

Nun müssen wir noch zeigen, dass (f_n) gleichmäßig gegen f konvergiert. Ist $\epsilon > 0$, so wähle $n_0 \in \mathbb{N}$ so, dass

$$\|f_n - f_m\| < \epsilon.$$

Dann gilt für alle $n, m \geq n_0$ und $x \in M$:

$$|f_n(x) - f(x)| \leq |f_n(x) - f_m(x)| + |f_m(x) - f(x)| < \epsilon + |f_m(x) - f(x)|.$$

Da für jedes feste $x \in M$ gilt: $\lim_{m \rightarrow \infty} |f_m(x) - f(x)| = 0$, folgt auch:

$$|f_n(x) - f(x)| < \epsilon,$$

für alle $x \in M$. □

Zum Schluss dieses Abschnittes wollen wir noch ohne Beweis den Weierstraßschen Approximationssatz festhalten. Er besagt, dass stetige Funktionen auf Intervallen sich durch Polynome gleichmäßig approximieren lassen. Genauer gilt:

Satz 6.1.4. (Weierstraßscher Approximationssatz, siehe Königsberger, Band 1, S.313)

Sei $f : [a, b] \rightarrow \mathbb{R}$ eine stetige Funktion. Dann existiert zu jedem $\epsilon > 0$ ein Polynom $p : [a, b] \rightarrow \mathbb{R}$ mit

$$\|f - p\| < \epsilon.$$

6.2 Gleichmäßige Konvergenz und Vertauschung von Grenzprozessen

Wir werden nun sehen, dass sich unter gleichmäßiger Konvergenz gewisse Eigenschaften von Funktionenfolgen auf ihre Grenzwerte übertragen lassen.

Satz 6.2.1. Sei $D \subset \mathbb{C}$ und $f_n : D \rightarrow \mathbb{C}$ eine gleichmäßig konvergente Folge stetiger Funktionen mit Grenzwert f . Dann ist $f : D \rightarrow \mathbb{C}$ stetig.

Beweis. Sei $a \in D$. Für jedes $x \in M$ gilt:

$$\begin{aligned} |f(x) - f(a)| &\leq |f(x) - f_n(x)| + |f_n(x) - f(a)| \\ &\leq |f(x) - f_n(x)| + |f_n(x) - f_n(a)| + |f_n(a) - f(a)|. \end{aligned}$$

Sei $\epsilon > 0$. Dann existiert wegen der gleichmäßigen Konvergenz der Folge f_n ein $n \in \mathbb{N}$, so dass

$$|f(y) - f_n(y)| < \epsilon/3$$

für alle $y \in D$. Wegen der Stetigkeit von f_n existiert ein $\delta > 0$, so dass

$$|f_n(x) - f_n(a)| < \epsilon/3$$

für alle $x \in B(a, \delta)$. Dann folgt:

$$|f(x) - f(a)| \leq |f(x) - f_n(x)| + |f_n(x) - f_n(a)| + |f_n(a) - f(a)| < \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon$$

für alle $x \in B(a, \delta)$. Somit ist f stetig in a . □

Bemerkungen.

- (a) Aus der Stetigkeit der Funktionen f_n und der Grenzfunktion f folgt, dass sich die Grenzprozesse $\lim_{x \rightarrow a} \lim_{n \rightarrow \infty}$ vertauschen lassen, denn

$$\lim_{x \rightarrow a} \lim_{n \rightarrow \infty} f_n(x) = \lim_{x \rightarrow a} f(x) = f(a) = \lim_{n \rightarrow \infty} f_n(a) = \lim_{n \rightarrow \infty} \lim_{x \rightarrow a} f_n(x).$$

- (b) Der Grenzwert einer punktweise konvergenten Folge stetiger Funktionen ist im allgemeinen nicht stetig, wie das Beispiel der Funktionenfolge $f_n : [0, 1] \rightarrow \mathbb{R}$ mit $f_n(x) = x^n$ zeigt.
- (c) Auf der anderen Seite existieren Beispiele von nur punktweise konvergenten Folge stetiger Funktionen, die gegen eine stetige Funktion konvergiert. Betrachte dazu die auf $M = [0, 2]$ durch

$$f_n(x) = \begin{cases} nx & \text{für } 0 \leq x \leq \frac{1}{n}, \\ 2 - nx & \text{für } \frac{1}{n} < x \leq \frac{2}{n}, \\ 0 & \text{für } \frac{2}{n} < x < 2 \end{cases}$$

definierte Funktionenfolge. Dann gilt: $\lim_{n \rightarrow \infty} f_n(x) = 0$ für alle $x \in M$. Denn ist $x > 0$, so existiert ein $n_0 \in \mathbb{N}$ mit $\frac{2}{n_0} < x$. Somit ist $f_n(x) = 0$ für alle $n \geq n_0$. Für $x = 0$ ist $f_n(x) = 0$ für alle $n \in \mathbb{N}$. Die Funktionenfolge (f_n) konvergiert *nicht* gleichmäßig gegen 0, denn für jedes $n \in \mathbb{N}$ gibt es $x = \frac{1}{n}$ mit $|f_n(\frac{1}{n}) - 0| = 1$.

Satz 6.2.2. Sei $f_n \in R([a, b])$ eine Folge von Regelfunktionen, die gleichmäßig gegen $f : [a, b] \rightarrow \mathbb{R}$ konvergiert. Dann ist $f \in R([a, b])$ und es gilt

$$\int f = \lim_{n \rightarrow \infty} \int f_n.$$

Beweis. Wir zeigen zunächst, dass f eine Regelfunktion ist, d.h. zu jedem $\epsilon > 0$ existiert eine Treppenfunktion φ mit $\|\varphi - f\| < \epsilon$. Dazu wähle $n \in \mathbb{N}$ so groß, dass $\|f - f_n\| < \frac{\epsilon}{2}$. Da f_n eine Regelfunktion ist, existiert eine Treppenfunktion φ mit $\|f_n - \varphi\| < \frac{\epsilon}{2}$. Also folgt:

$$\|f - \varphi\| \leq \|f - f_n\| + \|f_n - \varphi\| < \epsilon.$$

Außerdem gilt für alle Regelfunktionen $g \in R([a, b])$:

$$\left| \int g \right| \leq (b - a) \|g\|.$$

Betrachte dazu eine Folge von Treppenfunktionen $\varphi_n \in T([a, b])$ mit $\lim_{n \rightarrow \infty} \|\varphi_n - g\| = 0$. Dann folgt:

$$\left| \int g \right| \stackrel{Def}{=} \lim_{n \rightarrow \infty} \left| \int \varphi_n \right| \leq \lim_{n \rightarrow \infty} \|\varphi_n\| (b - a) = \|g\| (b - a).$$

Dabei folgt das letzte Gleichheitszeichen aus der Abschätzung

$$\left| \|\varphi_n\| - \|g\| \right| \leq \|\varphi_n - g\|.$$

Damit erhalten wir:

$$\left| \int f - \int f_n \right| = \left| \int (f - f_n) \right| \leq \|f - f_n\| (b - a) \rightarrow 0$$

für $n \rightarrow \infty$. □

□

Wir haben gesehen, dass eine gleichmäßig konvergente Funktionenfolge von stetigen Funktionen eine stetige Funktion als Grenzwert besitzt. In wie weit lassen sich entsprechende Aussagen für differenzierbare Funktionen beweisen?

Genauer: Sei f_n eine auf $[a, b]$ gleichmäßig konvergente Folge von differenzierbaren Funktionen. Gilt dann:

- (1) $\lim_{n \rightarrow \infty} f_n(x) = f(x)$ ist differenzierbar?
- (2) Ist f differenzierbar, ist dann $\lim_{n \rightarrow \infty} f'_n(x) = f'(x)$?

Beide Fragen sind mit nein zu beantworten:

- (1) z.B. lässt sich $f(x) = |x|$ auf $I = [-1, 1]$ gleichmäßig durch differenzierbare Funktionen, ja sogar durch Polynome (Weierstraßscher Approximationssatz 6.1.4) approximieren.
- (2) Sei $I = [0, \pi]$, $f_n(x) = \frac{\sin nx}{n}$. Dann gilt: $\|f_n\| = \frac{1}{n}$, d.h. f_n konvergiert gleichmäßig gegen 0. Die Funktionenfolge $f'_n(x) = \cos nx$ konvergiert nicht einmal punktweise (betrachte zum Beispiel $x = \frac{\pi}{2}$).

Es gilt der folgende Satz:

Satz 6.2.3. Sei $f_n \in C^1([a, b])$ eine Folge von stetig differenzierbaren Funktionen mit den Eigenschaften:

- (1) f'_n konvergiert gleichmäßig.
- (2) Es existiert ein $x_0 \in [a, b]$, so dass die Zahlenfolge $f_n(x_0)$ konvergiert.

Dann konvergiert die Funktionenfolge f_n gleichmäßig gegen eine differenzierbare Funktion f und es gilt

$$f'(x) = \left(\lim_{n \rightarrow \infty} f_n \right)'(x) = \lim_{n \rightarrow \infty} (f'_n(x)).$$

Beweis. Sei $g(s) := \lim_{n \rightarrow \infty} f'_n(s)$ für $s \in [a, b]$. Dann ist g eine stetige Funktion. Aus dem Hauptsatz der Differential- und Integralrechnung angewandt auf f_n folgt:

$$f_n(x) = \int_{x_0}^x f'_n(s) ds + f_n(x_0).$$

Da f'_n gleichmäßig gegen g konvergiert, folgt aus Satz 6.2.2:

$$\lim_{n \rightarrow \infty} f_n(x) = \int_{x_0}^x g(s) ds + \lim_{n \rightarrow \infty} f_n(x_0) := f(x).$$

Diese Konvergenz ist gleichmäßig, denn

$$\begin{aligned} |f_n(x) - f(x)| &= \left| \int_{x_0}^x f'_n(s) ds + f_n(x_0) - \int_{x_0}^x g(s) ds - \lim_{n \rightarrow \infty} f_n(x_0) \right| \\ &= \left| \int_{x_0}^x (f'_n(s) - g(s)) ds + f_n(x_0) - \lim_{n \rightarrow \infty} f_n(x_0) \right| \\ &\leq |x - x_0| \|f'_n - g\| + |f_n(x_0) - \lim_{n \rightarrow \infty} f_n(x_0)|. \end{aligned}$$

Da g stetig ist, ist f wegen des Hauptsatzes differenzierbar und es gilt:

$$f'(x) = g(x) = \lim_{n \rightarrow \infty} f'_n(x).$$

□

□

Bemerkung.

Dies liefert einen anderen Beweis dafür, dass man Potenzreihen gliedweise differenzieren darf. Die Ableitung der Partialsummenfolge einer Potenzreihe ist nämlich ebenfalls gleichmäßig konvergent.

Kapitel 7

Vektorräume

7.1 Grundlegende Definitionen

Eine wichtiger grundlegender Begriff in Physik und Mathematik ist der Begriff der Gruppe.

Definition 7.1.1. Eine *Gruppe* ist ein Paar (G, \cdot) , bestehend aus einer Menge G und einer Verknüpfung (oder Komposition) \cdot auf G , d.h. einer Abbildung

$$\cdot : G \times G \rightarrow G$$

mit folgenden Eigenschaften (*Gruppenaxiomen*):

1. Für alle $x, y, z \in G$ gilt $(x \cdot y) \cdot z = x \cdot (y \cdot z)$ (*Assoziativgesetz*).
2. Es gibt ein Element $e \in G$ (*neutrales Element* von G) mit $e \cdot x = x$ für alle $x \in G$.
3. Zu jedem $x \in G$ gibt es ein $x' \in G$ (*inverses Element* zu x) mit $x' \cdot x = e$.

Eine Gruppe heißt *abelsch*, falls $x \cdot y = y \cdot x$ für alle $x, y \in G$ (*Kommutativgesetz*).

Der Einfachheit halber werden wir oft xy statt $x \cdot y$ schreiben.

Satz 7.1.2. *Es sei (G, \cdot) eine Gruppe, so gilt:*

1. *Ist $e \in G$ ein neutrales Element, so gilt $xe = x$ für alle $x \in X$.*
2. *Das neutrale Element ist eindeutig bestimmt.*
3. *Ist x' invers zu x , so gilt auch $xx' = e$.*
4. *Zu jedem $x \in G$ gibt es genau ein inverses Element x' , das man mit x^{-1} bezeichnet.*

Beweis.

zu 3. $xx' = e(xx') = (x''x')(xx') = x''(x'(xx')) = x''((x'x)x') = x''(ex') = x''x' = e$, wobei $x'' = (x')'$ bezeichnet.

zu 1. $xe = x(x'x) = (xx')x = ex = x$.

zu 2. Sei \tilde{e} weiteres neutrales Element. Dann folgern wir aus Teil 1: $e\tilde{e} = e$. Da e neutrales Element ist, erhalten wir nach Definition: $e\tilde{e} = \tilde{e}$. Also ist $e = \tilde{e}$.

zu 4. Es seien x', \tilde{x} inverse Elemente zu x . Dann folgt:

$$\tilde{x} = \tilde{x}e \stackrel{3.}{=} \tilde{x}(xx') = (\tilde{x}x)x' = ex' = x'.$$

□

Beispiele für Gruppen:

1. Die Menge \mathbb{R} der reellen Zahlen ist eine abelsche Gruppe mit der Verknüpfung

$$(x, y) \rightarrow x + y.$$

Das neutrale Element ist 0, das inverse Element ist $x^{-1} = -x$.

2. $\mathbb{Q}^* = \mathbb{Q} \setminus \{0\}$ und $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$ mit der Verknüpfung

$$(x, y) \rightarrow x \cdot y$$

sind abelsche Gruppen mit neutralem Element 1 und inverse Element $x^{-1} = \frac{1}{x}$.

3. Sei M eine Menge und

$$S(M) = \{f \mid f : M \rightarrow M \text{ bijektive Abbildung}\}.$$

Dann ist $S(M)$ mit der Verknüpfung

$$(f, g) \rightarrow f \circ g$$

von Abbildungen eine Gruppe. Das neutrale Element ist die Identität id_M auf M , definiert durch $\text{id}_M(x) = x$ für alle $x \in M$. Das inverse Element zu f ist ihre inverse Abbildung f^{-1} . Die Gruppe $(S(M), \circ)$ heißt die *Gruppe der Permutationen* von M oder auch die *symmetrische Gruppe* von M . Ist $M = \{1, \dots, n\}$, so schreibt man $S_n = S(M)$. Die Gruppe S_n ist nicht abelsch, falls $n \geq 3$.

Definition 7.1.3. Sei (G, \cdot) eine Gruppe. Eine nichtleere Teilmenge $H \subset G$ heißt *Untergruppe* von G , falls Folgendes gilt:

1. H ist abgeschlossen bezüglich der Gruppenoperation auf G , d.h. für alle $h_1, h_2 \in H$ folgt $h_1 h_2 \in H$.
2. Für jedes $h \in H$ ist auch $h^{-1} \in H$.

Bemerkung. Jede Untergruppe H von G ist mit der Gruppenoperation auf G wieder eine Gruppe. Wegen 1.+2. mit $h \in H$ auch $e = hh^{-1} \in H$.

Beispiele.

1. Ist G eine Gruppe, so sind $\{e\}$ und G trivialerweise Untergruppen. Sie heißen *triviale Untergruppen*.
2. Die Menge \mathbb{Z} der ganzen Zahlen ist Untergruppe der reellen Zahlen.

3. Die Menge

$$D(\mathbb{C}) = \{f : \mathbb{C} \rightarrow \mathbb{C} \mid f(z) = e^{i\varphi}z, \varphi \in \mathbb{R}\}$$

ist eine Untergruppe von $(S(\mathbb{C}), \circ)$. Jede Abbildung $f \in D(\mathbb{C})$ ist nämlich bijektiv und für alle $f_1, f_2 \in D(\mathbb{C})$ mit $f_1(z) = e^{i\varphi_1}z$ und $f_2(z) = e^{i\varphi_2}z$ ist auch $f_1 \circ f_2 \in D(\mathbb{C})$, denn

$$f_1 \circ f_2(z) = e^{i\varphi_1}e^{i\varphi_2}z = e^{i(\varphi_1+\varphi_2)}z.$$

Außerdem ist für jedes $f \in S(\mathbb{C})$ mit $f(z) = e^{i\varphi}z$ auch $f^{-1} \in S(\mathbb{C})$, denn $f^{-1}(z) = e^{-i\varphi}z$.

Nun kommen wir zu dem wichtigen Begriff des Vektorraumes. In der Schule wurde die Vektorrechnung hauptsächlich zur Beschreibung von Geraden oder Ebenen im \mathbb{R}^3 benutzt. Jedoch findet der Vektorraumbegriff in Physik und Mathematik in einem weit größeren Bereich seine Anwendung. Wichtige Beispiele sind Lösungsmengen von linearen Differentialgleichungen. In der Quantenmechanik spielen gewisse “unendlichdimensionale” Vektorräume (Hilberträume) eine zentrale Rolle. Wir wollen nun den Begriff des Vektorraumes auf ein klares Fundament stellen und uns seiner mathematischen Beschreibung zuwenden.

Ein Vektorraum V trägt die Struktur einer abelschen Gruppe. Seine Elemente nennen wir Vektoren. Des Weiteren ist eine Multiplikation dieser Vektoren mit Elementen aus einem Körper definiert. Die Elemente aus dem Körper heißen *Skalare*. Genauer gilt:

Definition 7.1.4. Eine abelsche Gruppe $(V, +)$ heißt *Vektorraum über dem Körper K* , falls eine Abbildung (*skalare Multiplikation*)

$$\begin{aligned} K \times V &\rightarrow V \\ (\alpha, x) &\rightarrow \alpha \cdot x \end{aligned}$$

erklärt ist. Für die skalare Multiplikation sollen folgende Regeln gelten:

1. $(\alpha + \beta) \cdot x = \alpha \cdot x + \beta \cdot x$,
2. $\alpha \cdot (x + y) = \alpha \cdot x + \alpha \cdot y$,
3. $\alpha \cdot (\beta \cdot x) = (\alpha\beta) \cdot x$,
4. $1 \cdot x = x$,

für alle $\alpha, \beta \in K$, $x, y \in V$.

Bemerkung. In diesem Skriptum wird K meistens der Körper der reellen oder komplexen Zahlen sein. In gewissen Anwendungen, wie zum Beispiel in der Kryptographie, spielen auch endliche Körper eine wichtige Rolle.

Beispiele.

1. Sei K ein Körper. Sei $n \in \mathbb{N}$, so ist

$$K^n = \underbrace{K \times \dots \times K}_{n\text{-mal}} = \{(x_1, \dots, x_n) \mid x_i \in K, 1 \leq i \leq n\}$$

ein Vektorraum über K mit

$$x + y = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n)$$

und

$$\alpha x = (\alpha x_1, \alpha x_2, \dots, \alpha x_n)$$

für $\alpha \in K$ und $x, y \in K^n$. Insbesondere ist $K^1 = K$ auch ein Vektorraum über sich selbst.

2. Sei M eine Menge und V ein Vektorraum über K . Definiere

$$V^M := \{f \mid f : M \rightarrow V \text{ Abbildung}\}.$$

Definiere eine Addition auf V^M durch

$$(f + g)(x) := f(x) + g(x)$$

und eine skalare Multiplikation durch

$$(\alpha f)(x) := \alpha f(x)$$

für $\alpha \in K$ und $f, g \in V^M$. Dann ist V^M ein Vektorraum über K , denn

- a) V^M ist eine abelsche Gruppe. Diese Eigenschaft folgt unmittelbar aus der Definition der Addition. Die Gruppeneigenschaften von V übertragen sich auf V^M . Der Nullvektor ist die Nullabbildung.
- b) Die skalare Multiplikation erfüllt die oben genannten Axiome.

Bemerkung. Das Beispiel 1 lässt sich als Spezialfall von Beispiel 2 auffassen. Denn ist $M = \{1, \dots, n\}$, so bestimmt jede Abbildung $f : M \rightarrow K$ ein n -Tupel

$$(f(1), \dots, f(n)) \in K^n.$$

Auf der anderen Seite bestimmt jedes n -Tupel $(x_1 \dots x_n) \in K^n$ eine Abbildung $f \in V^M$ mit

$$f(1) = x_1, \dots, f(n) = x_n.$$

Die Addition der Abbildungen entspricht der Addition von n -Tupeln. Die skalare Multiplikation einer Abbildung entspricht der skalaren Multiplikation eines n -Tupels.

Aus den Vektorraumaxiomen lassen sich folgende Rechenregeln ableiten.

Satz 7.1.5. Sei V ein Vektorraum über K und mit 0_K bzw. 0_V seien die 0 in K bzw. V bezeichnet. Dann gilt:

1. $0_K \cdot x = 0_V$, für alle $x \in V$.
2. $\alpha \cdot 0_V = 0_V$, für alle $\alpha \in K$.
3. $(-1) \cdot x = -x$, für alle $x \in V$.
4. Ist $\alpha \cdot x = 0$ für $\alpha \in K$ und $x \in V$, so folgt: $\alpha = 0_K$ oder $x = 0_V$.

Bemerkung. Wenn es aus dem Zusammenhang klar ist, werden wir die 0 nicht zusätzlich kennzeichnen. Insbesondere werden wir im folgenden Beweis darauf verzichten.

Beweis. Es sei $x \in V, \alpha \in K$. Dann folgt:

1. $0 \cdot x = (0 + 0) \cdot x = 0 \cdot x + 0 \cdot x$ und damit ist $0 \cdot x = 0$.
2. $\alpha \cdot 0 = \alpha \cdot (0 + 0) = \alpha \cdot 0 + \alpha \cdot 0$.
3. $x + (-1) \cdot x = 1 \cdot x + (-1) \cdot x = (1 + (-1)) \cdot x = 0 \cdot x = 0$, d.h. $(-1) \cdot x = -x$.
4. Es sei $\alpha \cdot x = 0$ und $\alpha \neq 0$, so folgt: $x = 1 \cdot x = (\alpha^{-1} \cdot \alpha) \cdot x = \alpha^{-1} \cdot (\alpha \cdot x) = 0$.

□

Von großer Bedeutung ist der Begriff des Untervektorraumes (*Unterraum, Teilraum*).

Definition 7.1.6. Sei V ein Vektorraum über dem Körper K . Eine nichtleere Teilmenge $U \subset V$ heißt *Unterraum*, falls U abgeschlossen bezüglich der Addition und skalaren Multiplikation in V ist, d.h.

- a) für alle $x, y \in U$ gilt: $x + y \in U$.
- b) für alle $\alpha \in K$ und für alle $x \in U$ gilt: $\alpha x \in U$.

Bemerkungen.

1. Aus der Definition folgt, dass U mit der Struktur von V ein Vektorraum ist. Denn für jedes $y \in U$, ist auch das inverse Element $-y = (-1)y \in U$. Wegen a) ist somit auch der Nullvektor in U . Die übrigen Vektorraumaxiome gelten insbesondere in U , da sie in V gelten.
2. Die beiden definierenden Eigenschaften eines Unterraumes entsprechen genau dem Superpositionsprinzip in der Physik. Dies wird besonders in Beispiel 7) deutlich.

Beispiele.

1. Ist V ein Vektorraum, so sind $\{0\}$ und V trivialerweise Unterräume. Sie heißen *triviale Unterräume*.
2. Unterräume werden oft mittels gewisser linearer Gleichungssysteme definiert. Ist zum Beispiel $V = \mathbb{R}^2$, so ist

$$U = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 + x_2 = 0\}$$

ein Unterraum von \mathbb{R}^2 . Dagegen sind die Mengen

$$U = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_2 = x_1^2\} \text{ und } U' = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_2 \geq 0\}$$

keine Unterräume von \mathbb{R}^2 .

Viele wichtige Beispiele von Unterräumen stammen aus der Analysis.

3. Die Menge der Abbildungen

$$\mathbb{R}^{\mathbb{N}} = \{f \mid f : \mathbb{N} \rightarrow \mathbb{R}\}$$

ist nach obigem Beispiel ein Vektorraum über \mathbb{R} . Er heißt *Vektorraum der reellen Folgen*, denn jede Abbildung $f \in \mathbb{R}^{\mathbb{N}}$ definiert eine Folge reeller Zahlen durch $f(n) = x_n$. Umgekehrt definiert jede Folge eine Abbildung $f \in \mathbb{R}^{\mathbb{N}}$. Dann ist die Menge der Nullfolgen

$$\{f \in \mathbb{R}^{\mathbb{N}} \mid f(n) \rightarrow 0 \text{ für } n \rightarrow \infty\}$$

ein Unterraum von $\mathbb{R}^{\mathbb{N}}$, denn Summen und skalare Vielfache von Nullfolgen sind Nullfolgen.

4. Die Menge der stetigen Abbildungen

$$C^0(\mathbb{R}) = \{f : \mathbb{R} \rightarrow \mathbb{R} \mid f \text{ stetige Abbildung}\}$$

ist ein Unterraum aller Abbildungen $\mathbb{R}^{\mathbb{R}} := \{f \mid f : \mathbb{R} \rightarrow \mathbb{R}\}$. Ist $[a, b] \subset \mathbb{R}$ ein kompaktes Intervall, so ist die Menge $T([a, b])$ der Treppenfunktionen ein Unterraum des Vektorraumes aller Abbildungen $\mathbb{R}^{[a, b]} := \{f \mid f : [a, b] \rightarrow \mathbb{R}\}$. Genauso ist die Menge $R([a, b])$ der Regelfunktionen ein Unterraum des Vektorraumes aller Abbildungen $\mathbb{R}^{[a, b]} := \{f \mid f : [a, b] \rightarrow \mathbb{R}\}$. Sie umfassen alle Treppenfunktionen.

5. Die Menge der stetig differenzierbaren Abbildungen

$$C^1(\mathbb{R}) = \{f \in C^0(\mathbb{R}) \mid f \text{ ist differenzierbar und } f' \text{ ist stetig}\}$$

ist Unterraum von $C^0(\mathbb{R})$. Im allgemeinen definiert man für $k \in \mathbb{N}$:

$$C^k(\mathbb{R}) = \{f \in C^{k-1} \mid f \text{ ist } k \text{ - mal differenzierbar, die } k \text{ - te Ableitung } f^{(k)} \text{ ist stetig}\}.$$

Dann ist $C^k(\mathbb{R})$ ist Unterraum von $C^{k-1}(\mathbb{R})$. Die Menge der unendlich oft differenzierbaren Abbildungen

$$C^\infty(\mathbb{R}) = \{f \in C^0(\mathbb{R}) \mid f \text{ unendlich oft differenzierbar}\}$$

ist Unterraum von $C^k(\mathbb{R})$ für alle $k \in \mathbb{N}$.

6. Sei K ein Körper.

$$P_n(K) = \{p \in K^K \mid p(x) = \sum_{i=0}^n a_i x^i, a_i \in K\}$$

heißen *Polynomabbildungen* (oder *Polynome*) vom Grad $\leq n$. $P_n(K)$ ist Unterraum von K^K . $P_n(\mathbb{R})$ ist sogar Unterraum von $C^\infty(\mathbb{R})$.

7. Die Lösungsmenge von *homogenen linearen Differentialgleichungen* bilden einen Untervektorraum. Zum Beispiel ist

$$\{f \in C^2(\mathbb{R}) \mid f'' + 3f' + 2f = 0\}$$

ein Unterraum von $C^2(\mathbb{R})$, denn die Summe zweier Lösungen und das skalare Vielfache einer Lösung ist wieder eine Lösung (*Superpositionsprinzip*). Allgemeiner Typ einer homogenen linearen Differentialgleichung n-ter Ordnung (mit konstanten Koeffizienten) ist:

$$a_n f^{(n)} + a_{n-1} f^{(n-1)} + \dots + a_0 f = 0, a_i \in \mathbb{R}, f \in C^n(\mathbb{R}).$$

Durchschnitte von Unterräumen sind Unterräume. Genauer gilt:

Satz 7.1.7. Sei V ein Vektorraum über K und P eine beliebige Menge von Unterräumen von V . Dann ist ihr Durchschnitt

$$\bigcap_{U \in P} U = \{x \in V \mid x \in U \text{ für alle } U \in P\}$$

wieder ein Unterraum von V .

Bemerkung. Ist $P = \{U_i \subset V \mid U_i \text{ Unterraum, } i \in I\}$, wobei I eine Menge (Indexmenge) ist, so schreibt man auch

$$\bigcap_{i \in I} U_i \text{ statt } \bigcap_{U \in P} U.$$

Beweis. Es ist $\bigcap_{U \in P} U \neq \emptyset$, denn $0 \in U$ für alle $U \in P$.

Sei $x, y \in \bigcap_{U \in P} U$, d.h. $x, y \in U$ für alle $U \in P$. Da jedes $U \in P$ ein Unterraum von V ist, folgt auch $x + y \in U$ und $\alpha x \in U$ für alle $\alpha \in K$ und alle $U \in P$. Dies zeigt, dass $\bigcap_{U \in P} U$ abgeschlossen bezüglich der Addition und skalaren Multiplikation ist. \square

Ein wichtiges Beispiel dazu ist der Begriff des *Erzeugendensystems*:

Definition 7.1.8. Sei $E \subset V$ eine beliebige Teilmenge eines Vektorraumes V . Dann heißt

$$\text{Span}(E) := \bigcap_{\substack{E \subset U \\ U \text{ Unterraum von } V}} U$$

der von E aufgespannte (erzeugte) Unterraum. E heißt *Erzeugendensystem* von V , falls $\text{Span}(E) = V$.

Bemerkungen.

- $\text{Span}(E)$ ist der kleinste Unterraum, der die Menge E enthält. Denn für jeden Unterraum U von V mit $E \subset U$, folgt $\text{Span}(E) \subset U$, denn $\text{Span}(E)$ liegt im Durchschnitt aller Unterräume, die E umfassen.
- $\text{Span}(\emptyset) = 0$.

Nun wollen wir zeigen, dass $\text{Span}(E)$ mit dem Vektorraum, bestehend aus allen Linearkombinationen von E , übereinstimmt. Eine Linearkombination von E ist eine endliche Summe der Form

$$\sum_{e \in E} \alpha_e e.$$

Dabei sind die Koeffizienten $\alpha_e \in K$ mit $\alpha_e \neq 0$ für höchstens endlich viele $e \in E$. Ist

$$\{e_1, \dots, e_k\} = \{e \in E \mid \alpha_e \neq 0\},$$

so gilt

$$\sum_{e \in E} \alpha_e e = \alpha_1 e_1 + \dots + \alpha_k e_k,$$

falls wir $\alpha_i := \alpha_{e_i}$ setzen. Allerdings erweist sich die erste Schreibweise im Folgenden als zweckmäßiger.

Satz 7.1.9. Sei V ein Vektorraum über K und $E \subset V$. Definiere für $E \neq \emptyset$

$$[E] := \left\{ x \in V \mid x = \sum_{e \in E} \alpha_e e, \alpha_e \in K, \text{ mit } \alpha_e \neq 0 \right. \\ \left. \text{für höchstens endlich viele } e \in E \right\}$$

und definiere $[E] = \{0\}$, falls $E = \emptyset$. Dann ist $[E]$ ein Unterraum von V und es gilt:

$$\text{Span}(E) = [E].$$

Beweis. Die Menge $[E]$ ist abgeschlossen bezüglich der Addition. Betrachte dazu $x = \sum_{e \in E} \alpha_e e$ und $y = \sum_{e \in E} \beta_e e$ mit $\alpha_e, \beta_e \in K$ und $\alpha_e, \beta_e \neq 0$ für endlich viele $e \in E$. Dann ist

$$x + y = \sum_{e \in E} (\alpha_e + \beta_e) e \in [E],$$

denn $c_e = \alpha_e + \beta_e \neq 0$ für höchstens endlich viele $e \in E$.

Genauso folgt, dass $[E]$ abgeschlossen bezüglich der skalaren Multiplikation ist, denn ist $x = \sum_{e \in E} \alpha_e e$ eine Linearkombination von Vektoren aus E , so gilt dies auch für

$$\alpha x = \sum_{e \in E} \alpha \alpha_e e,$$

für alle $\alpha \in K$. Damit ist $[E]$ ein Unterraum von V , der die Menge E enthält.

Sei nun U ein beliebiger Unterraum von V , der E enthält, so enthält er auch alle Linearkombinationen von E . Also ist $[E] \subset U$. Damit ist $[E]$ der kleinste Vektorraum, der die Menge E enthält und somit ist $[E] = \text{Span}(E)$. \square

Beispiel. $V = \mathbb{R}^2$, $E = \{(1, 0), (0, 1)\}$. Dann ist

$$\text{Span}(E) = \{x \in \mathbb{R}^2 \mid x = \alpha_1(1, 0) + \alpha_2(0, 1), \alpha_1, \alpha_2 \in \mathbb{R}\} = \mathbb{R}^2.$$

Definition 7.1.10. Sei $k \in \mathbb{N}$ und seien U_1, \dots, U_k Unterräume von V , so heißt

$$\sum_{i=1}^k U_i = U_1 + \dots + U_k := \text{Span}(U_1 \cup U_2 \dots \cup U_k)$$

die *Summe der Unterräume* U_1, \dots, U_k von V . Allgemeiner definiert man: Ist P eine Menge von Unterräumen von V , so heißt

$$\sum_{U \in P} U := \text{Span} \left(\bigcup_{U \in P} U \right)$$

Summe der Unterräume P , wobei $\bigcup_{U \in P} U := \{x \in V \mid \text{es gibt ein } U \in P \text{ mit } x \in U\}$.

Übung:

a) $U_1 + \dots + U_k = \{x_1 + \dots + x_k \mid x_i \in U_i, 1 \leq i \leq k\}$.

Hinweis: Man beweise: die rechte Seite ist ein Unterraum von V . Außerdem ist er der kleinste Unterraum, der die Unterräume U_1, \dots, U_k enthält.

- b) Zeigen Sie: Sind U_1, U_2 Unterräume von V , so ist $U_1 \cup U_2$ im Allgemeinen kein Unterraum.

Auf den folgenden Begriff werden wir genauer in Abschnitt 7.3 eingehen.

Definition 7.1.11. Es sei U Unterraum von V . Ein Unterraum W von V heißt *Komplement* von U , falls $W \cap U = \{0\}$ und $U + W = V$. In diesem Falle schreibt man auch

$$V = U \oplus W$$

und nennt V die *direkte Summe* von U und W .

7.2 Basis und Dimension

Definition 7.2.1. Sei V ein Vektorraum über K . Eine Teilmenge E heißt *linear unabhängig* genau dann, falls der Nullvektor in V sich nur trivial aus Elementen von E kombinieren lässt, d.h. sind $\alpha_e \in K$, mit $\alpha_e \neq 0$ für höchstens endlich viele $e \in E$, so folgt aus

$$\sum_{e \in E} \alpha_e e = 0$$

$\alpha_e = 0$ für alle $e \in E$. Falls E nicht linear unabhängig ist, so heißt E *linear abhängig*.

Bemerkungen.

- a) Ist $E = \{0\} \subset V$, so ist E linear abhängig.
- b) Ist $E = \emptyset$, so ist E linear unabhängig, denn in diesem Falle existiert keine Linearkombination mit Elementen aus E , die den Nullvektor darstellt.
- c) Eine endliche Menge $E = \{e_1, \dots, e_m\}$ ist also genau dann linear unabhängig, falls für alle $\alpha_i \in K$, $i \in \{1, \dots, m\}$, aus

$$\sum_{i=1}^m \alpha_i e_i = 0$$

$\alpha_1 = \dots = \alpha_m = 0$ folgt.

Satz 7.2.2. Sei $E \subset V$ eine Teilmenge eines Vektorraumes V . Dann sind folgende Aussagen äquivalent:

- a) E ist linear unabhängig.
- b) Jeder Vektor $x \in \text{Span}(E)$ lässt sich auf genau eine Weise mit Vektoren aus E linear darstellen, d.h. es existieren eindeutig bestimmte Skalare α_e mit $\alpha_e \neq 0$ für höchstens endlich viele $e \in E$, so dass

$$x = \sum_{e \in E} \alpha_e e.$$

Beweis. “a) \Rightarrow b)”

Sei E linear unabhängig und $x \in \text{Span}(E)$. Seien

$$x = \sum_{e \in E} \alpha_e e = \sum_{e \in E} \beta_e e,$$

mit $\alpha_e, \beta_e \neq 0$ für höchstens endlich viele $e \in E$ zwei Linearkombinationen, die x darstellen. Dann ist

$$0 = \sum_{e \in E} (\alpha_e - \beta_e)e,$$

wobei die Skalare $c_e := \alpha_e - \beta_e$ für höchstens endlich viele $e \in E$ von null verschieden sind. Daher ist $c_e = \alpha_e - \beta_e = 0$ für alle $e \in E$.

“b) \Rightarrow a)”

Sei nun b) erfüllt. Betrachte $\sum_{e \in E} \alpha_e e = 0$. Auf der anderen Seite lässt sich der Nullvektor auch durch die triviale Linearkombination darstellen, d.h es gilt: $\sum_{e \in E} \beta_e e = 0$ mit $\beta_e = 0$ für alle $e \in E$. Also folgt aus b): $\alpha_e = \beta_e = 0$ für alle $e \in E$. \square

Bemerkung. Ist E linear abhängig, so ist jedes $E' \supset E$ ebenfalls linear abhängig. Ist E linear unabhängig, so ist auch $E' \subset E$ linear unabhängig, insbesondere ist $0 \notin E$.

Beispiele.

1. Sei V ein Vektorraum, $x \in V$. Dann gilt: $\{x\}$ ist genau dann linear unabhängig, falls $x \neq 0$, denn ist $x \neq 0$, so folgt aus $\alpha x = 0$, $\alpha = 0$. Ist $x = 0$, so folgt $\alpha x = 0$ für alle $\alpha \in K$, d.h. $\{x\}$ ist linear abhängig.
2. Sei K ein Körper und $V = K^n$. Dann ist

$$E = \{e_1, \dots, e_n\} \text{ mit } e_i = (0, \dots, 0, \overset{i\text{-te Stelle}}{1}, 0, \dots, 0)$$

linear unabhängig. Denn

$$\sum_{i=1}^n \alpha_i e_i = (\alpha_1, \dots, \alpha_n) = 0$$

genau dann, wenn $\alpha_1 = \dots = \alpha_n = 0$. Genauso sieht man, dass E ein Erzeugendensystem von K^n ist.

Es gibt eine weitere Kennzeichnung linear unabhängiger Mengen.

Satz 7.2.3. Sei V ein Vektorraum und $E \subset V$. Dann sind folgende Aussagen äquivalent:

- i) E ist linear unabhängig.
- ii) Aus $E' \subsetneq E$ folgt $\text{Span}(E') \subsetneq \text{Span}(E)$, d.h. E ist minimales Erzeugendensystem von $\text{Span}(E)$.

Beweis. “i) \Rightarrow ii)”:

Ist $x \in E \setminus E'$, so ist $x \notin \text{Span}(E')$, denn sonst ist $1x = x = \sum_{e \in E'} \alpha_e e$ und somit ist E linear abhängig.

“ii) \Rightarrow i)”:

Angenommen, E ist linear abhängig. Dann existiert eine Linearkombination $\sum_{e \in E} \alpha_e e = 0$ mit $\alpha_{e_0} \neq 0$ für ein $e_0 \in E$. Damit besitzt aber e_0 die Darstellung

$$e_0 = -\frac{1}{\alpha_{e_0}} \sum_{e \in E \setminus \{e_0\}} \alpha_e e \in \text{Span}(E \setminus \{e_0\}).$$

Insbesondere folgt: $\text{Span}(E \setminus \{e_0\}) = \text{Span}(E)$ und somit ist E nicht minimales Erzeugendensystem von $\text{Span}(E)$. \square

Lemma 7.2.4. Sei $E \subset V$ linear unabhängig, aber $\text{Span}(E) \neq V$. Dann ist $E \cup \{x\}$ für jedes $x \in V \setminus \text{Span}(E)$ linear unabhängig.

Beweis. Sei $\sum_{e \in E} \alpha_e e + \alpha x = 0$, für $\alpha \in K$. Da $x \notin \text{Span}(E)$ ist, ist $\alpha = 0$. Da E linear unabhängig ist, folgt dann $\alpha_e = 0$ für alle $e \in E$. \square

Definition 7.2.5. Sei V ein Vektorraum und $B \subset V$. Dann heißt B *Basis* von V genau dann, wenn B ein linear unabhängiges Erzeugendensystem von V ist, d.h. die Menge B ist linear unabhängig und $\text{Span}(B) = V$.

Beispiele.

1. Die leere Menge $E = \emptyset$ ist Basis von $V = \{0\}$.
2. Ist K ein Körper, so ist $\{e_1, \dots, e_n\} = B$ mit $e_i = (0, \dots, 0, \overset{i\text{-te Stelle}}{1}, 0, \dots, 0)$ eine Basis von K^n .
3. $\{(1, 2), (0, 1)\}$ ist eine Basis des \mathbb{R}^2 .
4. Sei

$$P(\mathbb{R}) = \{p : \mathbb{R} \rightarrow \mathbb{R} \mid p(x) = \sum_{i=0}^n a_i x^i, n \in \mathbb{N}_0, a_i \in \mathbb{R}\}$$

der Vektorraum der reellen Polynome. Dann ist die Menge

$$E = \{1, x, x^2, x^3, \dots, x^n, \dots\}$$

eine Basis von $P(\mathbb{R})$. Die Elemente in E heißen auch *Monome*.

Satz 7.2.6. Sei V ein Vektorraum und $B \subset V$. Dann sind folgende Aussagen äquivalent:

1. B ist Basis von V .
2. B ist minimales Erzeugendensystem von V , d.h. $\text{Span}(B) = V$ und keine kleinere Menge erzeugt V . Ist also $E \subsetneq B$, so ist $\text{Span}(E) \neq V$.
3. B ist maximale linear unabhängige Menge in V , d.h. ist E eine Teilmenge von V mit $B \subsetneq E$, so ist E linear abhängig.

Beweis. “1 \Leftrightarrow 2” folgt unmittelbar aus Satz 7.2.3. Betrachte dazu $V = \text{Span}(B)$.

“1 \Rightarrow 3”: Sei B eine Basis von V . Dann ist B eine maximale linear unabhängige Menge. Denn ist $B \subsetneq E \subset V$, so existiert ein $x \in E \setminus B$. Da B eine Basis ist, lässt sich x als Linearkombination von Elementen aus B darstellen. Also ist

$$x = \sum_{e \in B} \alpha_e e.$$

Insbesondere ist E linear abhängig.

“3 \Rightarrow 1”: Sei B maximale linear unabhängige Menge in V . Dann muss B eine Basis sein, denn sonst existiert ein $x \in V \setminus \text{Span}(B)$ und wegen Lemma 7.2.4 ist $B \cup \{x\} = E$ linear unabhängig. Somit wäre B nicht die größte linear unabhängige Teilmenge von V . \square

Nun wollen wir die Existenz einer Basis für einen beliebigen Vektorraum beweisen. Dafür erweist sich folgende Definition als sehr hilfreich.

Sei V ein Vektorraum, $F \subset V$ eine linear unabhängige Teilmenge und $E \supset F$ ein Erzeugendensystem von V . Man betrachte die Menge

$$P(F, E) = \{G \mid G \text{ linear unabhängig, } F \subset G \subset E\}$$

aller linear unabhängigen Teilmengen zwischen F und E . Es wird sich herausstellen, dass die “maximalen Elemente” in $P(F, E)$ Basen von V darstellen.

Definition 7.2.7. $B \in P(F, E)$ heißt *maximal* genau dann, wenn für alle $B' \in P(F, E)$ mit $B \subset B'$ folgt: $B = B'$.

Satz 7.2.8. Sei V ein Vektorraum, $F \subset V$ eine linear unabhängige Teilmenge und $E \supset F$ ein Erzeugendensystem von V . Ist B maximal in $P(F, E)$, so ist B eine Basis von V .

Beweis. Der Beweis dieser Aussage ist bis auf eine kleine Modifikation identisch mit dem Beweis der Implikation “3 \Rightarrow 1” in Satz 7.2.6.

Sei also $B \in P(F, E)$ ein maximales Element. Dann muss B eine Basis sein, denn andernfalls ist $\text{Span}(B) \subsetneq V$. Aus $\text{Span}(E) = V$ folgt dann die Existenz eines Elementes $x \in E \setminus \text{Span}(B)$, denn sonst wäre $\text{Span}(E) = \text{Span}(B)$. Wegen Lemma 7.2.4 ist $B' = B \cup \{x\}$ linear unabhängig. Aber dann ist B keine maximale linear unabhängige Teilmenge von E . \square

Satz 7.2.9. Sei $F \subset E \subset V$, V Vektorraum, F linear unabhängig und E Erzeugendensystem. Dann existiert eine Basis B mit $F \subset B \subset E$.

Beweis. Wir zeigen die Existenz eines maximalen Elementes. Betrachte $F \in P(F, E)$. Ist F nicht maximal, so gibt es $F' \in P(F, E)$ mit $F \subsetneq F' \subset E$. Ist F' nicht maximal, so gibt es ein $F'' \in P(F, E)$ mit $F' \subsetneq F'' \subset E$. Auf dieser Weise produzieren wir eine Kette

$$F \subsetneq F' \subsetneq F'' \subsetneq \dots \subset E.$$

Ist E endlich und $|E| = n$ die Anzahl ihrer Elemente, so muss diese Kette nach spätestens n Schritten abbrechen. Somit erhalten wir ein maximales Element. Ist E unendlich, so beweist man die Existenz eines maximalen Elementes mit dem Zornschen Lemma. Dies entspricht einem Axiom aus der Mengenlehre, dem sog. Auswahlaxiom. \square

Wichtige Folgerungen hieraus können wir in folgendem Satz zusammenfassen.

Satz 7.2.10. Sei V ein Vektorraum. Dann gilt:

1. V besitzt eine Basis.
2. Jede linear unabhängige Menge F kann zu einer Basis ergänzt werden, d.h. es gibt eine Basis B mit $F \subset B$.
3. Jedes Erzeugendensystem E enthält eine Basis.
4. (Basisergänzungssatz): Sei F eine linear unabhängige Teilmenge und E ein Erzeugendensystem von V . Dann kann F durch eine Teilmenge $B' \subset E$ zu einer Basis $B = F \cup B'$ ergänzt werden.

Beweis. Zum Beweis wählen wir in Satz 7.2.9 die linear unabhängige Teilmenge und das Erzeugendensystem in geeigneter Weise.

zu 1) Die leere Menge \emptyset ist linear unabhängig und der Vektorraum V ist trivialerweise ein Erzeugendensystem von V . Also ist jedes maximale Element in $B(\emptyset, V)$ eine Basis von V .

zu 2) Ist B ein maximales Element in $B(F, V)$, so ist B eine Basis, die F enthält.

zu 3) Ist B maximales Element in $B(\emptyset, E)$, so ist B eine Basis, die in E enthalten ist.

zu 4) Ist E ein Erzeugendensystem von V , so auch $F \cup E$. Also ist jedes maximale Element B in $B(F, F \cup E)$ eine Basis mit

$$F \subset B \subset F \cup E.$$

Also existiert eine Teilmenge $B \setminus F = B' \subset E$ mit $B = F \cup B'$.

□

Nun wollen wir zeigen, dass in einem endlich erzeugten Vektorraum verschiedene Basen gleich viele Elemente besitzen. Als Vorbereitung dazu dient das folgende Lemma. Es stellt in gewisser Weise eine Präzisierung von Satz 7.2.10(4) dar.

Lemma 7.2.11. (Austauschlemma)

Sei V ein Vektorraum, $B \subset V$ eine Basis und $E \subset V$ ein Erzeugendensystem. Dann folgt: für alle $b' \in B$ existiert ein $e \in E$, so dass

$$(B \setminus \{b'\}) \cup \{e\}$$

eine Basis ist.

Beweis. Ist $b' \in B$, so ist $B \setminus \{b'\}$ wegen Satz 7.2.6 kein Erzeugendensystem von V . Insbesondere ist $E \not\subset \text{Span}(B \setminus \{b'\})$. Dann existiert ein $e \in E$ mit $e \notin \text{Span}(B \setminus \{b'\})$ und daher ist wegen Lemma 7.2.4 die Menge $(B \setminus \{b'\}) \cup \{e\}$ linear unabhängig.

Nun zeigen wir, dass $\text{Span}((B \setminus \{b'\}) \cup \{e\})$ ein Erzeugendensystem von V ist: Es genügt zu zeigen: $b' \in \text{Span}((B \setminus \{b'\}) \cup \{e\})$, denn dann ist $\text{Span}((B \setminus \{b'\}) \cup \{e\}) = \text{Span}(B)$. Da $e \notin \text{Span}(B \setminus \{b'\})$, aber B eine Basis von V ist, wird e durch eine Linearkombination der Form

$$e = \sum_{b \in B \setminus \{b'\}} \alpha_b b + \alpha_{b'} b'$$

mit $\alpha_{b'} \neq 0$ dargestellt. Somit besitzt b' die Darstellung

$$b' = \frac{1}{\alpha_{b'}} e - \frac{1}{\alpha_{b'}} \sum_{b \in B \setminus \{b'\}} \alpha_b b \in \text{Span}((B \setminus \{b'\}) \cup \{e\}).$$

□

Im Folgenden bezeichnen wir mit $|M|$ die Anzahl der Elemente einer endlichen Menge M .

Satz 7.2.12. (Austauschsatz)

Sei V ein Vektorraum, $B \subset V$ eine Basis und E ein endliches Erzeugendensystem. Dann gibt es eine Teilmenge $B' \subset E$ mit B' ist Basis und $|B| = |B'|$. Insbesondere ist $|B| \leq |E|$.

Beweis. Sei B eine Basis von V . Wir werden sukzessive die Elemente von B durch die Elemente von E mit Hilfe des Austauschlemmas ersetzen.

Wegen des Austauschlemmas gibt es zu jedem $b_1 \in B$ ein $e_1 \in E$, so dass

$$B^{(1)} = (B \setminus \{b_1\}) \cup \{e_1\}$$

eine Basis ist. Wenden wir nun das Austauschlemma auf $B = B^{(1)}$ und E an, so können wir zu $b_2 \in B \setminus \{b_1\}$ ein $e_2 \in E$, $e_2 \neq e_1$ wählen, so dass auch

$$B^{(2)} = (B \setminus \{b_1, b_2\}) \cup \{e_1, e_2\}$$

eine Basis ist. Die iterative Anwendung dieses Verfahrens zeigt nun, dass die Basis B nicht mehr Elemente als E enthält. Denn hätte B mehr Elemente als E und ist $E = \{e_1, \dots, e_n\}$, so würden wir nach n -maliger Anwendung des Austauschlemmas eine Basis

$$B^{(n)} = (B \setminus \{b_1, \dots, b_n\}) \cup E$$

erhalten mit $B \setminus \{b_1, \dots, b_n\} \neq \emptyset$. Damit wäre wegen Satz 7.2.6 die Menge E als echte Teilmenge von $B^{(n)}$ kein Erzeugendensystem von V . Also gilt: $k := |B| \leq |E| = n$ und nach k -maliger Anwendung des Austauschlemmas erhalten wir eine Basis

$$B' := B^{(k)} = \{e_1, \dots, e_k\} \subset E.$$

□

Als unmittelbare Folgerung erhalten wir:

Satz 7.2.13. Sei V ein endlich erzeugter Vektorraum. Dann besitzt jede Basis die gleiche Anzahl von Elementen.

Beweis. Sei E ein endliches Erzeugendensystem von V und B eine Basis. Dann ist wegen des Austauschsatzes 7.2.12 $|B| \leq |E|$. Ist B' eine weitere Basis, so ist $|B'| \leq |B|$, denn B ist ein endliches Erzeugendensystem. Genauso gilt: $|B| \leq |B'|$, denn B' ist endliches Erzeugendensystem. □

Definition 7.2.14. Sei V ein endlich erzeugter Vektorraum und B eine Basis. Dann heißt

$$\dim V := |B|$$

die *Dimension* von V . Ist V nicht endlich erzeugt, so nennen wir V *unendlich dimensional* und schreiben

$$\dim V = \infty.$$

Beispiele.

1. Für den Nullraum $V = \{0\}$ ist die leere Menge eine Basis. Also ist $\dim V = 0$.
2. Sei K ein Körper und $V = K^n$. Dann ist $\dim V = n$, denn $B = \{e_1, \dots, e_n\}$ mit $e_i = (0, \dots, 0, \overset{i\text{-te Stelle}}{1}, 0, \dots, 0)$ ist eine Basis.
3. $V = C^0(\mathbb{R})$ ist unendlich dimensional, denn z.B. enthält $C^0(\mathbb{R})$ die Menge aller Polynome $P(\mathbb{R}) = \{p \in \mathbb{R}^{\mathbb{R}} \mid p(x) = \sum_{i=0}^n a_i x^i, a_i \in \mathbb{R}, n \in \mathbb{N}\}$ als Unterraum.

Satz 7.2.15. *Ist V ein Vektorraum mit $\dim V = n$, so gilt: Ist F linear unabhängig, so ist $|F| \leq n$. Ist E ein Erzeugendensystem, so ist $|E| \geq n$. Der Gleichheitsfall tritt nur dann ein, falls F bzw. E eine Basis sind.*

Beweis. Ist F linear unabhängig, so folgt aus Satz 7.2.10(2) die Existenz einer Basis mit $F \subset B$. Dann ist $|F| \leq n$ und die Gleichheit würde $F = B$ implizieren. Aus Satz 7.2.10(3) folgt, dass E eine Basis B enthält und damit gilt: $n \leq |E|$. Ist $n = |E|$ so folgt: $E = B$. \square

7.3 Lineare Komplemente und Dimensionsformel

Sei $U \subset W$ ein Unterraum in einem Vektorraum W . Wir hatten einen Unterraum $V \subset W$ ein Komplement (*lineares Komplement*) von U genannt, falls

$$U \cap V = \{0\} \text{ und } U + V = W .$$

Wir schreiben auch $U \oplus V = W$.

Lemma 7.3.1. *Sei $W = U \oplus V$. Dann ist für jedes $x \in W$ die Zerlegung $x = x_u + x_v$ mit $x_u \in U$ und $x_v \in V$ eindeutig bestimmt.*

Beweis. Die Existenz der Zerlegung ist schon gezeigt (siehe Übung, S. 146). Sei $x_u + x_v = x'_u + x'_v$ mit $x_u, x'_u \in U$ und $x_v, x'_v \in V$. Daraus folgt: $x_u - x'_u = x'_v - x_v \in U \cap V \Rightarrow x_u = x'_u$ und $x_v = x'_v$. \square

Wir zeigen nun, dass jeder Unterraum ein Komplement besitzt.

Satz 7.3.2. *Zu jedem Unterraum U von W existiert ein Komplement V .*

Beweis. Sei $B_1 \subset U$ eine Basis von U . Nach Satz 7.2.10(2) kann B_1 zu einer Basis B von W ergänzt werden, d.h. es existiert $B_2 \subset W$, so dass $B = B_1 \cup B_2$ eine Basis von W ist. Setze $V := \text{Span}(B_2)$. Dann gilt: $V \cap U = \{0\}$. Denn ist $x \in U \cap V$, so lässt sich x sowohl als Linearkombination mit Elementen aus B_1 als auch mit Elementen aus B_2 schreiben. Also folgt

$$x = \sum_{b \in B_1} \alpha_b b = \sum_{b \in B_2} \beta_b b .$$

Durch Subtraktion der Gleichungen erhalten wir dann

$$0 = \sum_{b \in B_1} \alpha_b b + \sum_{b \in B_2} (-\beta_b) b .$$

Da die Menge $B_1 \cup B_2$ linear unabhängig ist, sind alle Koeffizienten gleich 0. Also folgt auch $x = 0$. Wegen

$$W \supset U + V = \text{Span}(U \cup V) \supset \text{Span}(B_1 \cup B_2) = W$$

erhalten wir: $W = U + V$. \square

Nun beweisen wir die folgende Dimensionsformel für die Summe $U + V$ zweier endlich dimensionaler Unterräume von W .

Satz 7.3.3. (Dimensionsformel)

Es seien U, V endlich dimensionale Unterräume von W . Dann gilt:

$$\dim(U + V) = \dim U + \dim V - \dim(U \cap V) .$$

Beweis. Sei $B_d = \{b_1, \dots, b_r\}$ eine Basis von $U \cap V$. Da $U \cap V$ Unterraum von U und V ist, kann wegen Satz 7.2.10(2) B_d zu einer Basis

$$B_1 = \{b_1, \dots, b_r, x_1, \dots, x_n\}$$

von U und zu einer Basis

$$B_2 = \{b_1, \dots, b_r, y_1, \dots, y_m\}$$

von V ergänzt werden. Dann ist

$$B_1 \cup B_2 = \{b_1, \dots, b_r, x_1, \dots, x_n, y_1, \dots, y_m\}$$

Erzeugendensystem von $U + V$, denn ist $x \in U + V$, so existieren $x_u \in U$ und $x_v \in V$ mit $x = x_u + x_v \in \text{Span}(B_1 \cup B_2)$.

$B_1 \cup B_2$ ist linear unabhängig, denn ist

$$\sum_{i=1}^r \alpha_i b_i + \sum_{i=1}^n \beta_i x_i + \sum_{i=1}^m \gamma_i y_i = 0,$$

so folgt

$$\sum_{i=1}^r \alpha_i b_i + \sum_{i=1}^n \beta_i x_i = - \sum_{i=1}^m \gamma_i y_i \in V.$$

Da die linke Seite in U liegt, erhalten wir

$$\sum_{i=1}^m \gamma_i y_i \in V \cap U \quad \text{und somit auch} \quad \sum_{i=1}^n \beta_i x_i \in V \cap U.$$

Insbesondere existieren Koeffizienten $\gamma'_1, \dots, \gamma'_r$ mit $\sum_{i=1}^n \beta_i x_i = \sum_{i=1}^r \gamma'_i b_i$. Da B_1 als Basis linear unabhängig ist, gilt

$$\beta_1 = \dots = \beta_n = \gamma'_1 = \dots = \gamma'_r = 0.$$

Da B_2 ebenfalls eine Basis ist, gilt dann auch

$$\alpha_1 = \dots = \alpha_r = \gamma_1 = \dots = \gamma_m = 0.$$

Also folgt:

$$\dim(U + V) = m + n + r = (m + r) + (n + r) - r = \dim U + \dim V - \dim(U \cap V).$$

□

Eine Anwendung der Dimensionsformel ist das folgende Korollar.

Korollar 7.3.4. In einem endlich dimensionalen Vektorraum W sei V Komplement von U . Dann gilt:

$$\dim W = \dim U + \dim V.$$

Beweis. Dies folgt aus Satz 7.3.3, denn $U + V = W$ und $U \cap V = \{0\}$.

□

Bemerkung. Insbesondere haben in einem endlich dimensionalen Vektorraum W alle Komplemente eines Unterraumes U die gleiche Dimension, nämlich $\dim W - \dim U$.

Zum Ende dieses Abschnittes wollen wir noch den Begriff der direkten Summe auf eine beliebige Anzahl von Unterräumen ausdehnen.

Definition 7.3.5. Es seien $U_1 \dots U_n$ Unterräume eines Vektorraumes V . Die in 7.1.10 definierte Summe

$$U_1 + \dots + U_n$$

der Vektorräume U_i heißt *direkt*, falls aus

$$x_1 + \dots + x_n = 0$$

mit $x_i \in U_i$ folgt:

$$x_1 = x_2 = \dots = x_n = 0.$$

Ist die Summe der U_i direkt, so schreiben wir

$$U_1 \oplus \dots \oplus U_n.$$

Bemerkungen.

1. Jede der beiden folgenden Eigenschaften ist äquivalent zur Direktheit der Summe

$$U := U_1 + \dots + U_n.$$

- (a) Zu jedem $x \in U$ existieren eindeutig bestimmte Vektoren $x_i \in U_i$ mit

$$x = x_1 + \dots + x_n.$$

- (b) Für jeden Unterraum U_i gilt:

$$U_i \cap (U_1 + \dots + U_{i-1} + U_{i+1} + \dots + U_n) = \{0\}.$$

Insbesondere ist die Summe zweier Unterräume $U_1, U_2 \subset V$ direkt, falls gilt:

$$U_1 \cap U_2 = \{0\}.$$

2. Aus der Dimensionsformel folgt

$$\dim(U_1 + \dots + U_n) \leq \dim U_1 + \dots + \dim U_n.$$

Wie das nächste Korollar zeigt, gilt Gleichheit nur dann, falls die Summe direkt ist.

Korollar 7.3.6. *Es seien U_1, \dots, U_n endlich dimensionale Unterräume eines Vektorraumes V und $U = U_1 + \dots + U_n$. Dann sind folgende Aussagen äquivalent:*

- 1.

$$U = U_1 \oplus \dots \oplus U_n,$$

d.h. die Summe der U_i ist direkt.

2.

$$\dim U = \dim U_1 + \dots + \dim U_n$$

Beweis. Der Beweis der Äquivalenz wird mit Hilfe der vollständigen Induktion über n geführt. Für $n = 1$ ist nichts zu zeigen. Wir nehmen also an, dass die Aussage schon für $n \in \mathbb{N}$ bewiesen sei. Betrachte nun $(n + 1)$ endlich dimensionale Unterräume U_1, \dots, U_{n+1} und es sei $U = U_1 + \dots + U_{n+1}$. Für $i \in \{1, \dots, n + 1\}$ definiere

$$W_i := U_1 + \dots + U_{i-1} + U_{i+1} + \dots + U_{n+1}.$$

Dann folgt aus der Dimensionsformel

$$\dim U = \dim(U_i + W_i) = \dim U_i + \dim W_i - \dim(U_i \cap W_i). \quad (*)$$

Ist die Summe $U = U_1 + \dots + U_{n+1}$ direkt, so gilt nach obiger Bemerkung: $U_i \cap W_i = \{0\}$ und somit ist $\dim(U_i \cap W_i) = 0$. Außerdem ist dann auch die Summe

$$W_i := U_1 + \dots + U_{i-1} + U_{i+1} + \dots + U_{n+1}$$

direkt und nach Induktionsvoraussetzung folgt

$$\dim U = \dim U_1 + \dots + \dim U_{n+1}.$$

Ist nun umgekehrt diese Gleichung erfüllt, so folgt aus (*) und aus Teil 2 der obigen Bemerkung

$$\begin{aligned} \dim(U_i \cap W_i) &= \dim W_i + \dim U_i - \dim U \\ &= \dim W_i - (\dim U_1 + \dots + \dim U_{i-1} + \dim U_{i+1} + \dots + \dim U_{n+1}) \\ &\leq 0. \end{aligned}$$

Damit ist $U_i \cap W_i = \{0\}$ und die Summe ist direkt. □

Kapitel 8

Lineare Abbildungen

Wir betrachten nun Abbildungen, die mit der (linearen) Struktur der Vektorräume kompatibel sind. Diese Abbildungen werden wir lineare Abbildungen nennen.

8.1 Grundlegendes über lineare Abbildungen

Definition 8.1.1. Es seien V, W K -Vektorräume, d.h. Vektorräume über einem Körper K . Eine Abbildung $f : V \rightarrow W$ heißt *linear* (K -linear), wenn gilt:

1. $f(x + y) = f(x) + f(y)$, für alle $x, y \in V$.
2. $f(\alpha x) = \alpha f(x)$, für alle $\alpha \in K$ und $x \in V$.

Bemerkungen.

1. Aus der Definition folgt, dass unter einer linearen Abbildung der Nullvektor auf den Nullvektor abgebildet wird, denn ist $0 \in V$, so folgt:

$$f(0) = f(0 + 0) = f(0) + f(0)$$

und somit ist $f(0) = 0 \in W$.

2. Im Falle unendlich dimensionaler Räume werden lineare Abbildungen auch manchmal *Operatoren* genannt.

Beispiele.

1. Die *Nullabbildung* $0 : V \rightarrow W$ mit $x \mapsto 0$ und die *Identität* $\text{id}_V : V \rightarrow V$ mit $x \mapsto x$ sind lineare Abbildungen.
2. Seien K ein Körper, $V = K^n$ und $W = K$. Für $i \in \{1, \dots, n\}$ ist $\text{pr}_i : K^n \rightarrow K$ mit $\text{pr}_i(x_1, \dots, x_n) = x_i$ linear.
3. Es seien $U, V \subset W$ Unterräume mit $W = U \oplus V$. Dann lässt sich wegen Lemma 7.3.1 jeder Vektor $x \in W$ eindeutig in der Form $x = x_u + x_v$ mit $x_u \in U$ und $x_v \in V$ schreiben. Die zugehörigen Abbildungen $\text{pr}_U : W \rightarrow U$ mit $x \rightarrow x_u$ und $\text{pr}_V : W \rightarrow V$ mit $x \rightarrow x_v$ heißen *Projektionen* auf U bzw. V . Beide Projektionen sind linear, denn

aus $x = x_u + x_v$ und $y = y_u + y_v$ folgt $x + y = (x_u + y_u) + (x_v + y_v)$ und somit erhalten wir sowohl

$$\text{pr}_U(x + y) = (x + y)_u = x_u + y_u = \text{pr}_U(x) + \text{pr}_U(y)$$

als auch

$$\text{pr}_V(x + y) = (x + y)_v = x_v + y_v = \text{pr}_V(x) + \text{pr}_V(y).$$

Genauso sieht man $\text{pr}_U(\alpha x) = \alpha \text{pr}_U(x)$ bzw. $\text{pr}_V(\alpha x) = \alpha \text{pr}_V(x)$ für alle $\alpha \in K$ und $x \in W$.

Aus der Analysis stammen folgende Beispiele:

4. Sei $C^1(\mathbb{R}) = \{f \in \mathbb{R}^{\mathbb{R}} \mid f \text{ stetig differenzierbar}\}$ der \mathbb{R} -Vektorraum der stetig differenzierbaren Abbildungen. Dann ist die Ableitung

$$D : C^1(\mathbb{R}) \rightarrow C^0(\mathbb{R})$$

mit $D(f)(x) = f'(x)$ eine lineare Abbildung. Dies folgt aus den Rechenregeln für die Ableitung (siehe Satz 4.3.1). D heißt auch *Differentialoperator*.

5. Sei $R([a, b])$ der Vektorraum der Regelfunktionen auf dem Intervall $[a, b] \subset \mathbb{R}$. Dann ist

$$\int : R([a, b]) \rightarrow \mathbb{R}$$

mit $f \mapsto \int f$ eine lineare Abbildung.

6. Sei M eine nichtleere Menge, V ein K -Vektorraum. Ist $f : M \rightarrow V$ eine beliebige Abbildung, so ist $F : V^M \rightarrow V^M$ mit $\phi \mapsto \phi \circ f$ eine lineare Abbildung (Übung).

Nun zeigen wir, dass die Hintereinanderschaltung von linearen Abbildungen wieder linear ist.

Satz 8.1.2. *Es seien U, V, W K -Vektorräume und $f : U \rightarrow V$ sowie $g : V \rightarrow W$ lineare Abbildungen. Dann ist auch ihre Komposition $F = g \circ f : U \rightarrow W$ linear.*

Beweis. Seien $x, y \in U$, so gilt

$$(g \circ f)(x + y) = g(f(x) + f(y)) = g(f(x)) + g(f(y)) = (g \circ f)(x) + (g \circ f)(y).$$

Ist $x \in U$ und $\alpha \in K$, so gilt

$$(g \circ f)(\alpha x) = g(f(\alpha x)) = g(\alpha f(x)) = \alpha g(f(x)) = \alpha (g \circ f)(x).$$

□

Definition 8.1.3. Eine bijektive lineare Abbildung $f : V \rightarrow W$ heißt *Isomorphismus*. Ist $V = W$, so nennt man f auch *Automorphismus*. Zwei Vektorräume V und W heißen *isomorph* (man schreibt $V \cong W$), falls ein Isomorphismus $f : V \rightarrow W$ existiert.

Satz 8.1.4. *Ist $f : V \rightarrow W$ ein Isomorphismus, so auch $f^{-1} : W \rightarrow V$.*

Beweis. Zu zeigen ist die Linearität der Umkehrabbildung $f^{-1} : W \rightarrow V$. Seien $y_1, y_2 \in W$ und $x_1, x_2 \in V$ die eindeutig bestimmten Elemente in V mit $f(x_1) = y_1$ und $f(x_2) = y_2$. Dann gilt:

$$f^{-1}(y_1 + y_2) = f^{-1}(f(x_1) + f(x_2)) = f^{-1}(f(x_1 + x_2)) = x_1 + x_2 = f^{-1}(y_1) + f^{-1}(y_2).$$

Ist $\alpha \in K$ und $y = f(x) \in W$, so folgt:

$$f^{-1}(\alpha y) = f^{-1}(\alpha f(x)) = f^{-1}(f(\alpha x)) = \alpha x = \alpha f^{-1}(y).$$

□

Korollar 8.1.5. Sei V ein Vektorraum über K . Dann heißt

$$\text{GL}(V) = \{f : V \rightarrow V \mid f \text{ linearer Automorphismus}\}$$

die allgemeine lineare Gruppe (*general linear group*). Sie ist eine Untergruppe von $S(V)$ bezüglich der Komposition von Abbildungen.

Beweis. Wegen der Sätze 8.1.2 und 8.1.4 sind mit $f, g \in \text{GL}(V)$ sowohl $g \circ f$ als auch f^{-1} Elemente in $\text{GL}(V)$. □

Nun zeigen wir, dass Bilder und Urbilder von Unterräumen unter linearen Abbildungen wieder Unterräume sind.

Satz 8.1.6. Es seien V, W K -Vektorräume und $f : V \rightarrow W$ eine lineare Abbildung. Sind $U \subset V$ und $U' \subset W$ Unterräume, so auch das Bild $f(U) \subset W$ und das Urbild $f^{-1}(U') \subset V$.

Beweis.

1. Zunächst zeigen wir, dass $f(U) \subset W$ abgeschlossen bezüglich der Addition ist. Denn sind $x_1 = f(u_1)$, $x_2 = f(u_2)$ mit $u_1, u_2 \in U$, so ist $x_1 + x_2 = f(u_1 + u_2)$. Da U Unterraum von V ist, folgt $f(u_1 + u_2) \in f(U)$.

$f(U) \subset W$ ist auch abgeschlossen bezüglich der skalaren Multiplikation, denn ist $x = f(u)$, $u \in U$, so ist $\alpha x = \alpha f(u) = f(\alpha u) \in f(U)$, für $\alpha \in K$.

2. $f^{-1}(U')$ ist abgeschlossen bezüglich der Addition. Denn sind $x, y \in f^{-1}(U')$, so sind $f(x), f(y) \in U'$. Da U' Unterraum von W ist, folgt mittels der Linearität von f :

$$f(x + y) = f(x) + f(y) \in U',$$

also ist $x + y \in f^{-1}(U')$. Außerdem ist $f^{-1}(U')$ abgeschlossen bezüglich der skalaren Multiplikation. Denn ist $x \in f^{-1}(U')$ und $\alpha \in K$, so folgt $f(\alpha x) = \alpha f(x) \in U'$. Somit ist $\alpha x \in f^{-1}(U')$.

□

Korollar 8.1.7. Sei $f : V \rightarrow W$ eine lineare Abbildung, so ist

$$\text{Kern } f = f^{-1}(\{0\}) = \{x \in V \mid f(x) = 0\}$$

ein Unterraum von V und

$$\text{Bild } f = f(V) = \{y \in W \mid f(x) = y, x \in V\}$$

ein Unterraum von W .

Bemerkungen.

1. $f : V \rightarrow W$ ist nach Definition genau dann surjektiv, falls $\text{Bild } f = W$.
2. $f : V \rightarrow W$ ist genau dann injektiv, falls $\text{Kern } f = \{0\}$. Denn ist f injektiv und ist $x \in \text{Kern } f$, so folgt $f(x) = f(0) = 0$ und somit auch $x = 0$. Ist umgekehrt $\text{Kern } f = \{0\}$ und $f(x) = f(y)$, so folgt $f(x - y) = 0$ und somit auch $x - y = 0$.

Beispiele.

1. $0 : V \rightarrow W$ sei die Nullabbildung. Dann ist $\text{Kern } 0 = V$ und $\text{Bild } 0 = \{0\}$.
Sei $\text{id}_V : V \rightarrow V$ die Identität. Dann ist $\text{Kern } \text{id}_V = \{0\}$ und $\text{Bild } \text{id}_V = V$.
2. Ist $\text{pr}_i : K^n \rightarrow K$ die Projektion auf die i -te Komponente, so ist

$$\text{Kern } \text{pr}_i = \{(x_1, \dots, x_n) \mid x_i = 0\}$$

und $\text{Bild } \text{pr}_i = K$.

3. Sei $W = U \oplus V$ und $\text{pr}_U : W \rightarrow U$ mit $x \rightarrow x_u$ die Projektionsabbildung, so gilt:
 $\text{Kern } \text{pr}_U = V$ und $\text{Bild } \text{pr}_U = U$.
4. Sei $D : C^1(\mathbb{R}) \rightarrow C^0(\mathbb{R})$ die Ableitung, so besteht $\text{Kern } D$ aus der Menge der konstanten Abbildungen. Außerdem ist $\text{Bild } D = C^0(\mathbb{R})$, denn ist $g \in C^0(\mathbb{R})$, so ist wegen des Hauptsatzes der Differential- und Integralrechnung (siehe Satz 5.2.1) die Funktion $G(x) := \int_0^x g(s) ds$ in $C^1(\mathbb{R})$ und $D(G)(x) = g(x)$.

Satz 8.1.8. *Es seien V, W K -Vektorräume und $f : V \rightarrow W$ eine lineare Abbildung. Sei U ein zu $\text{Kern } f$ komplementärer Unterraum von V . Dann ist*

$$f|_U : U \rightarrow \text{Bild } f$$

ein Isomorphismus.

Beweis. Wir müssen zeigen, dass die Abbildung $f|_U : U \rightarrow \text{Bild } f$ bijektiv ist. Sei also $y \in \text{Bild } f$, so existiert ein $x \in V$ mit $f(x) = y$. Da

$$V = U \oplus \text{Kern } f,$$

existiert eine Zerlegung $x = x_1 + x_2$, mit $x_1 \in U$ und $x_2 \in \text{Kern } f$. Dann gilt:

$$f(x_1) = f(x_1) + f(x_2) = f(x_1 + x_2) = f(x) = y$$

und somit folgt die Surjektivität. Die Abbildung $f|_U : U \rightarrow \text{Bild } f$ ist aber auch injektiv, denn ist $x \in U$ mit $f(x) = 0$, so ist $x \in \text{Kern } f \cap U = \{0\}$. \square

Definition 8.1.9. (*Lineare Gleichungen*)

Es sei $f : V \rightarrow W$ eine lineare Abbildung. Ist $b \in W$, so heißt

$$f(x) = b$$

eine *lineare Gleichung*.

$$\mathcal{L}_{f,b} = \{x \in V \mid f(x) = b\}$$

heißt *Lösungsmenge der Gleichung*. Ist $b = 0$, so heißt die Gleichung *homogen*, und es gilt:

$$\mathcal{L}_{f,0} = \text{Kern } f \ni \{0\}.$$

Ist $b \neq 0$, so heißt die Gleichung *inhomogen*.

Satz 8.1.10. *Sei $f : V \rightarrow W$ eine lineare Abbildung. Dann folgt: Ist $\mathcal{L}_{f,b} \neq \emptyset$ und ist $f(x_0) = b$ eine spezielle Lösung (partikuläre Lösung), so gilt*

$$\mathcal{L}_{f,b} = x_0 + \text{Kern } f := \{x_0 + x \mid x \in \text{Kern } f\}.$$

Bemerkung. Man sagt: Die Menge der Lösungen einer inhomogenen linearen Gleichung besteht aus einer speziellen Lösung und dem Raum aller Lösungen der zugehörigen homogenen Gleichung. Insbesondere ist die Lösung nur dann eindeutig, falls $\text{Kern } f = \{0\}$.

Beweis. Jedes Element $z = x_0 + x$ mit $x \in \text{Kern } f$ ist enthalten in $\mathcal{L}_{f,b}$, denn

$$f(z) = f(x_0) + f(x) = f(x_0) = b.$$

Ist umgekehrt $z \in \mathcal{L}_{f,b}$, d.h ist $f(z) = b$, so folgt $f(z - x_0) = f(z) - f(x_0) = b - b = 0$. Somit ist $z - x_0 = x \in \text{Kern } f$. \square

Als Beispiele betrachten wir die folgenden wichtigen Typen von linearen Gleichungen.

Beispiele.

1. Lineare Differentialgleichungen

- (a) Wie in Beispiel 4 nach Korollar 8.1.7 sei $D : C^1(\mathbb{R}) \rightarrow C^0(\mathbb{R})$, $Df = f'$ die Ableitung. Zu $g \in C^0(\mathbb{R})$ betrachte die Gleichung $Df = g$. Dann ist $G(x) = \int_0^x g(s)ds$ eine spezielle Lösung dieser Gleichung und

$$\mathcal{L}_{D,g} = \{G + c \mid c \in \mathbb{R}\}.$$

- (b) (*Schwingungsgleichung*)

Sei $L : C^2(\mathbb{R}) \rightarrow C^0(\mathbb{R})$ der lineare Operator definiert durch

$$L(f) = f'' + a_1 f' + a_0 f,$$

mit $a_0, a_1 \in \mathbb{R}$. Die zu $g \in C^0(\mathbb{R})$ gehörige Gleichung

$$L(f) = f'' + a_1 f' + a_0 f = g$$

heißt Schwingungsgleichung. Wegen ihre grossen Bedeutung in der Physik (siehe z.B die Vorlesung zur theoretischen Physik) werden wir den linearen Differentialgleichungen ein eigenes Kapitel widmen.

2. Lineare skalare Gleichungssysteme

Die folgenden linearen Gleichungen treten zumindest in Spezialfällen schon in der Schule auf. Sei K ein Körper und $\{a_1, \dots, a_n\}$ seien n Vektoren in K^m . Dann ist $A : K^n \rightarrow K^m$ mit

$$A(x_1, \dots, x_n) := \sum_{j=1}^n x_j a_j$$

eine lineare Abbildung (der einfache Beweis dieser Aussage sei zur Übung überlassen). Also ist für jedes $b \in K^m$ die Gleichung $A(x) = b$ linear. Schreibt man

$$a_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} \quad \text{und} \quad b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix},$$

so ist diese Gleichung äquivalent zu

$$\begin{pmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

Dieses sind m skalare Gleichungen mit n Unbekannten. Das zugehörige quadratische Schema

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

nennt man auch Matrix. Jede solche Matrix definiert eine lineare Abbildung $A : K^n \rightarrow K^m$, durch

$$A(x) := \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n \end{pmatrix}.$$

Die Lösungsmenge solcher Gleichungssysteme lässt sich durch sogenannte elementare Umformungen berechnen. Das wichtigste Verfahren dazu ist der *Gauss-Algorithmus*. Wir werden später genauer darauf eingehen.

8.2 Erzeugendensysteme, Basissysteme und lineare Abbildungen

Satz 8.2.1. Sei $f : V \rightarrow W$ eine lineare Abbildung und E Erzeugendensystem von V . Dann ist $f(E)$ Erzeugendensystem von $\text{Bild } f$, d.h. $\text{Span}(f(E)) = \text{Bild } f$.

Beweis. Sei $y \in \text{Bild } f$, d.h. $y = f(x)$ mit $x \in V$. Da $x \in \text{Span}(E)$, gilt:

$$x = \sum_{e \in E} \alpha_e e,$$

mit $\alpha_e \in K$ und $\alpha_e \neq 0$ für höchstens endlich viele $e \in E$. Also folgt:

$$f(x) = \sum_{e \in E} \alpha_e f(e)$$

und somit ist $y \in \text{Span}(f(E))$. □

Satz 8.2.2. *Sei E ein Erzeugendensystem des Vektorraumes V und $f : V \rightarrow W$ eine lineare Abbildung. Dann ist f durch $f|_E$ schon eindeutig festgelegt.*

Beweis. Sei $g : V \rightarrow W$ eine weitere lineare Abbildung mit $g|_E = f|_E$. Ist $x \in \text{Span}(E) = V$, so ist $x = \sum_{e \in E} \alpha_e e$. Dann gilt:

$$f(x) = f\left(\sum_{e \in E} \alpha_e e\right) = \sum_{e \in E} \alpha_e f(e) = \sum_{e \in E} \alpha_e g(e) = g\left(\sum_{e \in E} \alpha_e e\right) = g(x).$$

Eine beliebige Abbildung von einer Basis in einen Vektorraum lässt sich linear fortsetzen. Genauer gilt:

Satz 8.2.3. *Sei B eine Basis eines K -Vektorraumes V und $F : B \rightarrow W$ eine beliebige Abbildung in einen K -Vektorraum W . Dann gibt es eine eindeutig bestimmte lineare Abbildung $f : V \rightarrow W$ mit $f|_B = F$.*

Beweis. Die Eindeutigkeit gilt schon nach Satz 8.2.2 für ein Erzeugendensystem. Existenz: Da $B \subset V$ eine Basis ist, gibt es für $x \in V$ genau eine Darstellung $x = \sum_{b \in B} \alpha_b b$.

Setze $f(x) = \sum_{b \in B} \alpha_b F(b)$. Diese Abbildung ist linear, denn ist $y = \sum_{b \in B} \beta_b b$, so gilt:

$$\begin{aligned} f(x+y) &= f\left(\sum_{b \in B} (\alpha_b + \beta_b) b\right) = \sum_{b \in B} (\alpha_b + \beta_b) F(b) \\ &= \sum_{b \in B} \alpha_b F(b) + \sum_{b \in B} \beta_b F(b) = f(x) + f(y). \end{aligned}$$

Ist $\alpha \in K$, so gilt:

$$f(\alpha x) = f\left(\sum_{b \in B} \alpha \alpha_b b\right) = \sum_{b \in B} \alpha \alpha_b F(b) = \alpha \left(\sum_{b \in B} \alpha_b F(b)\right) = \alpha f(x).$$

□

Wann ist das Bild einer Basis unter einer linearen Abbildung wieder eine Basis?

Satz 8.2.4. *Sei $f : V \rightarrow W$ linear und $B \subset V$ eine Basis. Dann gilt: f ist injektiv $\Leftrightarrow f(B)$ ist eine Basis von $\text{Bild} f$.*

Beweis. Wegen Satz 8.2.1 ist $f(B)$ immer ein Erzeugendensystem von $\text{Bild} f$ für alle linearen Abbildungen f .

“ \Rightarrow ” Es gilt für jede Linearkombination mit Elementen aus der Menge $f(B)$:

$$0 = \sum_{b \in B} \alpha_b f(b) = f\left(\sum_{b \in B} \alpha_b b\right).$$

Da f injektiv ist, ist wegen obiger Bemerkung der Kern von f trivial und somit ist $\sum_{b \in B} \alpha_b b = 0$.

Da B eine Basis ist, gilt $\alpha_b = 0$, für alle $b \in B$.

“ \Leftarrow ” Wegen Bemerkung 2 nach dem Korollar 8.1.7 genügt es zu zeigen: Kern $f = 0$. Ist $x \in \text{Kern } f$, so folgt:

$$x = \sum_{b \in B} \alpha_b b \quad \text{und}$$

$$0 = f(x) = f\left(\sum_{b \in B} \alpha_b b\right) = \sum_{b \in B} \alpha_b f(b).$$

Da $f(B)$ eine Basis ist, gilt: $\alpha_b = 0$, für alle $b \in B$, und somit $x = 0$. □

Korollar 8.2.5. Sei $f : V \rightarrow W$ eine lineare Abbildung und B eine Basis von V , so gilt: f ist genau dann ein Isomorphismus, falls $f(B)$ eine Basis von W ist.

Beweis. Sei zunächst $f : V \rightarrow W$ ein Isomorphismus und B eine Basis von V . Dann ist wegen Satz 8.2.4 $f(B)$ eine Basis von $\text{Bild } f = W$.

Sei nun B eine Basis von V und $f(B)$ eine Basis von W . Aus Satz 8.2.4 folgt die Injektivität von f . Da $W = \text{Span}(f(B)) = \text{Bild } f$, folgt auch die Surjektivität. □

Korollar 8.2.6. Es seien V, W K -Vektorräume und $f : V \rightarrow W$ ein Isomorphismus. Dann gilt:

$$\dim V = \dim W.$$

Ist insbesondere V endlich dimensional, so auch W .

Beweis. Sei V endlich dimensional und B eine Basis von V . Da f bijektiv ist, gilt $|B| = |f(B)|$. Da wegen Korollar 8.2.5 $f(B)$ eine Basis von W ist, gilt: $\dim V = |B| = \dim W$.

Ist V unendlich dimensional und B eine Basis, so hat B unendlich viele Elemente. Dies gilt dann auch für $f(B)$. □

Bemerkung. Dies zeigt insbesondere, dass die Dimension eine Isomorphie-Invariante ist. Haben zwei K -Vektorräume verschiedene Dimensionen, so sind sie nicht isomorph. Nun erhalten wir folgende wichtige Dimensionsformel für lineare Abbildungen.

Satz 8.2.7. (Dimensionsformel für lineare Abbildungen)

Sei $f : V \rightarrow W$ eine lineare Abbildung und U ein beliebiges Komplement von Kern f . Dann gilt:

$$\dim U = \dim \text{Bild } f.$$

Ist $\dim V < \infty$, so folgt daraus:

$$\dim \text{Kern } f + \dim \text{Bild } f = \dim V.$$

Beweis. Aus Satz 8.1.8 folgt, dass die Abbildung $f|_U : U \rightarrow \text{Bild } f$ ein Isomorphismus ist. Also folgt aus Korollar 8.2.6: $\dim \text{Bild } f = \dim U$.

Ist $\dim V < \infty$, so folgt aus $V = U \oplus \text{Kern } f$ mit Korollar 7.3.4

$$\dim V = \dim U + \dim \text{Kern } f.$$

Korollar 8.2.8. Sei $f : V \rightarrow W$ eine lineare Abbildung. Dann gilt:

1. Ist f injektiv, so folgt $\dim V \leq \dim W$.
2. f surjektiv, so folgt: $\dim V \geq \dim W$.
3. Sei $\dim V = \dim W < \infty$ und es sei f injektiv oder surjektiv. Dann ist f ein Isomorphismus.

Beweis.

zu 1. Sei f injektiv und B eine Basis von V . Dann ist wegen Satz 8.2.4 die Menge $f(B)$ eine Basis von $\text{Bild}(f)$ und es gilt: $|B| = \dim V = |f(B)| \leq \dim W$.

zu 2. Sei f surjektiv und B eine Basis von V . Dann ist wegen Satz 8.2.1 die Menge $f(B)$ ein Erzeugendensystem von $\text{Bild}(f)$ und es gilt: $|B| = \dim V \geq |f(B)| \geq \dim \text{Bild } f = \dim W$.

zu 3. Sei f injektiv und B eine Basis von V . Dann ist wegen Satz 8.2.4 die Menge $f(B)$ eine Basis von $\text{Bild}(f)$. Dann ist aber $f(B)$ sogar eine Basis von W . Wäre dies nicht der Fall, so könnten wir die in W linear unabhängige Menge $f(B)$ zu einer Basis von W ergänzen. Da $\dim V = |B| < \infty$, wäre dann $\dim V < \dim W$ im Widerspruch zur Annahme. Also ist $f(B)$ eine Basis von W und wegen $\text{Bild } f = \text{Span } f(B) = W$ ist f auch surjektiv.

Sei nun f surjektiv, so gilt: $\dim \text{Kern } f = \dim V - \dim \text{Bild } f = 0$. Also ist f auch injektiv.

□

Nun zeigen wir, dass endlich dimensionale K -Vektorräume bis auf Isomorphie durch ihre Dimension bestimmt sind. Dazu benötigen wir den Begriff der *geordneten (numerierten) Basis*. Sei V ein n -dimensionaler K -Vektorraum und B eine Basis. Dann heißt eine bijektive Abbildung $\{1, \dots, n\} \rightarrow B$ eine *geordnete bzw. numerierte Basis*. Man schreibt diese Abbildung als n -Tupel (b_1, \dots, b_n) und bezeichnet sie wieder mit B .

Dann gilt der folgende Satz:

Satz 8.2.9. Sei V ein n -dimensionaler K -Vektorraum, so ist V isomorph zu K^n . Genauer bestimmt jede geordnete Basis $B = (b_1, \dots, b_n)$ einen Isomorphismus $\phi_B : V \rightarrow K^n$, wobei ϕ_B eindeutig durch

$$\phi_B(b_i) = e_i = (0, \dots, 0, \overset{i\text{-te Stelle}}{1}, 0, \dots, 0)$$

festgelegt ist. Umgekehrt bestimmt jeder Isomorphismus $\phi : V \rightarrow K^n$ genau eine geordnete Basis $(\phi^{-1}(e_1), \dots, \phi^{-1}(e_n))$ von V .

Bemerkung. Die Menge der Isomorphismen $\phi : V \rightarrow K^n$ steht also in eineindeutiger Beziehung zu den geordneten Basen.

Beweis. Sei $B = (b_1, \dots, b_n)$ eine (geordnete) Basis von V . Dann besitzt $\phi_B : B \rightarrow K^n$ mit $\phi_B(b_i) = e_i$ wegen Satz 8.2.3 genau eine lineare Erweiterung $\phi_B : V \rightarrow K^n$. Da $\phi_B(B)$ eine Basis von K^n ist, ist $\phi_B : V \rightarrow K^n$ wegen Korollar 8.2.5 ein Isomorphismus.

Ist umgekehrt $\phi : V \rightarrow K^n$ ein Isomorphismus, so ist $\phi^{-1} : K^n \rightarrow V$ ein Isomorphismus und wegen Korollar 8.2.5 $(\phi^{-1}(e_1), \dots, \phi^{-1}(e_n))$ eine geordnete Basis von V . □

Korollar 8.2.10. Sei V ein endlich dimensionaler K -Vektorraum. Ist W ein K -Vektorraum, so gilt:

$$V \cong W \Leftrightarrow \dim V = \dim W .$$

Beweis. “ \Rightarrow ” Diese Richtung folgt aus Korollar 8.2.6.

“ \Leftarrow ” Ist $\dim V = \dim W = n$, so folgt aus Satz 8.2.9 : $V \cong K^n$ und $W \cong K^n$ und somit $V \cong W$. Sind insbesondere $B = (b_1, \dots, b_n)$ und $C = (c_1, \dots, c_n)$ geordnete Basen von V bzw. W , so ist $\phi_C^{-1} \circ \phi_B : V \rightarrow W$ ein Isomorphismus.

$$\begin{array}{ccc} V & \xrightarrow{\phi_C^{-1} \circ \phi_B} & W \\ \phi_B \searrow & & \swarrow \phi_C \\ & K^n & \end{array}$$

□

Definition 8.2.11. Ist $B = (b_1, \dots, b_n)$ eine geordnete Basis des K -Vektorraumes V , so heißt der Isomorphismus $\phi_B : V \rightarrow K^n$ auch ein *Koordinatensystem* von V . Ist $x \in V$, so heißen

$$(x_1, \dots, x_n) = \phi_B(x)$$

auch die *Koordinaten* von x bezüglich B . Sind $B = (b_1, \dots, b_n)$ und $C = (c_1, \dots, c_n)$ zwei Basen, so heißt der Isomorphismus $T_{B,C} := \phi_C \circ \phi_B^{-1} : K^n \rightarrow K^n$ auch *Koordinatentransformation*.

$$\begin{array}{ccc} & V & \\ \phi_B \swarrow & & \searrow \phi_C \\ K^n & \xrightarrow{T_{B,C}} & K^n \end{array}$$

8.3 Vektorräume linearer Abbildungen und Dualräume

Definition 8.3.1. Es seien V, W Vektorräume über dem Körper K . Dann bezeichnen wir mit

$$L(V, W) := \{f : V \rightarrow W \mid f \text{ linear}\}$$

die Menge der linearen Abbildungen. Ist $V = W$, so nennen wir die Elemente in $L(V, V)$ auch *Endomorphismen*. Wir schreiben dann auch $\text{End}(V)$ statt $L(V, V)$.

Ist M eine beliebige Menge und W ein K -Vektorraum, so haben wir gesehen:

$$W^M = \{f \mid f : M \rightarrow W\}$$

ist wieder ein Vektorraum über K , mit $(f + g)(x) = f(x) + g(x)$ und $(\alpha f)(x) = \alpha f(x)$ für $\alpha \in K$ und $x \in W$. Nun ist die Summe zweier linearer Abbildungen und das skalare Vielfache einer linearen Abbildung wieder linear. Daher gilt:

Satz 8.3.2. *Es seien V, W Vektorräume über dem Körper K . Dann ist $L(V, W)$ Unterraum von W^V . Sind darüberhinaus die Vektorräume endlich dimensional, so gilt:*

$$\dim L(V, W) = \dim V \cdot \dim W.$$

Beweis. Zum Beweis der zweiten Aussage zeige man: Sind (v_1, \dots, v_n) bzw. (w_1, \dots, w_m) Basen von V bzw. W , so bilden die linearen Abbildungen $f_{ij} : V \rightarrow W$, $i \in \{1, \dots, n\}, j \in \{1, \dots, m\}$, mit

$$f_{ij}(v_k) = \begin{cases} 0 & k \neq i \\ w_j & k = i \end{cases}$$

eine Basis von $L(V, W)$. □

Von besonderer Bedeutung ist der Spezialfall $L(V, K)$.

Definition 8.3.3. Sei V ein K -Vektorraum, so heißt der Vektorraum $L(V, K)$ auch *Dualraum* von V . Abkürzend schreiben wir V^* statt $L(V, K)$. Seine Elemente heißen auch *Linearformen*.

Satz 8.3.4. *Sei (v_1, \dots, v_n) eine Basis von V . Dann definieren die Linearformen (v_1^*, \dots, v_n^*) mit*

$$v_i^*(v_j) = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$

eine Basis von V^* . Diese Basis heißt auch die zu (v_1, \dots, v_n) duale Basis.

Bemerkung. Der Satz zeigt insbesondere, dass die Vektorräume V und V^* isomorph sind, falls V endlich dimensional ist.

Beweis. Die duale Basis entspricht genau der Basis im Beweis von Satz 8.3.2 im Spezialfall $L(V, K)$, wobei in K als Basis die 1 gewählt wird. Ohne von diesem Satz Gebrauch zu machen sieht man jedoch ganz leicht, dass (v_1^*, \dots, v_n^*) eine Basis von V^* darstellt. Denn ist

$$\sum_{i=1}^n x_i v_i^* = 0,$$

so folgt durch Anwendung dieser Gleichung auf v_j

$$0 = \sum_{i=1}^n x_i v_i^*(v_j) = x_j.$$

Also sind die Linearformen (v_1^*, \dots, v_n^*) linear unabhängig. Sie erzeugen aber auch V^* , denn ist $\varphi \in V^*$, so folgt

$$\varphi = \sum_{i=1}^n \varphi(v_i) v_i^*,$$

da beide Seiten übereinstimmen, wenn man sie auf die Elemente der Basis (v_1, \dots, v_n) anwendet. □

Ist nun $f : V \rightarrow W$ eine lineare Abbildung, so kann man ihr in natürlicher Weise eine lineare Abbildung $f^t : W^* \rightarrow V^*$ zuordnen. Sie heißt *transponierte Abbildung*. Genauer definiert man:

Definition 8.3.5. Sei $f : V \rightarrow W$ linear, so heißt $f^t : W^* \rightarrow V^*$ mit $f^t(\varphi) = \varphi \circ f$ die zu f *transponierte Abbildung*. Die Definition lässt sich auch durch folgendes Diagramm leicht illustrieren:

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ & \searrow \varphi \circ f & \downarrow \varphi \\ & & K \end{array}$$

Bemerkungen.

- $\varphi \circ f : V \rightarrow K$ ist in V^* , da nach Satz 8.1.2 Verknüpfungen linearer Abbildungen linear sind.
- Die Abbildung $\langle \cdot, \cdot \rangle : V^* \times V \rightarrow K$ mit

$$\langle \varphi, x \rangle = \varphi(x)$$

nennt man auch *kanonische Paarung* von V und V^* . Ist $f : V \rightarrow W$ eine lineare Abbildung, so folgt für die transponierte Abbildung $f^t : W^* \rightarrow V^*$:

$$\langle f^t(\varphi), x \rangle_1 = \langle \varphi, f(x) \rangle_2,$$

wobei $\langle \cdot, \cdot \rangle_1 : V^* \times V \rightarrow K$ bzw. $\langle \cdot, \cdot \rangle_2 : W^* \times W \rightarrow K$ die kanonischen Paarungen von V und V^* bzw. W und W^* bezeichnen.

Nun zeigen wir, dass f^t auch eine lineare Abbildung ist.

Satz 8.3.6. Sei $f : V \rightarrow W$ linear, so ist auch $f^t : W^* \rightarrow V^*$ eine lineare Abbildung.

Beweis. Seien $\varphi_1, \varphi_2 \in W^*$, so gilt:

$$f^t(\varphi_1 + \varphi_2) = (\varphi_1 + \varphi_2) \circ f = \varphi_1 \circ f + \varphi_2 \circ f = f^t(\varphi_1) + f^t(\varphi_2).$$

Genauso folgt für alle $\alpha \in K$ und $\varphi \in W^*$:

$$f^t(\alpha\varphi) = (\alpha\varphi) \circ f = \alpha f^t(\varphi).$$

□

Kapitel 9

Lineare Differentialgleichungen I

Als erste wichtige Anwendung der linearen Algebra wollen wir uns nun mit linearen Differentialgleichungen befassen.

9.1 Lineare Differentialgleichungen n -ter Ordnung

Sei $I \subset \mathbb{R}$ ein Intervall. Im Folgenden bezeichnen wir für $n \in \mathbb{N} \cup \{0\}$ mit

$$C^n(I, \mathbb{C}) = \{z : I \rightarrow \mathbb{C} \mid z \text{ ist } n\text{-mal stetig differenzierbar}\}$$

die Menge der auf I n -mal stetig differenzierbaren komplexwertigen Funktionen. Dabei heißt die Abbildung $z : I \rightarrow \mathbb{C}$ mit $z(t) = x(t) + iy(t)$ *differenzierbar*, wenn ihre Real- und Imaginärteil $x : I \rightarrow \mathbb{R}$ und $y : I \rightarrow \mathbb{R}$ differenzierbar sind. Ist dies der Fall, so heißt $z'(t) = x'(t) + iy'(t)$ die *Ableitung* von z . Sind die Real- und Imaginärteile $x : I \rightarrow \mathbb{R}$ und $y : I \rightarrow \mathbb{R}$ n -mal stetig differenzierbar, so heißt auch $z : I \rightarrow \mathbb{C}$ *n -mal stetig differenzierbar*. Mit $C^\infty(I, \mathbb{C})$ werden die unendlich oft differenzierbaren Funktionen und mit $C^0(I, \mathbb{C})$ die stetigen Funktionen bezeichnet.

Diese Mengen stellen Vektorräume über \mathbb{C} dar. Wie schon erwähnt wurde, definiert die Ableitung

$$D : C^1(I, \mathbb{C}) \rightarrow C^0(I, \mathbb{C})$$

mit $D(z) = z'$ eine lineare Abbildung. In diesem Fall nennt man D auch *linearer Operator* oder auch *linearer Differentialoperator 1. Ordnung*. Allgemein heißt $L : C^n(I, \mathbb{C}) \rightarrow C^0(I, \mathbb{C})$ ein *linearer Differentialoperator n -ter Ordnung*, falls er von der Form

$$L(z) = z^{(n)} + a_{n-1}z^{(n-1)} + \dots + a_1z' + a_0z$$

ist, wobei die Koeffizienten $a_i \in C^0(I, \mathbb{C})$ aus stetigen Funktionen bestehen. Da $L(z_1 + z_2) = L(z_1) + L(z_2)$ und $L(\alpha z_1) = \alpha L(z_1)$ für $z_1, z_2 \in C^n(I, \mathbb{C})$ und $\alpha \in \mathbb{C}$ gilt, ist L eine lineare Abbildung. Ist $b \in C^0(I, \mathbb{C})$, so nennen wir die Gleichung

$$L(z) = z^{(n)} + a_{n-1}z^{(n-1)} + \dots + a_0z = b$$

lineare Differentialgleichung n -ter Ordnung. Aus der Linearität von L folgt wegen Satz 8.1.10, dass die Menge der Lösungen dieser Differentialgleichung von der Form

$$\mathcal{L}_{L,b} = \text{Kern } L + z_0$$

ist. Dabei ist z_0 eine spezielle Lösung (partikuläre Lösung) der Differentialgleichung. Mittels des Operators D lässt sich L auch in der Form

$$L = D^n + a_{n-1}D^{n-1} + \dots + a_1z' + a_0\text{id}$$

schreiben. Dabei ist $D^k = \underbrace{D \circ \dots \circ D}_{k\text{-mal}}$ die k -fache Komposition von D mit sich selbst.

Wir wollen nun eine Abschätzung der Dimension des Kerns von L angeben. Dafür benötigen wir den folgenden Eindeutigkeitsatz für die Lösung der Gleichung $L(z) = b$ bei vorgegebenen Anfangswerten $z(t_0), z'(t_0), \dots, z^{(n-1)}(t_0)$ an der Stelle t_0 .

Satz 9.1.1. (Eindeutigkeitsatz)

Es seien $z_1, z_2 : I \rightarrow \mathbb{C}$ zwei Lösungen der Gleichung $L(z) = b$ mit den Anfangswerten

$$z_1(t_0) = z_2(t_0), \dots, z_1^{(n-1)}(t_0) = z_2^{(n-1)}(t_0)$$

für ein $t_0 \in I$. Dann gilt:

$$z_1(t) = z_2(t)$$

für alle $t \in I$.

Der Beweis ergibt sich aus folgendem einfachen Lemma.

Lemma 9.1.2. Es sei $f : I \rightarrow \mathbb{C}$ eine auf einem Intervall $I \subset \mathbb{R}$ differenzierbare Funktion. Existiert eine Konstante $c > 0$ mit

$$|f'(t)| \leq c|f(t)|$$

für alle $t \in I$ und gilt $f(t_0) = 0$, so folgt $f(t) = 0$ für alle $t \in I$.

Beweis. Sei zunächst f eine reellwertige Funktion mit $f(t) \geq 0$. Man betrachte die Funktion $g(t) = f(t) \cdot e^{-ct}$. Dann ist $g'(t) = f'(t)e^{-ct} - cf(t)e^{-ct} \leq 0$. Also ist $g(t)$ monoton fallend. Da $g(t_0) = 0$ und $g(t) \geq 0$, folgt $g(t) = 0$ für $t \geq t_0$ und somit auch $f(t) = 0$ für $t \geq t_0$. Betrachtet man $g(t) = f(t)e^{ct}$, so folgt $g'(t) = f'(t)e^{ct} + cf(t)e^{ct} \geq 0$. Also ist $g(t)$ monoton steigend und wir erhalten auch $f(t) = 0$ für $t \leq t_0$.

Sei nun f eine beliebige komplexwertige Funktion mit $|f'(t)| \leq c|f(t)|$. Betrachtet man nun $h(t) = f(t)\bar{f}(t)$, so gilt

$$\begin{aligned} |h'(t)| &= |f'(t)\bar{f}(t) + f(t)\bar{f}'(t)| \leq |f'(t)\bar{f}(t)| + |f(t)\bar{f}'(t)| \\ &= 2|f'(t)\bar{f}(t)| \leq 2c|f(t)\bar{f}(t)| = 2c|h(t)|. \end{aligned}$$

Ist nun $f(t_0) = 0$, so folgt aus $h(t_0) = 0$ wegen des ersten Teils: $h(t) = 0$ und somit $f(t) = 0$ für alle $t \in I$.

Beweis des Eindeutigkeitsatzes:

Sei nun $J \subset I$ ein kompaktes Intervall, welches $t_0 \in I$ enthält. Wegen der Stetigkeit der Funktionen a_i existiert eine Konstante A mit $|a_i(t)| \leq A$ für alle $t \in J$ und $i \in \{0, \dots, n-1\}$. Sei $z(t) = z_1(t) - z_2(t)$ und

$$f(t) = z(t) \cdot \bar{z}(t) + z'(t)\bar{z}'(t) + \dots + z^{(n-1)}(t)\bar{z}^{(n-1)}(t).$$

Insbesondere ist dann $z^{(k)}(t)\bar{z}^{(k)}(t) \leq f(t)$ für $k \leq n-1$ und somit auch $|z^{(k)}(t)| \leq \sqrt{f(t)}$.
Da

$$f'(t) = z'(t)\bar{z}(t) + z(t)\bar{z}'(t) + \dots + z^{(n)}(t) \cdot \bar{z}^{(n-1)}(t) + z^{(n-1)}(t)\bar{z}^{(n)}(t),$$

folgt:

$$|z^{(n)}(t)| = \left| \sum_{k=0}^{n-1} a_k(t) z^{(k)}(t) \right| \leq \sum_{k=0}^{n-1} |a_k(t)| |z^{(k)}(t)| \leq \sum_{k=0}^{n-1} A\sqrt{f} \leq nA\sqrt{f}.$$

Insgesamt erhalten wir

$$\begin{aligned} |f'(t)| &\leq \sum_{k=0}^{n-1} (|z^{(k+1)}(t)\bar{z}^{(k)}(t)| + |z^{(k)}(t)\bar{z}^{(k+1)}(t)|) \\ &= \sum_{k=0}^{n-2} \underbrace{(|z^{(k+1)}(t)||\bar{z}^{(k)}(t)| + |z^{(k)}(t)||\bar{z}^{(k+1)}(t)|)}_{\leq \sqrt{f(t)} \cdot \sqrt{f(t)} + \sqrt{f(t)} \cdot \sqrt{f(t)}} + \underbrace{(|z^{(n)}(t)||\bar{z}^{(n-1)}(t)| + |z^{(n-1)}(t)||\bar{z}^{(n)}(t)|)}_{\leq nA\sqrt{f(t)} \cdot \sqrt{f(t)} + \sqrt{f(t)} \cdot nA\sqrt{f(t)}} \\ &\leq 2(n-1)f(t) + 2Anf(t) = (2n(1+A) - 2)f(t) \\ &= cf(t). \end{aligned}$$

Ist nun $z_1^{(k)}(t_0) = z_2^{(k)}(t_0)$ für $k \in \{0, \dots, n-1\}$, so ist $f(t_0) = 0$ und somit nach Lemma 9.1.2
 $f(t) = 0$ für alle $t \in J$. □

□

Satz 9.1.3. Sei $L : C^n(I, \mathbb{C}) \rightarrow C^0(I, \mathbb{C})$ ein linearer Differentialoperator n -ter Ordnung. Ist $t_0 \in I$, so ist $\phi : \text{Kern } L \rightarrow \mathbb{C}^n$ mit

$$\phi(z) = (z(t_0), z'(t_0), \dots, z^{(n-1)}(t_0))$$

eine lineare injektive Abbildung. Insbesondere ist $\dim \text{Kern } L \leq n$.

Beweis. Die Abbildung ϕ ist offensichtlich linear. Wegen des Eindeutigkeitsatzes 9.1.1 ist ϕ auch injektiv. Wegen Korollar 8.2.8 folgt $\dim \text{Kern } L \leq \dim \mathbb{C}^n = n$. □

Bemerkung.

- (a) Die Abbildung ϕ ist auch surjektiv, d.h. zu jedem n -Tupel (z_0, \dots, z_{n-1}) komplexer Zahlen existiert eine Funktion $z \in \text{Kern } L$ mit den Anfangswerten

$$z(t_0) = z_0, z'(t_0) = z_1, \dots, z^{(n-1)}(t_0) = z_{n-1}.$$

Wir werden diese Aussage gleich im Falle konstanter Koeffizienten beweisen. Der allgemeine Fall benötigt etwas mehr Theorie und wird später behandelt.

- (b) Der Satz besagt also: Es gibt n Lösungen $z_1, \dots, z_n : I \rightarrow \mathbb{C}$ der Gleichung $Lz = 0$, so dass jede beliebige Lösung $z : I \rightarrow \mathbb{C}$ sich als Linearkombination von z_1, \dots, z_n schreiben lässt. Die Lösungen heißen auch *Fundamentalsystem* der Gleichung $Lz = 0$.

9.2 Homogene lineare Differentialgleichungen mit komplexen Koeffizienten

Nun wollen wir im Falle konstanter Koeffizienten $a_0, a_1, \dots, a_{n-1} \in \mathbb{C}$ ein Fundamentalsystem für die homogene Differentialgleichung

$$Lz = z^{(n)} + a_{n-1}z^{(n-1)} + \dots + a_0z = 0$$

mit $z \in C^n(\mathbb{R}, \mathbb{C})$ explizit bestimmen. Dazu ist es zweckmäßig, den Differentialoperator folgendermaßen zu interpretieren. Man betrachte das komplexe Polynom

$$p(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_0.$$

Substituieren wir in diesem Polynom die Variable λ durch den Operator D , so erhalten wir:

$$p(D) = D^n + a_{n-1}D^{n-1} + \dots + a_0\text{id} = L.$$

Hierbei wurde $a_0 = a_0 \cdot 1$ durch $a_0\text{id}$ ersetzt, wobei id die Identität ist. Sind nun p und q zwei Polynome (der Koeffizient zur höchsten Potenz muss dabei nicht notwendigerweise gleich 1 sein), so erhalten wir:

$$p \cdot q(D) = p(D) \circ q(D) = q(D) \circ p(D). \quad (*)$$

Sei nun

$$p(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_0$$

ein Polynom mit komplexen Koeffizienten, so lässt es sich wegen Satz 3.3.10 in Linearfaktoren zerlegen, d.h. es gibt $m \leq n$ verschiedene Nullstellen $\lambda_1, \dots, \lambda_m$, so dass

$$p(z) = (\lambda - \lambda_1)^{k_1} \dots (\lambda - \lambda_m)^{k_m},$$

wobei $k_i \in \mathbb{N}$ und $k_1 + \dots + k_m = n$ gilt. Der Faktorisierung des Polynoms entspricht wegen (*) eine Faktorisierung des Differentialoperators $L = p(D) = \sum_{j=0}^n a_j D^j$, d.h.

$$p(D) = (D - \lambda_1 \text{id})^{k_1} \circ \dots \circ (D - \lambda_m \text{id})^{k_m}.$$

Da die Operatoren $(D - \lambda_j \text{id})$ paarweise kommutieren, kann die Reihenfolge der Kompositionen beliebig vertauscht werden. Daraus ergibt sich sofort

$$\text{Kern}(D - \lambda_1 \text{id})^{k_1} + \dots + \text{Kern}(D - \lambda_m \text{id})^{k_m} \subset \text{Kern } p(D).$$

Nun zeigen wir, dass die Summe direkt ist. Als Vorbereitung dazu dient das folgende Lemma:

Lemma 9.2.1. 1. Sei $f : I \rightarrow \mathbb{C}$ eine k -mal differenzierbare Funktion und $\lambda \in \mathbb{C}$. Dann gilt:

$$(D - \lambda \text{id})^k (f(t)e^{\lambda t}) = f^{(k)}(t)e^{\lambda t}.$$

2. Es seien die Realteil und Imaginärteil von $h : I \rightarrow \mathbb{C}$ Polynome und $\lambda \neq \mu \in \mathbb{C}$. Dann gilt für $k \in \mathbb{N}$:

$$(D - \lambda \text{id})^k h(t)e^{\mu t} = \tilde{h}(t)e^{\mu t},$$

wobei \tilde{h} ein Polynom ist. Dabei ist \tilde{h} von null verschieden, falls h von null verschieden ist.

3. Seien h_1, \dots, h_n Polynome und $\lambda_1, \dots, \lambda_n$ paarweise verschiedene komplexe Zahlen. Ist

$$h_1 e^{\lambda_1 t} + \dots + h_n e^{\lambda_n t} = 0,$$

so folgt $h_1 = h_2 = \dots = h_n = 0$.

Beweis.

zu 1. Die Anwendung des Operators $D - \lambda \text{id}$ auf $f(t)e^{\lambda t}$ impliziert

$$(D - \lambda \text{id})f(t)e^{\lambda t} = f'(t)e^{\lambda t} + \lambda f(t)e^{\lambda t} - \lambda f(t)e^{\lambda t} = f'(t)e^{\lambda t}.$$

Nach k -facher Anwendung von $(D - \lambda \text{id})$ erhalten wir somit $f^{(k)}(t)e^{\lambda t}$.

zu 2. Die Anwendung des Operators $D - \lambda \text{id}$ auf $h(t)e^{\mu t}$ impliziert

$$(D - \lambda \text{id})h(t)e^{\mu t} = h'(t)e^{\mu t} + \mu h(t)e^{\mu t} - \lambda h(t)e^{\mu t} = ((\mu - \lambda)h(t) + h'(t))e^{\mu t} = h_1(t)e^{\mu t}.$$

Dabei ist $h_1(t) = (\mu - \lambda)h(t) + h'(t)$ ein von null verschiedenes Polynom, falls h von null verschieden ist, denn der Grad des Polynoms h' (d.h. der Wert der in h höchsten auftretenden Potenz) ist echt kleiner als der Grad von h . Wäre $(\mu - \lambda)h(t) + h'(t)$ das Nullpolynom, so würde wegen $(\mu - \lambda) \neq 0$ der Grad von h mit dem Grad von h' übereinstimmen. Wenden wir nun den Operator $(D - \lambda \text{id})^k$ auf $h(t)e^{\mu t}$ an, so erhalten wir

$$(D - \lambda \text{id})^k(h(t)e^{\mu t}) = h_k(t)e^{\mu t}.$$

Dabei ist $\tilde{h} := h_k$ ein von null verschiedenes Polynom, falls h von null verschieden ist.

zu 3. Der Beweis dieser Aussage folgt durch Induktion über n . Für $n = 1$ ist nichts zu zeigen. Wir nehmen also an, dass die Aussage schon für $n \in \mathbb{N}$ bewiesen ist. Sei also

$$h_1 e^{\lambda_1 t} + \dots + h_n e^{\lambda_n t} + h_{n+1} e^{\lambda_{n+1} t} = 0, \quad (+)$$

wobei h_1, \dots, h_{n+1} Polynome und $\lambda_1, \dots, \lambda_{n+1}$ paarweise verschiedene komplexe Zahlen sind. Ist h_{n+1} ein Polynom vom Grade k , so folgt aus 1.

$$(D - \lambda_{n+1} \text{id})^k(h_{n+1} e^{\lambda_{n+1} t}) = 0$$

und aus 2.

$$(D - \lambda_{n+1} \text{id})^k(h_j e^{\lambda_j t}) = \tilde{h}_j(t) e^{\lambda_j t}$$

für $j \leq n$. Durch Anwendung des Operators $(D - \lambda_{n+1} \text{id})^k$ auf (+) erhalten wir daher:

$$\tilde{h}_1 e^{\lambda_1 t} + \dots + \tilde{h}_n e^{\lambda_n t} = 0.$$

Nach Induktionsannahme ist $\tilde{h}_1 = \tilde{h}_2 = \dots = \tilde{h}_n = 0$ und somit wegen 2. auch $h_1 = h_2 = \dots = h_n = 0$.

□

Satz 9.2.2. Sei $L : C^n(\mathbb{R}, \mathbb{C}) \rightarrow C^0(\mathbb{R}, \mathbb{C})$ der Differentialoperator mit

$$L = D^n + a_{n-1}D^{n-1} + \dots + a_0 \text{id},$$

wobei $a_i \in \mathbb{C}$. Dann ist

$$\text{Kern}(D - \lambda_1 \text{id})^{k_1} \oplus \dots \oplus \text{Kern}(D - \lambda_m \text{id})^{k_m} = \text{Kern } p(D),$$

wobei für $i \in \{1, \dots, m\}$

$$\text{Kern}(D - \lambda_i \text{id})^{k_i} = \text{Span}\{e^{\lambda_i t}, t e^{\lambda_i t}, t^2 e^{\lambda_i t}, \dots, t^{k_i-1} e^{\lambda_i t}\}$$

gilt.

Beweis. Sei $z_i \in \text{Kern}(D - \lambda_i \text{id})^{k_i}$, so definiere $f(t) = z(t)e^{-\lambda_i t}$. Dann folgt mit Hilfe des ersten Teils von Lemma 9.2.1 :

$$0 = (D - \lambda_i \text{id})^{k_i} z(t) = f^{(k_i)}(t) e^{\lambda_i t}.$$

Also ist $f^{(k_i)}(t) = 0$ und somit sind die Real- und Imaginärteil von f Polynome. Die Summe der Vektorräume $\text{Kern}(D - \lambda_i \text{id})^{k_i}$ ist nach Definition 7.3.5 direkt, falls für jede Linearkombination

$$0 = z_1 + \dots + z_m$$

mit $z_i \in \text{Kern}(D - \lambda_i \text{id})^{k_i}$ folgt: $z_1 = z_2 = \dots = z_m = 0$. Nun ist jedes $z_i \in \text{Kern}(D - \lambda_i \text{id})^{k_i}$ von der Form $z_i(t) = f_i(t)e^{\lambda_i t}$, wobei f_i ein Polynom darstellt. Daher gilt

$$0 = f_1(t)e^{\lambda_1 t} + \dots + f_m(t)e^{\lambda_m t}.$$

Damit sind wegen des dritten Teils von Lemma 9.2.1 die Polynome f_i und somit auch die Funktionen z_i identisch null. Also ist die Summe direkt.

Wegen Teil 1 des Lemmas 9.2.1 gilt auch:

$$\text{Span}\{e^{\lambda_i t}, t e^{\lambda_i t}, t^2 e^{\lambda_i t}, \dots, t^{k_i-1} e^{\lambda_i t}\} \subset \text{Kern}(D - \lambda_i \text{id})^{k_i}.$$

Die Elemente sind linear unabhängig, denn ist

$$c_1 e^{\lambda_i t} + c_2 t e^{\lambda_i t} + \dots + c_{k_i} t^{k_i-1} e^{\lambda_i t} = 0,$$

so erhalten wir durch Auswertung der linken Seite an der Stelle $t = 0$, dass der Koeffizient c_1 gleich null ist. Teilen wir diese Gleichung durch t für $t \neq 0$, so folgt

$$c_2 e^{\lambda_i t} + \dots + c_{k_i} t^{k_i-2} e^{\lambda_i t} = 0$$

und somit ist auch $c_2 = 0$. Mehrmaliges Wiederholen dieses Verfahrens zeigt dann, dass alle Koeffizienten gleich null sind. Insbesondere haben wir gezeigt:

$$\dim \text{Kern}(D - \lambda_i \text{id})^{k_i} \geq k_i.$$

Damit folgt mit Satz 9.1.3

$$n = k_1 + \dots + k_m \leq \dim \text{Kern}(D - \lambda_1 \text{id})^{k_1} + \dots + \dim \text{Kern}(D - \lambda_m \text{id})^{k_m} \leq \dim \text{Kern } p(D) \leq n.$$

Also ist

$$\dim \text{Kern}(D - \lambda_i \text{id})^{k_i} = k_i$$

und somit ist

$$\text{Kern}(D - \lambda_1 \text{id})^{k_1} \oplus \dots \oplus \text{Kern}(D - \lambda_m \text{id})^{k_m} = \text{Kern } p(D).$$

□

9.3 Homogene lineare Differentialgleichungen mit reellen Koeffizienten

Nun wollen wir uns den linearen Differentialgleichungen mit reellen Koeffizienten zuwenden. Sei also nun

$$Lz = z^{(n)} + a_{n-1}z^{(n-1)} + \dots + a_0z = 0 \quad (*)$$

ein Differentialoperator n -ter Ordnung mit reellen Koeffizienten $a_i \in \mathbb{R}$. Bei solchen Gleichungen ist man meistens nur an den reellen Lösungen interessiert. Die Idee besteht nun darin, die reellen Lösungen aus den komplexen Lösungen zu gewinnen. Ist $z : I \rightarrow \mathbb{C}$ eine Lösung der Gleichung (*), so sind sowohl der Realteil $\operatorname{Re}(z) = x$ als auch der Imaginärteil $\operatorname{Im}(z) = y$ Lösungen dieser Gleichung. Denn ist

$$L(z) = L(x + iy) = L(x) + iL(y) = 0,$$

so folgt auch $L(x) = 0$ und $L(y) = 0$. Man betrachte nun das zu L gehörige Polynom

$$p(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_0.$$

Ist λ eine reelle Nullstelle der Vielfachheit k , so erhalten wir mit Satz 9.2.2 die k reellen Lösungen

$$e^{\lambda t}, te^{\lambda t}, t^2e^{\lambda t}, \dots, t^{k-1}e^{\lambda t}.$$

Man bezeichne mit

$$V_\lambda := \operatorname{Span}_{\mathbb{R}}\{e^{\lambda t}, te^{\lambda t}, t^2e^{\lambda t}, \dots, t^{k-1}e^{\lambda t}\} := \{u_1e^{\lambda t} + u_2te^{\lambda t} + \dots + u_kt^{k-1}e^{\lambda t} \mid u_j \in \mathbb{R}\}$$

den reellen Vektorraum, der durch diese Elemente erzeugt wird. Das Erzeugendensystem ist linear unabhängig, denn ihre Elemente sind auch über \mathbb{C} linear unabhängig. Somit definiert V_λ einen reellen Vektorraum der Dimension k . Offensichtlich ist V_λ Teilmenge des komplexen Vektorraumes

$$\operatorname{Kern}(D - \lambda \operatorname{id})^k = \{c_1e^{\lambda t} + c_2te^{\lambda t} + \dots + c_kt^{k-1}e^{\lambda t} \mid c_j \in \mathbb{C}\}.$$

Ist $\mu = \alpha + i\beta$ eine nicht reelle Nullstelle der Vielfachheit m , mit $\beta > 0$, so ist auch $\bar{\mu} = \alpha - i\beta$ eine Nullstelle der Vielfachheit m . In diesem Falle gilt wegen Satz 9.2.2

$$\operatorname{Kern}(D - \mu \operatorname{id})^m = \{z_0, \dots, z_{m-1}\} \quad \text{und} \quad \operatorname{Kern}(D - \bar{\mu} \operatorname{id})^m = \{\bar{z}_0, \dots, \bar{z}_{m-1}\}$$

mit

$$z_j(t) = t^j e^{\mu t} = t^j e^{(\alpha+i\beta)t} = t^j e^{\alpha t} (\cos \beta t + i \sin \beta t)$$

und

$$\bar{z}_j(t) = t^j e^{\bar{\mu} t} = t^j e^{(\alpha-i\beta)t} = t^j e^{\alpha t} (\cos \beta t - i \sin \beta t).$$

Dann gilt

$$\begin{aligned} \operatorname{Re} z_j &= t^j e^{\alpha t} \cos \beta t = \frac{1}{2}(z_j + \bar{z}_j) \in \operatorname{Kern}(D - \mu \operatorname{id})^m \oplus \operatorname{Kern}(D - \bar{\mu} \operatorname{id})^m, \\ \operatorname{Im} z_j &= t^j e^{\alpha t} \sin \beta t = \frac{1}{2i}(z_j - \bar{z}_j) \in \operatorname{Kern}(D - \mu \operatorname{id})^m \oplus \operatorname{Kern}(D - \bar{\mu} \operatorname{id})^m. \end{aligned}$$

Definiere durch

$$V_{(\mu, \bar{\mu})} := \operatorname{Span}_{\mathbb{R}}\{\operatorname{Re} z_0, \operatorname{Im} z_0, \dots, \operatorname{Re} z_{m-1}, \operatorname{Im} z_{m-1}\}$$

den reellen Vektorraum. Dieser Vektorraum ist in $\operatorname{Kern}(D - \mu \operatorname{id})^m \oplus \operatorname{Kern}(D - \bar{\mu} \operatorname{id})^m$ enthalten. Wir zeigen nun, dass die Dimension des Vektorraumes $V_{(\mu, \bar{\mu})}$ $2m$ beträgt.

Lemma 9.3.1. Sei also $\mu = \alpha + i\beta$ mit $\beta > 0$ eine Nullstelle des Polynoms

$$p(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_0$$

der Vielfachheit m . Dann sind die $2m$ Funktionen

$$\{\operatorname{Re} z_0, \operatorname{Im} z_0, \dots, \operatorname{Re} z_{m-1}, \operatorname{Im} z_{m-1}\},$$

mit $\operatorname{Re} z_j(t) = t^j e^\alpha \cos \beta t$ und $\operatorname{Im} z_j(t) = t^j e^\alpha \sin \beta t$ reelle Lösungen von

$$Lz = z^{(n)} + a_{n-1}z^{(n-1)} + \dots + a_0z = 0$$

und linear unabhängig im reellen Vektorraum $C^\infty(\mathbb{R}, \mathbb{R})$.

Beweis. Zu zeigen bleibt die lineare Unabhängigkeit. Man betrachte dazu die Linearkombination

$$\sum_{j=0}^{m-1} (u_j \operatorname{Re} z_j + v_j \operatorname{Im} z_j) = 0$$

mit $u_j, v_j \in \mathbb{R}$. Nun gilt aber

$$\begin{aligned} u_j \operatorname{Re} z_j + v_j \operatorname{Im} z_j &= \frac{u_j}{2}(z_j + \bar{z}_j) + \frac{v_j}{2i}(z_j - \bar{z}_j) \\ &= \frac{u_j}{2}(z_j + \bar{z}_j) - \frac{iv_j}{2}(z_j - \bar{z}_j) \\ &= \frac{u_j - iv_j}{2}z_j + \frac{u_j + iv_j}{2}\bar{z}_j. \end{aligned}$$

Also folgt

$$0 = \sum_{j=0}^{m-1} (u_j \operatorname{Re} z_j + v_j \operatorname{Im} z_j) = \sum_{j=0}^{m-1} \left(\frac{u_j - iv_j}{2}z_j + \frac{u_j + iv_j}{2}\bar{z}_j \right).$$

Da die Funktionen $\{z_0, \dots, z_{m-1}\} \cup \{\bar{z}_0, \dots, \bar{z}_{m-1}\}$ im komplexen Vektorraum $C^\infty(\mathbb{R}, \mathbb{C})$ linear unabhängig sind, folgt $u_j - iv_j = u_j + iv_j = 0$ und somit $u_j = v_j = 0$, für alle $j \in \{1, \dots, m\}$. \square

Zusammenfassend erhalten wir nun den folgenden Satz über den Lösungsraum von reellen homogenen linearen Differentialgleichungen mit konstanten Koeffizienten.

Satz 9.3.2. Sei $L : C^n(\mathbb{R}, \mathbb{R}) \rightarrow C^0(\mathbb{R}, \mathbb{R})$ der lineare reelle Differentialoperator mit

$$Ly = y^{(n)} + a_{n-1}y^{(n-1)} + \dots + a_0y.$$

Man betrachte das zugehörige Polynom

$$p(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_0.$$

Für $j \in \{1, \dots, k\}$ seien λ_j alle seine reellen Nullstellen mit Vielfachheit k_j . Für $j \in \{1, \dots, m\}$ seien $\mu_j = \alpha_j + i\beta_j$, $\beta_j > 0$ und $\bar{\mu}_j = \alpha_j - i\beta_j$ alle nicht reellen Nullstellen jeweils mit Vielfachheit m_j . Man betrachte die reellen Unterräume

$$V_{\lambda_j} := \operatorname{Span}_{\mathbb{R}} \{e^{\lambda_j t}, te^{\lambda_j t}, t^2 e^{\lambda_j t}, \dots, t^{k_j-1} e^{\lambda_j t}\}$$

sowie

$$V_{(\mu_j, \bar{\mu}_j)} := \text{Span}_{\mathbb{R}}\{e^{\alpha_j t} \cos \beta_j t, e^{\alpha_j t} \sin \beta_j t, \dots, t^{m_j-1} e^{\alpha_j t} \cos \beta_j t, t^{m_j-1} e^{\alpha_j t} \sin \beta_j t\}.$$

Dann gilt:

$$\text{Kern } L = V_{\lambda_1} \oplus \dots \oplus V_{\lambda_k} \oplus V_{(\mu_1, \bar{\mu}_1)} \oplus \dots \oplus V_{(\mu_m, \bar{\mu}_m)}.$$

Beweis. Die reellen Vektorräume V_{λ_j} und $V_{(\mu_j, \bar{\mu}_j)}$ sind in Kern L enthalten. Ihre Summe ist direkt, denn wie wir oben gesehen haben, ist jeder Summand in einem komplexen Unterraum enthalten. Da

$$\dim V_{\lambda_j} = k_j \quad \text{und} \quad \dim V_{(\mu_j, \bar{\mu}_j)} = 2m_j,$$

folgt

$$\dim V_{\lambda_1} \oplus \dots \oplus V_{\lambda_k} \oplus V_{(\mu_1, \bar{\mu}_1)} \oplus \dots \oplus V_{(\mu_m, \bar{\mu}_m)} = \sum_{j=1}^k k_j + \sum_{j=1}^m 2m_j = \text{Grad}(p) = n.$$

Da wegen des Satzes 9.1.3 die Abschätzung $\dim \text{Kern } L \leq n$ gilt, folgt die Behauptung. \square

Beispiele.

1. Homogene Schwingungsgleichung.

Bei der homogenen Schwingungsgleichung handelt es sich um eine homogene lineare Differentialgleichung 2. Ordnung mit reellen Koeffizienten. Sie ist also von der Form

$$L(y) = y'' + a_1 y' + a_0 y = 0,$$

mit $a_1, a_0 \in \mathbb{R}$. Der Lösungsraum Kern L der homogenen Gleichung ist 2-dimensional. In Abhängigkeit von den Nullstellen des zugehörigen Polynoms

$$p(\lambda) = \lambda^2 + a_1 \lambda + a_0$$

gegeben durch

$$-\frac{a_1}{2} \pm \sqrt{\Delta}$$

mit

$$\Delta = \frac{a_1^2 - 4a_0}{4}$$

unterscheidet man drei Fälle:

($\Delta > 0$): Dann sind die Nullstellen λ_1, λ_2 von p reell und verschieden. Es folgt

$$\text{Kern } L = V_{\lambda_1} \oplus V_{\lambda_2}$$

mit $V_{\lambda_j} = \text{Span}\{e^{\lambda_j t}\}$. Also existieren zu jeder Lösung $y \in \text{Kern } L$ Konstanten $c_1, c_2 \in \mathbb{R}$ mit

$$y(t) = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t}.$$

($\Delta = 0$): Die Nullstellen λ_1, λ_2 von p sind reell und es gilt $\lambda_1 = \lambda_2 = \lambda = -\frac{a_1}{2}$. Es folgt

$$\text{Kern } L = V_\lambda = \text{Span}\{e^{\lambda t}, te^{\lambda t}\}.$$

Also existieren zu jeder Lösung $y \in \text{Kern } L$ Konstanten $c_1, c_2 \in \mathbb{R}$ mit

$$y(t) = c_1 e^{\lambda t} + c_2 t e^{\lambda t} = e^{-\frac{a_1}{2}t} (c_1 + c_2 t).$$

($\Delta < 0$): Die Nullstellen sind nicht reell. In diesem Falle sind sie von der Form $\mu = \alpha + i\beta$ und $\bar{\mu} = \alpha - i\beta$ mit $\beta = \sqrt{|\Delta|} > 0$ und $\alpha = -\frac{a_1}{2}$. Es folgt

$$\text{Kern } L = V_{(\mu, \bar{\mu})} = \text{Span}\{e^{\alpha t} \cos \beta t, e^{\alpha t} \sin \beta t\}.$$

Also existieren zu jeder Lösung $y \in \text{Kern } L$ Konstanten $c_1, c_2 \in \mathbb{R}$ mit

$$y(t) = c_1 e^{\alpha t} \cos \beta t + c_2 e^{\alpha t} \sin \beta t = e^{-\frac{a_1}{2}t} \left(c_1 \cos \left(\sqrt{|\Delta|} t \right) + c_2 \sin \left(\sqrt{|\Delta|} t \right) \right).$$

2. Man betrachte die lineare Differentialgleichung 4. Ordnung gegeben durch

$$L(y) = y^{(4)} + 8y^{(2)} + 16y = 0.$$

Das zugehörige Polynom ist durch

$$\lambda^4 + 8\lambda^2 + 16 = (\lambda^2 + 4)^2 = (\lambda - 2i)^2 (\lambda + 2i)^2$$

gegeben. Also sind $\mu = 2i$, $\bar{\mu} = -2i$ jeweils Nullstellen der Multiplizität 2 und es gilt:

$$\text{Kern } L = V_{(\mu, \bar{\mu})} = \text{Span}\{\cos 2t, \sin 2t, t \cos 2t, t \sin 2t\}.$$

Also existieren zu jeder Lösung $y \in \text{Kern } L$ Konstanten $c_1, c_2, c_3, c_4 \in \mathbb{R}$ mit

$$y(t) = c_1 \cos 2t + c_2 \sin 2t + c_3 t \cos 2t + c_4 t \sin 2t.$$

9.4 Inhomogene lineare Differentialgleichungen mit reellen und komplexen Koeffizienten

Wir betrachten zunächst einen wichtigen Spezialfall, bei dem sich eine spezielle Lösung der inhomogenen linearen Differentialgleichung leicht herleiten lässt. Man betrachte den linearen Differentialoperator $L : C^n(\mathbb{R}, \mathbb{C}) \rightarrow C^0(\mathbb{R}, \mathbb{C})$ mit

$$Ly = y^{(n)} + a_{n-1}y^{(n-1)} + \dots + a_0y$$

und komplexen Koeffizienten $a_j \in \mathbb{C}$. Für jedes $\mu \in \mathbb{C}$ wollen wir nun eine spezielle Lösung der inhomogenen Differentialgleichung $L(y) = e^{\mu t}$ berechnen. Dazu ist folgendes Lemma hilfreich:

Lemma 9.4.1. Sei $p(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_0$ ein Polynom und $p(D) = L$ der zugehörige Differentialoperator. Dann gilt:

$$L(e^{\mu t}) = p(\mu)e^{\mu t}.$$

Ist μ eine Nullstelle von p der Ordnung k und q das Polynom mit

$$p(\lambda) = (\lambda - \mu)^k q(\lambda),$$

so ist $q(\mu) \neq 0$ und es gilt:

$$L(t^k e^{\mu t}) = k! q(\mu) e^{\mu t} = p^{(k)}(\mu) e^{\mu t}.$$

Beweis. Es gilt:

$$\begin{aligned} L(e^{\mu t}) &= D^n(e^{\mu t}) + a_{n-1}D^{n-1}(e^{\mu t}) + \dots + a_0e^{\mu t} \\ &= (\mu^n + a_{n-1}\mu^{n-1} + \dots + a_0)e^{\mu t} \\ &= p(\mu)e^{\mu t}. \end{aligned}$$

Ist nun μ eine Nullstelle von p der Ordnung k und q das Polynom mit

$$p(\lambda) = (\lambda - \mu)^k q(\lambda),$$

so folgt aus Lemma 9.2.1

$$\begin{aligned} L(t^k e^{\mu t}) &= q(D)(D - \mu \text{id})^k (t^k e^{\mu t}) = q(D)(k! e^{\mu t}) \\ &= k! q(\mu) e^{\mu t} = p^{(k)}(\mu) e^{\mu t}. \end{aligned}$$

□

Aus diesem Lemma ergibt sich sofort der folgende Satz:

Satz 9.4.2. Sei μ eine Nullstelle von p der Ordnung k mit $k \in \mathbb{N}$. Ist $p(\mu) \neq 0$, so setze $k = 0$. Dann besitzt für $A \in \mathbb{C}$ die Differentialgleichung

$$p(D)z = Ae^{\mu t}$$

die Funktion

$$z(t) = \frac{A}{p^{(k)}(\mu)} t^k e^{\mu t}$$

als spezielle Lösung.

Damit erhalten wir für Differentialgleichungen mit reellen Koeffizienten den folgenden Satz:

Satz 9.4.3. Sei $p(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_0$ ein Polynom und $p(D) = L$ der zugehörige Differentialoperator. Betrachte zu $\omega, A \in \mathbb{R}$ die inhomogene Differentialgleichung

$$p(D)y = A \cos(\omega t).$$

Sei $i\omega$ eine Nullstelle der Ordnung k von p mit $k \in \mathbb{N}$. Ist $p(i\omega) \neq 0$, so setze $k = 0$. Dann existiert ein $\varphi \in [0, 2\pi]$, so dass

$$y(t) = \frac{|A|}{|p^{(k)}(i\omega)|} t^k \cos(\omega t + \varphi)$$

eine spezielle Lösung ist. Dabei ist φ durch die Darstellung

$$\frac{A}{p^{(k)}(i\omega)} = \frac{|A|}{|p^{(k)}(i\omega)|} e^{i\varphi}$$

bestimmt.

Beweis. Betrachte die inhomogene Differentialgleichung

$$p(D)z = Ae^{i\omega t}.$$

Dann ist wegen obigem Satz

$$z(t) = \frac{A}{p^{(k)}(i\omega)} t^k e^{i\omega t}$$

eine spezielle Lösung dieser Gleichung und $y(t) = \operatorname{Re}(z(t))$ ist eine Lösung der Gleichung

$$p(D)y = A \cos(\omega t).$$

Schreiben wir nun $\frac{A}{p^{(k)}(i\omega)}$ in Polarkoordinaten (siehe Satz 3.2.8), d.h.

$$\frac{A}{p^{(k)}(i\omega)} = \frac{|A|}{|p^{(k)}(i\omega)|} e^{i\varphi},$$

mit $\varphi \in [0, 2\pi]$, so folgt:

$$\begin{aligned} y(t) = \operatorname{Re}(z(t)) &= \operatorname{Re}\left(\frac{|A|}{|p^{(k)}(i\omega)|} t^k e^{i(\omega t + \varphi)}\right) \\ &= \frac{|A|}{|p^{(k)}(i\omega)|} t^k \cos(\omega t + \varphi). \end{aligned}$$

□

Bemerkungen.

- (a) Für eine Inhomogenität $b(t) = A \sin(\omega t)$ lässt sich analog eine partikuläre Lösung der Form

$$y(t) = \frac{|A|}{|p^{(k)}(i\omega)|} t^k \sin(\omega t + \varphi)$$

bestimmen.

- (b) Hat die Inhomogenität die Form $b(t) = b_1(t) + b_2(t)$, so bestimme man die partikulären Lösungen $y_1(t)$ und $y_2(t)$ bezüglich der Inhomogenitäten $b_1(t)$ bzw. $b_2(t)$. Dann ist $y(t) = y_1(t) + y_2(t)$ eine partikuläre Lösung zu der Inhomogenität $b(t)$.

Als nächstes wollen wir uns dem Problem der Lösung einer inhomogenen linearen Differentialgleichung im allgemeinen Fall zuwenden. Hierbei kann die Inhomogenität im Prinzip beliebig sein. Wir werden dabei den reellen und komplexen Fall simultan behandeln.

Sei also $L : C^n(I, K) \rightarrow C^0(I, K)$ mit

$$Ly = y^{(n)} + a_{n-1}y^{(n-1)} + \dots + a_0y,$$

wobei K der Körper der reellen oder komplexen Zahlen und I ein Intervall bezeichnen. Sei $b \in C^0(I, K)$ und man betrachte die inhomogene lineare Differentialgleichung mit

$$Ly = y^{(n)} + a_{n-1}y^{(n-1)} + \dots + a_0y = b.$$

Ist y eine spezielle Lösung dieser Gleichung (partikuläre Lösung), so ist die Menge aller Lösungen durch

$$\mathcal{L}_{L,b} = \text{Kern } L + y_0$$

gegeben. Das Auffinden dieser Lösung ist mit der Methode der Variation der Konstanten möglich. Ist $\{y_1, \dots, y_n\}$ eine Basis von $\text{Kern } L$, so werden wir zeigen, dass n Funktionen $\{f_1, \dots, f_n\}$ existieren, so dass

$$y = f_1 y_1 + \dots + f_n y_n$$

eine spezielle Lösung der inhomogenen Gleichung darstellt. Dazu benötigen wir folgendes Lemma:

Lemma 9.4.4. *Sei $\{y_1, \dots, y_n\}$ eine Basis von $\text{Kern } L$. Dann sind für jedes $t_0 \in I$ die Vektoren $a_j(t_0) \in K^n$ mit $a_j(t_0) = (y_j(t_0), y_j'(t_0), \dots, y_j^{(n-1)}(t_0))$ linear unabhängig.*

Beweis. Sei also $x_1 a_1(t_0) + \dots + x_n a_n(t_0) = 0$ mit $x_i \in K$ eine Linearkombination des Nullvektors in K^n . Dann ist die Funktion $z(t) = x_1 y_1(t) + \dots + x_n y_n(t)$ im Kern von L enthalten. Daraus folgt aber

$$z(t_0) = z'(t_0) = \dots = z^{(n)}(t_0) = 0$$

und wegen Satz 9.1.1 ist z identisch null. Da die Funktionen $\{y_1, \dots, y_n\}$ linear unabhängig sind, folgt $x_1 = x_2 = \dots = x_n = 0$. \square

Damit erhalten wir folgenden Satz:

Satz 9.4.5. (Variation der Konstanten)

Sei $b : I \rightarrow K$ eine stetige Funktion und $\{y_1, \dots, y_n\}$ eine Basis von $\text{Kern } L$. Dann existieren Funktionen $\{u_1, \dots, u_n\}$ mit

$$u_1(t) \begin{pmatrix} y_1(t) \\ y_1'(t) \\ \vdots \\ y_1^{(n-1)}(t) \end{pmatrix} + u_2(t) \begin{pmatrix} y_2(t) \\ y_2'(t) \\ \vdots \\ y_2^{(n-1)}(t) \end{pmatrix} + \dots + u_n(t) \begin{pmatrix} y_n(t) \\ y_n'(t) \\ \vdots \\ y_n^{(n-1)}(t) \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ b(t) \end{pmatrix}.$$

Sind U_1, \dots, U_n Stammfunktionen von u_1, \dots, u_n , so ist

$$y(t) = U_1(t)y_1(t) + \dots + U_n(t)y_n(t)$$

eine spezielle Lösung der inhomogenen Differentialgleichung $L(y) = b$.

Beweis. Da für jedes $t \in I$ die Vektoren $a_j(t) = (y_j(t), y_j'(t), \dots, y_j^{(n-1)}(t))$ linear unabhängig sind, folgt die Existenz der Funktionen $\{u_1, \dots, u_n\}$. (Durch Auflösen des linearen Gleichungssystems kann man zeigen, dass die Funktionen u_1, \dots, u_n stetig sind.) Dann erhal-

ten wir

$$\begin{aligned}
 y &= U_1 y_1 + \dots + U_n y_n \\
 y' &= u_1 y_1 + \dots + u_n y_n + U_1 y_1' + \dots + U_n y_n' \\
 &= U_1 y_1' + \dots + U_n y_n' \\
 y'' &= u_1 y_1' + \dots + u_n y_n' + U_1 y_1'' + \dots + U_n y_n'' \\
 &= U_1 y_1'' + \dots + U_n y_n'' \\
 &\vdots \\
 y^{(n)} &= u_1 y_1^{(n-1)} + \dots + u_n y_n^{(n-1)} + U_1 y_1^{(n)} + \dots + U_n y_n^{(n)} \\
 &\quad b + U_1 y_1^{(n)} + \dots + U_n y_n^{(n)}.
 \end{aligned}$$

Somit folgt:

$$\begin{aligned}
 L(y) &= y^{(n)} + a_{n-1} y^{(n-1)} + \dots + a_0 y \\
 &= U_1 (y_1^{(n)} + a_{n-1} y_1^{(n-1)} + \dots + a_0 y_1) + \\
 &\quad + U_2 (y_2^{(n)} + a_{n-1} y_2^{(n-1)} + \dots + a_0 y_2) + \dots + \\
 &\quad + U_n (y_n^{(n)} + a_{n-1} y_n^{(n-1)} + \dots + a_0 y_n) + b \\
 &= b.
 \end{aligned}$$

□

Bemerkung. Betrachten wir den Beweis des obigen Satzes, so stellen wir fest, dass wir die Konstanz der Koeffizienten des Operators L mit $Ly = y^{(n)} + a_{n-1} y^{(n-1)} + \dots + a_0 y$ nicht benutzt haben. Haben wir also ein Fundamentalsystem (y_1, \dots, y_n) , d.h. eine Basis von Kern L bestimmt, so können wir, mit dem im obigen Satz beschriebenen Verfahren, eine spezielle Lösung finden. Allerdings ist das Auffinden eines Fundamentalsystems im Falle variabler Koeffizienten und Ordnung strikt grösser als eins ein schwieriges Problem. Sie lassen sich im Allgemeinen nicht in geschlossener Form angeben.

Kapitel 10

Matrizen

Die meisten konkreten Rechnungen in der linearen Algebra werden mit Matrizen durchgeführt. Sie spielen z.B. bei Darstellungen linearer Abbildungen zwischen endlich dimensionalen Vektorräumen sowie bei linearen Gleichungssystemen eine wichtige Rolle.

10.1 Grundlegendes über Matrizen

Sei $A : K^n \rightarrow K^m$ eine lineare Abbildung und (e_1, \dots, e_n) die kanonische Basis von K^n . Betrachte die Vektoren $\{A(e_1), \dots, A(e_n)\}$ mit

$$A(e_j) = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} \in K^m.$$

Dann gilt für $x = \sum_{j=1}^n x_j e_j$:

$$A(x) = \sum_{j=1}^n x_j A(e_j) = \sum_{j=1}^n x_j \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n \end{pmatrix}.$$

Definieren wir

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n \end{pmatrix},$$

so erhalten wir für die lineare Abbildung $A : K^n \rightarrow K^m$ die Darstellung

$$A(x) = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Definition 10.1.1. Sei K ein Körper und seien $a_{ij} \in K$ mit $i \in \{1, \dots, m\}$ und $j \in \{1, \dots, n\}$, $n, m \in \mathbb{N}$. Dann heißt das Schema

$$(a_{ij}) = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

bestehend aus m Zeilen und n Spalten eine $m \times n$ -Matrix mit Koeffizienten (Werten) in K . Die Zahl i heißt *Zeilenindex*, die Zahl j heißt *Spaltenindex* von (a_{ij}) . Die Zeilen nennt man auch *Zeilenvektoren*, die Spalten *Spaltenvektoren*. Mit $M(m \times n, K)$ bezeichnen wir die Menge aller $m \times n$ -Matrizen mit Koeffizienten in K . Ist $m = n$, so schreiben wir $M(n, K)$ statt $M(n \times n, K)$.

Bemerkungen.

1. Wie oben erläutert können wir jede $m \times n$ -Matrix auch als lineare Abbildung $A : K^n \rightarrow K^m$ auffassen und jeder linearen Abbildung $A : K^n \rightarrow K^m$ entspricht genau eine Matrix bezüglich der kanonischen Basen in K^n und K^m . Ihre n Spaltenvektoren bestehen aus den Vektoren $(A(e_1), \dots, A(e_n))$ in K^m . Insbesondere entspricht der Identität $\text{id} : K^n \rightarrow K^n$ die Matrix E_n , bestehend aus den Spaltenvektoren (e_1, \dots, e_n) . Die Matrix E_n heißt auch *Einheitsmatrix*. Wir werden im nächsten Abschnitt lineare Abbildungen auch bezüglich beliebiger Basen darstellen.
2. Wir werden im Folgenden nicht mehr zwischen Matrizen $(a_{ij}) \in M(m \times n, K)$ und den zugehörigen linearen Abbildungen $A : K^n \rightarrow K^m$ unterscheiden.
3. Die Menge $M(m \times n, K)$ bildet einen Vektorraum über K . Die Addition ist durch

$$(a_{ij}) + (b_{ij}) := (a_{ij} + b_{ij})$$

gegeben und die skalare Multiplikation durch

$$\alpha(a_{ij}) := (\alpha a_{ij})$$

für $\alpha \in K$. Die Addition und skalare Multiplikation in $M(m \times n, K)$ entspricht der Addition und skalaren Multiplikation im Vektorraum der linearen Abbildungen $L(K^n, K^m)$.

Es seien nun $A : K^n \rightarrow K^m$ und $B : K^\ell \rightarrow K^n$ lineare Abbildungen und $(a_{ij}) \in M(m \times n, K)$ und $(b_{ij}) \in M(n \times \ell, K)$ die zu A und B gehörigen Matrizen. Dann ist auch die Verknüpfung $C := A \circ B : K^\ell \rightarrow K^m$ eine lineare Abbildung und wegen den obigen Überlegungen entspricht C genau einer Matrix $(c_{ij}) \in M(m \times \ell, K)$.

Wir berechnen die Matrix c_{ij} , indem wir C auf die kanonische Basis $e_j \in K^\ell$ anwenden. Wir erhalten dann

$$\begin{pmatrix} c_{1j} \\ \vdots \\ c_{mj} \end{pmatrix} = C(e_j) = A(B(e_j)) = A \begin{pmatrix} b_{1j} \\ \vdots \\ b_{nj} \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} b_{1j} \\ \vdots \\ b_{nj} \end{pmatrix},$$

d.h. $c_{ij} = \sum_{k=1}^n a_{ik} b_{kj}$ (c_{ij} entsteht also aus der Multiplikation der i -ten Zeile von A mit der j -ten Spalte von B).

Definition 10.1.2. Seien $A = (a_{ij}) \in M(m \times n, K)$ und $B = (b_{ij}) \in M(n \times \ell, K)$ Matrizen, so heißt die Matrix $A \cdot B := C = (c_{ij}) \in M(m \times \ell, K)$ mit

$$c_{ij} = \sum_{k=1}^n a_{ik} b_{kj}$$

das *Produkt der Matrizen* A und B .

Bemerkung. Das Matrizenprodukt $A \cdot B$ entspricht also der Verknüpfung $A \circ B$ der linearen Abbildungen $B : K^\ell \rightarrow K^n$ und $A : K^n \rightarrow K^m$. Es ist genau dann definiert, falls die Zeilenzahl von B mit der Spaltenzahl von A übereinstimmt.

Wie können wir an einer Matrix A ablesen, ob A injektiv, bijektiv bzw. surjektiv ist?

Satz 10.1.3. Sei $A \in M(m \times n, K)$, $A = (a_{ij})$. Dann gilt:

- a) $A : K^n \rightarrow K^m$ ist injektiv \Leftrightarrow die Spaltenvektoren $\{A(e_1), \dots, A(e_n)\} \in K^m$ sind linear unabhängig. Insbesondere ist dann $n \leq m$.
- b) $A : K^n \rightarrow K^m$ ist surjektiv \Leftrightarrow die Spaltenvektoren bilden ein Erzeugendensystem von K^m . Insbesondere ist dann $n \geq m$.
- c) A ist bijektiv \Leftrightarrow die Spaltenvektoren bilden eine Basis von K^m . Insbesondere ist dann $m = n$.

Beweis. Sei e_1, \dots, e_n die kanonische Basis von K^n . Dann folgt aus Satz 8.2.4: $A : K^n \rightarrow K^m$ ist genau dann injektiv, falls $\{A(e_1), \dots, A(e_n)\} \subset K^m$ linear unabhängig ist. Insbesondere ist dann $n \leq m$.

Aus Satz 8.2.1 folgt: $A : K^n \rightarrow K^m$ ist genau dann surjektiv, falls $\{A(e_1), \dots, A(e_n)\} \subset K^m$ ein Erzeugendensystem von K^m ist. Insbesondere ist dann $n \geq m$.

Teil c) folgt aus a) und b). □

Bemerkungen.

1. Sei $A \in M(n, K)$ mit A bijektiv. Dann existiert eine Matrix $B \in M(n, K)$ (sie entspricht der zu A inversen Abbildung) mit

$$B \cdot A = A \cdot B = E_n.$$

Die Matrix $A^{-1} := B$ heißt die zu A *inverse Matrix*. Die inverse Matrix existiert wegen Satz 10.1.3 genau dann, wenn die n Spaltenvektoren von A linear unabhängig sind und somit eine Basis von K^n bilden. Wie wir bald sehen werden, ist dies äquivalent zur linearen Unabhängigkeit der n Zeilenvektoren.

2. Die Menge der invertierbaren Matrizen

$$\text{GL}(n, K) := \{A \in M(n, K) \mid A \text{ invertierbar}\}$$

bildet bezüglich der Matrizenmultiplikation eine Gruppe. Dies folgt aus der entsprechenden Aussage in Korollar 8.1.5 für lineare Abbildungen.

10.2 Matrizen und lineare Abbildungen

Unter Benutzung von Koordinatensystemen können wir einer beliebigen linearen Abbildung eine Matrix zuordnen. Die Zuordnung ist nicht kanonisch, da in allgemeinen Vektorräumen keine ausgezeichneten Basen existieren.

Definition 10.2.1. Seien V, W K -Vektorräume mit Basen $B = (b_1, \dots, b_n) \subset V$ und $C = (c_1, \dots, c_m) \subset W$. Es seien $\phi_B : V \rightarrow K^n$ und $\phi_C : W \rightarrow K^m$ die durch B und C definierten Koordinatensysteme. Sei $f : V \rightarrow W$ eine lineare Abbildung, so heißt die lineare Abbildung

$$\phi_C \circ f \circ \phi_B^{-1} : K^n \rightarrow K^m$$

die *Koordinatendarstellung* von f .

Die zugehörige Matrix $A = M_{B,C}(f) := \phi_C \circ f \circ \phi_B^{-1} \in M(m \times n, K)$ heißt auch die *Matrixdarstellung* von f bezüglich der Basen B und C von V bzw. W . Man kann sich die Definition auch durch das folgende kommutative Diagramm veranschaulichen:

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ \phi_B \downarrow & & \downarrow \phi_C \\ K^n & \xrightarrow{A} & K^m \end{array}$$

Bemerkung. In manchen Büchern (z.B. Fischer) findet man auch die Notation $A = M_B^C(f)$.

Übung: Die Abbildung

$$M_{B,C} : \text{Hom}_K(V, W) \rightarrow M(m \times n, K)$$

mit $f \mapsto \phi_C \circ f \circ \phi_B^{-1}$ ist ein Vektorraumisomorphismus.

Bemerkung. Ist $V = K^n, W = K^m$ und sind $B = (e_1, \dots, e_n)$ und $C = (e_1, \dots, e_m)$ die kanonischen Basen von K^n und K^m , so ist $\phi_B = \text{id}_{K^n}$ und $\phi_C = \text{id}_{K^m}$. Somit folgt $M_{B,C}(A) = A$.

Die Matrixdarstellung von f bezüglich B und C können wir wie folgt explizit beschreiben.

Satz 10.2.2. Sei $f : V \rightarrow W$ eine lineare Abbildung und $B = (b_1, \dots, b_n)$ sowie $C = (c_1, \dots, c_m)$ seien Basen von V bzw. W . Sei

$$(a_{ij}) = A = M_{B,C}(f) \in M(m \times n, K)$$

die Matrixdarstellung von f bezüglich B und C , so erfüllen die Koeffizienten a_{ij} die Gleichungen

$$f(b_j) = \sum_{i=1}^m a_{ij} c_i, \quad (*)$$

mit $j \in \{1, \dots, n\}$ und werden durch sie eindeutig festgelegt.

Bemerkung. Der j -te Spaltenvektor von A besteht aus den Koordinaten von $f(b_j)$ bezüglich der Basis $C = (c_1, \dots, c_m)$.

Beweis. Die Eindeutigkeit folgt natürlich daraus, dass C eine Basis von W ist. Die Koeffizienten (a_{ij}) der Matrix A erfüllen $(*)$, denn

$$\begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} = \phi_C \circ f \circ \phi_B^{-1}(e_j) = \phi_C(f(b_j)).$$

Also folgt

$$f(b_j) = \phi_C^{-1}(a_{1j}, \dots, a_{mj}) = \sum_{i=1}^m a_{ij} c_i.$$

□

Bemerkung. Man kann a_{ij} mittels der dualen Basis (c_1^*, \dots, c_m^*) von C auch folgendermaßen darstellen:

$$a_{ij} = c_i^*(f(b_j)) = \langle c_i^*, f(b_j) \rangle,$$

wobei \langle, \rangle die kanonische Paarung auf W^* und W bezeichnet (siehe Bemerkung nach Definition 8.3.5).

Nun wollen wir die Matrix zu einer gegebenen Koordinatentransformation bestimmen.

Satz 10.2.3. Es seien $B = (b_1, \dots, b_n)$ und $B' = (b'_1, \dots, b'_n)$ Basen des K -Vektorraumes V . Seien $\phi_B : V \rightarrow K^n$ und $\phi_{B'} : V \rightarrow K^n$ die zu B und B' gehörigen Koordinatensysteme. Dann erfüllt die zur Koordinatentransformation $T_{B,B'} = \phi_{B'} \circ \phi_B^{-1} : K^n \rightarrow K^n$ gehörige Matrix

$$(t_{ij}) = T = T_{B,B'}$$

die Gleichungen

$$b_j = \sum_{i=1}^n t_{ij} b'_i$$

für $j \in \{1, \dots, n\}$. Die Matrix T heißt auch Transformationsmatrix.

Beweis. Nach Definition 10.2.1 ist $T_{B,B'} = M_{B,B'}(\text{id}_V)$. Damit folgt die Behauptung aus Satz 10.2.2. □

Bemerkungen.

a) Hat also ein Vektor $v = \sum_{i=1}^n x_i b_i$ die Koordinaten $x = (x_1, \dots, x_n)$ bezüglich B , so hat er die Koordinaten $x' = (x'_1, \dots, x'_n)$ bezüglich B' , wobei $x' = T_{B,B'}(x)$.

b) Es gilt: $(T_{B,B'})^{-1} = (\phi_{B'} \circ \phi_B^{-1})^{-1} = (\phi_B^{-1})^{-1} \circ \phi_{B'}^{-1} = \phi_B \circ \phi_{B'}^{-1} = T_{B',B}$.

Beispiel. $V = K^3$, $B = ((1, 0, 0), (0, 1, 0), (0, 0, 1))$ sei die kanonische Basis und $B' = ((1, 0, 0), (1, 1, 0), (1, 1, 1))$ eine weitere Basis von K^3 . Sei $(t_{ij}) = T_{B,B'}$, so gilt:

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = t_{11} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + t_{21} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + t_{31} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$\begin{aligned} &\Rightarrow t_{11} = 1, \quad t_{21} = 0, \quad t_{31} = 0 \\ \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} &= t_{12} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + t_{22} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + t_{32} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \\ &\Rightarrow t_{12} = -1, \quad t_{22} = 1, \quad t_{32} = 0 \\ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} &= t_{13} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + t_{23} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + t_{33} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \\ &\Rightarrow t_{13} = 0, \quad t_{23} = -1, \quad t_{33} = 1. \end{aligned}$$

Also:

$$T_{B,B'} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Genauso sieht man:

$$T_{B',B} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} = (T_{B,B'})^{-1}.$$

Wie hängen nun die Matrixdarstellungen einer linearen Abbildung bezüglich verschiedener Koordinatensysteme zusammen?

Satz 10.2.4. Sei $f : V \rightarrow W$ eine lineare Abbildung, $B = (b_1, \dots, b_n)$, $B' = (b'_1, \dots, b'_n)$ Basen von V und $C = (c_1, \dots, c_m)$, $C' = (c'_1, \dots, c'_m)$ Basen von W . Ist

$$M_{B,C}(f) \in M(m \times n, K)$$

die Matrixdarstellung von f bezüglich B und C und

$$M_{B',C'}(f) \in M(m \times n, K)$$

die Matrixdarstellung von f bezüglich B' und C' , so gilt:

$$M_{B',C'}(f) = T_{C,C'} \cdot M_{B,C}(f) \cdot T_{B,B'}^{-1},$$

wobei $T_{C,C'} = \phi_{C'} \circ \phi_C^{-1}$ und $T_{B,B'} = \phi_{B'} \circ \phi_B^{-1}$ die zugehörigen Koordinatentransformationen bezeichnen.

Beweis. Es gilt:

$$M_{B,C}(f) = \phi_C \circ f \circ \phi_B^{-1} \quad \text{sowie} \quad M_{B',C'}(f) = \phi_{C'} \circ f \circ \phi_{B'}^{-1}.$$

Daraus folgt:

$$M_{B',C'}(f) = \phi_{C'} \circ \phi_C^{-1} \circ \phi_C \circ f \circ \phi_B^{-1} \circ \phi_B \circ \phi_{B'}^{-1} = T_{C,C'} \circ M_{B,C}(f) \circ (T_{B,B'})^{-1}.$$

□

Die Matrixdarstellung einer Verknüpfung zweier linearer Abbildungen entspricht natürlich der Verknüpfung der Matrixdarstellungen. Genauer gilt:

Satz 10.2.5. *Es seien $f : U \rightarrow V$ und $g : V \rightarrow W$ lineare Abbildungen zwischen endlich dimensionalen Vektorräumen. Es seien $B \subset U, C \subset V, D \subset W$ Basen. Dann folgt:*

$$M_{B,D}(g \circ f) = M_{C,D}(g) \cdot M_{B,C}(f) .$$

Beweis. Es seien $\phi_B : U \rightarrow K^n, \phi_C : V \rightarrow K^m, \phi_D : W \rightarrow K^\ell$ Koordinatendarstellungen, so hat man das folgende kommutative Diagramm

$$\begin{array}{ccccccc} U & \xrightarrow{f} & V & \xrightarrow{g} & W & & \\ \phi_B \downarrow & & \downarrow \phi_C & & \downarrow \phi_D & \cdot & \\ K^n & \xrightarrow{M_{B,C}(f)} & K^m & \xrightarrow{M_{C,D}(g)} & K^\ell & & \end{array}$$

Daraus folgt:

$$M_{B,D}(g \circ f) = \phi_D \circ (g \circ f) \circ \phi_B^{-1} = (\phi_D \circ g \circ \phi_C^{-1}) \cdot (\phi_C \circ f \circ \phi_B^{-1}) = M_{C,D}(g) \cdot M_{B,C}(f).$$

□

10.3 Der Rang einer linearen Abbildung

Eine wichtige Zahl, die wir einer linearen Abbildung zuordnen können, ist ihr Rang.

Definition 10.3.1. Sei $f : V \rightarrow W$ eine lineare Abbildung, so heißt die Zahl

$$\text{Rang } f = \dim \text{Bild } f$$

der *Rang der Abbildung* f .

Bemerkungen.

- Da Bild f ein Unterraum von W ist, folgt $\text{Rang } f \leq \dim W$. Es gilt aber auch $\text{Rang } f \leq \dim V$. Ist $\dim V = \infty$, so ist dies automatisch erfüllt. Ist $\dim V = n$ und $B = (b_1, \dots, b_n)$ eine Basis von V , so ist wegen Satz 8.2.1 die Menge $E = \{f(b_1), \dots, f(b_n)\}$ ein Erzeugendensystem von Bild f und daher ist $\text{Rang } f \leq n = \dim V$.
- Ist $B = (b_1, \dots, b_n)$ eine Basis von V und F eine linear unabhängige Teilmenge von $E = \{f(b_1), \dots, f(b_n)\}$, so folgt $|F| \leq \dim \text{Bild } f$ aus Satz 7.2.15. Auf der anderen Seite enthält E wegen des Austauschsatzes 7.2.12 eine Basis von Bild f . Damit gilt also:

$$\dim \text{Bild } f = \max\{|F| \mid F \subset E, F \text{ linear unabhängig}\},$$

d.h. $\text{Rang } f$ ist die Maximalzahl linear unabhängiger Vektoren in $\{f(b_1), \dots, f(b_n)\}$.

- Sei $A \in M(m \times n, K)$, so sind die Spaltenvektoren $\{A(e_1), \dots, A(e_n)\}$ ein Erzeugendensystem von Bild A und somit ist $\text{Rang } A$ die Maximalzahl linear unabhängiger Spaltenvektoren.

Definition 10.3.2. Sei $A \in M(m \times n, K)$, so heißt die Maximalzahl linear unabhängiger Spaltenvektoren auch *Spaltenrang* der Matrix A . Der *Zeilenrang* ist die Maximalzahl linear unabhängiger Zeilenvektoren von A .

Bemerkung. Damit ist also der Spaltenrang einer Matrix A gleich dem Rang der linearen Abbildung A . Wir werden bald sehen, dass der Spaltenrang mit dem Zeilenrang übereinstimmt.

Der Rang einer linearen Abbildung ändert sich nicht, wenn wir die lineare Abbildung mit Isomorphismen verknüpfen, d.h. er lässt sich aus einer beliebigen Matrixdarstellung ablesen. Genauer gilt:

Satz 10.3.3. Seien $f : V \rightarrow W$ eine lineare Abbildung und $\varphi : W \rightarrow W'$ und $\psi : V' \rightarrow V$ Isomorphismen. Dann gilt:

$$\text{Rang } f = \text{Rang}(\varphi \circ f \circ \psi).$$

Sind insbesondere $B = (b_1, \dots, b_n)$ bzw. $C = (c_1, \dots, c_m)$ Basen von V bzw. W und $\phi_B : V \rightarrow K^n$ bzw. $\phi_C : W \rightarrow K^m$ die zugehörigen Koordinatensysteme, so gilt:

$$\text{Rang } f = \text{Rang}(\phi_C \circ f \circ \phi_B^{-1}).$$

Somit kann also der Rang von f mit Hilfe einer beliebigen Matrixdarstellung bestimmt werden.

Beweis. Da $\varphi : W \rightarrow W'$ und $\psi : V' \rightarrow V$ Isomorphismen sind, folgt

$$\dim \text{Bild}(\varphi \circ f) = \dim \text{Bild } f \quad \text{und} \quad \dim \text{Bild}(f \circ \psi) = \dim \text{Bild } f.$$

□

Ziel ist es, Koordinaten zu finden, für die eine lineare Abbildung eine besonders einfache Gestalt annimmt.

Satz 10.3.4. Sei $f : V \rightarrow W$ eine lineare Abbildung mit $\text{Rang } f = r$, $\dim V = n$ und $\dim W = m$. Dann gibt es Basen $B = (b_1, \dots, b_n)$ und $C = (c_1, \dots, c_m)$ von V bzw. W , so dass

$$A = M_{B,C}(f) = \left(\begin{array}{c|c} E_r & 0_{r \times (n-r)} \\ \hline & \\ \hline 0_{(m-r) \times r} & 0_{(m-r) \times (n-r)} \end{array} \right),$$

wobei E_r die Einheitsmatrix in $M(r, K)$ und $0_{i \times j}$ die Nullmatrix in $M(i \times j, K)$ bezeichnet.

Beweis. Sei $f : V \rightarrow W$ linear und $U \subset V$ ein Komplement zu $\text{Kern } f$, d.h. $U \oplus \text{Kern } f = V$. Dann folgt aus Satz 8.1.8, dass $f : U \rightarrow \text{Bild } f$ ein Isomorphismus ist. Sei (c_1, \dots, c_r) eine Basis von $\text{Bild } f$, so ist $f^{-1}(c_j) = b_j$ eine Basis von U . Da $\dim \text{Kern } f = n - r$, können wir eine Basis (b_{r+1}, \dots, b_n) von $\text{Kern } f$ auswählen. Dies definiert dann eine Basis $(b_1, \dots, b_r, b_{r+1}, \dots, b_n)$ von V . Man ergänze (c_1, \dots, c_r) zu einer Basis (c_1, \dots, c_m) von W . Dann folgt:

$$c_j = f(b_j) = \sum_{i=1}^m a_{ij} c_i$$

für $1 \leq j \leq r$. Daraus folgt:

$$a_{ij} = \delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases}$$

für $1 \leq j \leq r, 1 \leq i \leq m$. Ist $j \geq r$, so gilt:

$$0 = f(b_j) = \sum_{i=1}^m a_{ij} c_i.$$

Daraus folgt: $a_{ij} = 0$ für $r < j \leq n$ und $1 \leq i \leq n$. □

Dieser Satz hat folgende Anwendungen in der Theorie der Matrizen.

Korollar 10.3.5. Sei $A \in M(m \times n, K)$ mit $\text{Rang } A = r$. Dann existieren Matrizen $T \in GL(n, K)$ und $S \in GL(m, K)$ mit

$$S \cdot A \cdot T^{-1} = \left(\begin{array}{c|c} E_r & 0_{r \times (n-r)} \\ \hline 0_{(m-r) \times r} & 0_{(m-r) \times (n-r)} \end{array} \right).$$

Beweis. Sei $A : K^n \rightarrow K^m$ die lineare Abbildung, die der Matrix A entspricht. Dann gibt es wegen Satz 10.3.4 Basen B von K^n und C von K^m mit zugehörigen Koordinatensystemen $\phi_B : K^n \rightarrow K^n, \phi_C : K^m \rightarrow K^m$, so dass

$$\phi_C \circ A \circ \phi_B^{-1} = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Die Matrizen T bzw. S , die ϕ_B bzw. ϕ_C repräsentieren sind invertierbar, d.h. $T \in GL(n, K), S \in GL(m, K)$, und es gilt:

$$S \cdot A \cdot T^{-1} = \phi_C \circ A \circ \phi_B^{-1} = M_{B,C}(A).$$

□

Bemerkung. Aus diesem Korollar folgt: Sind $A, B \in M(m \times n, K)$, so gilt

$$\text{Rang } A = \text{Rang } B \Leftrightarrow \text{Es existieren } S \in GL(m, K), T \in GL(n, K) \text{ mit } B = S \cdot A \cdot T^{-1}.$$

Denn ist $\text{Rang } A = \text{Rang } B = r$, so existieren $S_1, S_2 \in GL(m, K)$ und $T_1, T_2 \in GL(n, K)$ mit

$$S_1 A T_1^{-1} = S_2 B T_2^{-1} = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Dies impliziert

$$B = S A T^{-1}$$

mit $S = S_2^{-1} S_1$ und $T = T_1^{-1} T_2$.

Definition 10.3.6. Sei $A \in M(m \times n, K)$, so heißt die Matrix $A^t \in M(n \times m, K)$ mit

$$(A^t)_{ij} = a_{ji} \text{ für } 1 \leq i \leq n, 1 \leq j \leq m$$

die *transponierte Matrix*, d.h. aus der j -ten Zeile von A wird die j -te Spalte von A^t .

Beispiel.

$$\begin{pmatrix} 2 & 3 & 1 & 0 \\ 0 & 4 & 3 & 2 \end{pmatrix}^t = \begin{pmatrix} 2 & 0 \\ 3 & 4 \\ 1 & 3 \\ 0 & 2 \end{pmatrix}.$$

Wir wollen nun untersuchen wie die Matrixdarstellung einer linearen Abbildung $f : V \rightarrow W$ zu ihrer transponierten Abbildung $f^t : W^* \rightarrow V^*$ in Beziehung steht.

Satz 10.3.7. Sei $F : V \rightarrow W$ eine lineare Abbildung und $f^t : W^* \rightarrow V^*$ die transponierte Abbildung. Seien $B = (b_1, \dots, b_n)$ bzw. $C = (c_1, \dots, c_m)$ Basen von V bzw. W und $B^* = (b_1^*, \dots, b_n^*)$ bzw. $C^* = (c_1^*, \dots, c_m^*)$ die dualen Basen von V^* bzw. W^* . Dann gilt:

$$M_{C^*, B^*}(f^t) = (M_{B, C}(f))^t.$$

Beweis. Sei $f^t(c_j^*) = c_j^* \circ f = \sum_{k=1}^n \tilde{a}_{kj} b_k^*$, so erhält man durch Anwendung auf den Vektor b_i :

$$\tilde{a}_{ij} = f^t(c_j^*)(b_i) = c_j^*(f(b_i)).$$

Nun gilt: $f(b_i) = \sum_{k=1}^m a_{ki} c_k$ und somit folgt aus der Definition der dualen Basis

$$\tilde{a}_{ij} = c_j^*(f(b_i)) = \sum_{k=1}^m a_{ki} c_j^*(c_k) = a_{ji}.$$

□

Satz 10.3.8. Es seien $A \in M(m \times n, K)$, $B \in M(n \times \ell, K)$, so gilt:

$$(A \cdot B)^t = B^t \cdot A^t.$$

Beweis. Sind $(A)_{ij} = a_{ij}$ und $(B)_{ij} = b_{ij}$, so gilt: $(A^t)_{ij} = a_{ji}$ und $(B^t)_{ij} = b_{ji}$. Ist $C = AB$, so folgt $c_{ij} = \sum_{k=1}^n a_{ik} b_{kj}$ und somit

$$(AB)^t_{ij} = c_{ji} = \sum_{k=1}^n a_{jk} b_{ki}.$$

Auf der anderen Seite ist

$$(B^t A^t)_{ij} = \sum_{k=1}^n (B^t)_{ik} (A^t)_{kj} = \sum_{k=1}^n b_{ki} a_{jk} = \sum_{k=1}^n a_{jk} b_{ki}.$$

□

Bemerkung. Ist $A \in GL(n, K)$, so folgt $A^t \in GL(n, K)$ mit $(A^t)^{-1} = (A^{-1})^t$, denn $E_n = E_n^t = (A \cdot A^{-1})^t = (A^{-1})^t \cdot A^t$.

Satz 10.3.9. Ist $A \in M(m \times n, K)$, so gilt

$$\text{Rang } A = \text{Rang } A^t .$$

Bemerkung. Da den Spalten von A^t die Zeilen von A entsprechen, folgt daraus:
Zeilenrang = Spaltenrang.

Beweis. Ist $\text{Rang } A = r$, so erhalten wir aus Korollar 10.3.5 die Existenz von Matrizen $T \in GL(n, K)$, $S \in GL(m, K)$ mit $S \cdot A \cdot T^{-1} = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$. Dann folgt aus Satz 10.3.8

$$\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} = (SAT^{-1})^t = (T^{-1})^t(SA)^t = (T^{-1})^t A^t S^t = (T^t)^{-1} A^t S^t .$$

10.4 Gaußscher Algorithmus

Nun werden wir ein Verfahren kennenlernen, um gewisse Standardprobleme in der linearen Algebra explizit zu lösen, wie z.B.:

- a) Lineare Gleichungssysteme;
- b) Berechnung des Rangs einer Matrix;
- c) Berechnung der inversen Matrix.

Sei zunächst V ein Vektorraum über K und

$$U = \text{Span}(b_1, \dots, b_n) = \{x_1 b_1 + \dots + x_n b_n \mid x_i \in K\}$$

der von den Vektoren $b_i \in V$ aufgespannte Unterraum. Wir wollen nun einige elementare Umformungen der Vektoren (b_1, \dots, b_n) angeben, ohne dass sich das Erzeugnis ändert. Wir bemerken:

Typ 1:

$$U = \text{Span}(b_1, \dots, b_j, \dots, b_i, \dots, b_n) \quad (\text{Vertauschung von } b_i \text{ mit } b_j).$$

Typ 2:

$$U = \text{Span}(b_1, \dots, b_{i-1}, b_i + \lambda b_j, b_{i+1}, \dots, b_n) \quad (\text{Addition des } \lambda - \text{fachen von } b_j \text{ zu } b_i),$$

für $\lambda \in K$ und $i \neq j$. Dabei ändert sich das Erzeugnis nicht, denn wegen $b_i = b_i + \lambda b_j - \lambda b_j$ und $i \neq j$ ist $b_i \in \text{Span}(b_1, \dots, b_{i-1}, b_i + \lambda b_j, b_{i+1}, \dots, b_n)$.

Typ 3:

$$U = \text{Span}(b_1, \dots, b_{i-1}, \lambda b_i, b_{i+1}, \dots, b_n) \quad (\text{Multiplikation von } b_i \text{ mit } \lambda),$$

für $\lambda \neq 0$.

Diese Umformungen werden wir nun anwenden, um den Zeilenrang einer Matrix zu berechnen.

Sei $A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \in M(m \times n, K)$. Der Zeilenrang von A ist gegeben durch $\dim \text{Span}(a_1, \dots, a_m)$ mit den Zeilen $a_i = (a_{i1}, \dots, a_{in}) \in K^n$.

Satz 10.4.1. Sei $A \in M(m \times n, K)$, so kann A durch elementare Zeilenumformungen vom Typ 1 und Typ 2 auf folgende Form (Zeilenstufenform)

$$\begin{pmatrix} 0 \dots 0 & b_{1j_1} \dots & & & \dots & b_{1n} \\ \vdots & \vdots & 0 \dots 0 & b_{2j_2} \dots & & \dots & b_{2n} \\ \vdots & \vdots & \vdots & \vdots & 0 \dots 0 & b_{3j_3} \dots & \dots & b_{3n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 0 \dots 0 & \dots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \dots & b_{rj_r} \dots & b_{rn} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 \dots 0 & 0 \dots 0 & 0 \dots 0 & 0 \dots 0 & & & 0 & \dots & 0 \end{pmatrix}$$

gebracht werden mit $1 \leq j_1 < j_2 < \dots < j_r \leq n$ und $b_{ij_i} \neq 0$.

Beweis. 1. Die erste Spalte der Matrix A , für die nicht alle Elemente null sind, hat den Index j_1 . Durch Vertauschen der Zeilen bringen wir ein solches, von null verschiedenes, Element in die erste Zeile.

$$A \xrightarrow{\text{Typ1}} \begin{pmatrix} 0 & \dots & 0 & b_{1j_1} & \dots & b_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & b_{mj_1} & \dots & b_{mn} \end{pmatrix} \text{ und } b_{1j_1} \neq 0.$$

2. Für $i = 2, \dots, m$ addieren wir zur i -ten Zeile das $\left(-\frac{b_{ij_1}}{b_{1j_1}}\right)$ -fache der ersten Zeile.

$$\xrightarrow{\text{Typ2}} \begin{pmatrix} 0 & \dots & 0 & b_{1j_1} & \dots & b_{1n} \\ \vdots & & \vdots & 0 & & \vdots \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & 0 & & \vdots \end{pmatrix} \cdot$$

Wir wenden 1 und 2 auf A' an und erhalten:

$$\begin{pmatrix} 0 & \dots & 0 & b_{1j_1} & \dots & & \dots & b_{1n} \\ \vdots & & \vdots & 0 & \dots & 0 & b_{2j_2} & \dots & b_{2n} \\ \vdots & & \vdots & \vdots & & \vdots & 0 & & \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & & A'' \\ 0 & & 0 & 0 & & 0 & 0 & & \end{pmatrix}.$$

Man gehe weiter gleichermaßen vor, bis man nach endlich vielen Schritten (offensichtlich weniger als die Spaltenzahl) zu einer Matrix der angegebenen Form gelangt. \square

Bemerkung. Dieses Verfahren nennt man auch “Gaußscher Algorithmus”.

Beispiel.

$$A = \begin{pmatrix} 0 & 0 & 0 & -4 & -3 \\ 0 & -4 & 2 & 2 & 3 \\ 0 & 6 & -3 & 1 & 1 \\ 0 & 2 & -1 & 1 & 0 \end{pmatrix} \xrightarrow{\text{Typ1}} \begin{pmatrix} 0 & 2 & -1 & 1 & 0 \\ 0 & -4 & 2 & 2 & 3 \\ 0 & 6 & -3 & 1 & 1 \\ 0 & 0 & 0 & -4 & -3 \end{pmatrix} \xrightarrow{\text{Typ2}} \begin{pmatrix} 0 & 2 & -1 & 1 & 0 \\ 0 & 0 & 0 & 4 & 3 \\ 0 & 0 & 0 & -2 & 1 \\ 0 & 0 & 0 & -4 & -3 \end{pmatrix} \xrightarrow{\text{Typ2}} \begin{pmatrix} 0 & 2 & -1 & 1 & 0 \\ 0 & 0 & 0 & 4 & 3 \\ 0 & 0 & 0 & 0 & 2,5 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Bemerkung. Die Zeilenstufenform ist nicht eindeutig.

Korollar 10.4.2. Sei $A \in M(m \times n, K)$ und sei $B \in M(m \times n, K)$ eine zugehörige Zeilenstufenform. Sei U der Vektorraum, der von den Zeilen von A aufgespannt wird, so gilt: Die von null verschiedenen Zeilen von B bilden eine Basis von U . Insbesondere gilt:

$$\text{Rang } A = \text{Anzahl der von null verschiedenen Zeilen von } B.$$

Beweis. Da die Zeilenvektoren von B durch elementare Umformungen der Zeilen von A entstanden sind, bilden sie ein Erzeugendensystem von U . Dass die von null verschiedenen Zeilen von B linear unabhängig in K^n sind, sieht man wie folgt: Sei

$$\alpha_1(0, \dots, 0, b_{1j_1}, \dots, b_{1j_n}) + \alpha_2(0, \dots, 0, b_{2j_2}, \dots, b_{2j_n}) + \dots + \alpha_r(0, \dots, 0, b_{rj_r}, \dots, b_{rj_n}) = (0, \dots, 0),$$

so folgt aus $j_1 < j_2 < \dots < j_r$ und $b_{ij_i} \neq 0$, $\alpha_1 = \alpha_2 = \dots = \alpha_r = 0$. \square

Wann ist eine $n \times n$ Matrix A invertierbar? Es gilt zum Beispiel

$$A \text{ invertierbar} \Leftrightarrow \text{Rang } A = \dim \text{Bild } A = n,$$

denn wegen Korollar 8.2.8 ist $A : K^n \rightarrow K^n$ surjektiv genau dann, wenn A bijektiv ist.

Korollar 10.4.3. $A \in M(n, K)$ ist invertierbar \Leftrightarrow ihre Zeilenstufenform hat die Gestalt

$$\begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & \ddots & \vdots \\ 0 & \dots & b_{nn} \end{pmatrix}.$$

Bemerkung. Die Zeilenstufenform impliziert natürlich $b_{ii} \neq 0$.

Beweis. A invertierbar \Leftrightarrow Rang $A =$ Zeilenrang $= n = r \Leftrightarrow$

$$B = \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & \ddots & \vdots \\ 0 & \dots & b_{nn} \end{pmatrix}, \text{ denn aus } 1 \leq j_1 < j_2 < \dots < j_n \leq n \text{ folgt:}$$

$j_1 = 1, j_2 = 2, \dots, j_n = n$. Dabei ist $b_{ii} \neq 0$. □

Bemerkung. Korollar 10.4.2 kann man z.B. benutzen, um folgende Aufgabe zu lösen. Es seien $v_1, \dots, v_m \in K^n$ und $U = \text{Span}\{v_1, \dots, v_m\}$. Gesucht wird eine Basis von U . Die Lösung erhält man wie folgt:

Man schreibe $v_i = (a_{i1}, \dots, a_{in}) \in K^n$ und betrachte diese Vektoren als Zeilen der Matrix A . Diese Matrix bringe man auf Zeilenstufenform. Die von null verschiedenen Zeilen bilden dann eine Basis von U .

Allgemeiner gilt: Sei V ein beliebiger Vektorraum und $B = (b_1, \dots, b_n)$ eine Basis von V und $\phi_B : V \rightarrow K^n$ zugehöriger Koordinatenisomorphismus. Sei $U = \text{Span}\{v_1, \dots, v_m\}$ ein Unterraum von V , so finden wir eine Basis in U wie folgt. Betrachte den Unterraum $\phi_B(U) = \text{Span}\{\phi_B(v_1), \dots, \phi_B(v_m)\} \subset K^n$. Nach obigem Verfahren können wir eine Basis B' von $\phi_B(U)$ gewinnen. Dann folgt: $\phi_B^{-1}(B')$ ist eine Basis von U (denn ϕ_B ist ein Isomorphismus).

Wir wollen nun die oben angegebenen elementaren Umformungen durch Matrizen (Element-

armatrizen) beschreiben. Sei $A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \in M(m \times n, K)$ eine $m \times n$ Matrix. Ist $P \in M(m, K)$, so ist $P \cdot A = \left(P \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix}, \dots, P \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} \right)$, denn $PA(e_j) = P \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$.

Sei $i \neq j$, so entspricht der Vertauschung der i -ten und der j -ten Zeile (Typ 1) eine Matrix

Lemma 10.4.4. Die obigen Elementarmatrizen sind invertierbar und es gilt:

$$i) (P^{i,j})^{-1} = P^{i,j}.$$

$$ii) (Q^{i,j}(\lambda))^{-1} = Q^{i,j}(-\lambda).$$

$$iii) S^i(\lambda)^{-1} = S^i(\lambda^{-1}) \text{ für } \lambda \neq 0.$$

Es stellt sich nun heraus, dass sich jede invertierbare Matrix als endliches Produkt der Elementarmatrizen darstellen lässt.

Satz 10.4.5. Jede Matrix $A \in GL(n, K)$ ist endliches Produkt von Elementarmatrizen.

Bemerkung. Man sagt: Die Gruppe $GL(n, K)$ wird durch die Elementarmatrizen erzeugt.

Beweis. Wegen Korollar 10.4.3 hat die Zeilenstufenform von A die Gestalt

$$B = \begin{pmatrix} b_{11} & & b_{1n} \\ & \ddots & \\ 0 & & b_{nn} \end{pmatrix},$$

d.h. $B = B_k \dots B_1 A$ mit B_i Elementarmatrizen vom Typ 1 bzw. 2.

Wenden wir weitere Zeilenumformungen an, so können wir zunächst erreichen, dass alle Koeffizienten in der n -ten Spalte oberhalb b_{nn} verschwinden (durch Typ 1 bzw. Typ 2 Umformungen). Durch weitere Zeilenumformungen erhalten wir, dass alle Koeffizienten oberhalb der Diagonalen verschwinden. Das Ergebnis ist eine Diagonalmatrix

$$D = \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix} = B_s \dots B_{r+1} B_r \dots B_1 A$$

mit $d_i \neq 0$. Nun normieren wir die Diagonale noch zu 1, indem wir die i -ten Zeilen mit $\frac{1}{d_i}$ multiplizieren, d.h. wir multiplizieren D mit den Matrizen $S^i(\frac{1}{d_i})$. Dies impliziert $E_n = B_\ell \dots B_1 A$, wobei B_i Elementarmatrizen darstellen. Da die Inversen der Elementarmatrizen wieder Elementarmatrizen sind, gilt: $A = B_1^{-1} \dots B_\ell^{-1}$ ist Produkt von Elementarmatrizen. \square

Dieser Satz ergibt ein praktisches Verfahren, eine Matrix zu invertieren. Sei $A \in M(n \times n, K)$, so bringe A mit Elementarmatrizen B_1, \dots, B_k auf Zeilenstufenform, d.h. Zeilenstufenform $= B_k \dots B_1 A$. Ist der Rang $A < n$, so ist A nicht invertierbar und wir können die Rechnung beenden. Ist Rang $A = n$, so führe man weitere Zeilenumformungen durch, bis aus A die Einheitsmatrix E_n wird. Schematisch lässt sich dies wie folgt durchführen:

$$\begin{array}{ccc|ccc} & & A & & & E_n \\ & & B_1 A & & & B_1 E_n \\ & & \vdots & & & \vdots \\ E_n = B_\ell \dots B_1 A & & & & A^{-1} = B_\ell \dots B_1 E_n & \end{array}$$

d.h. man führt simultan Zeilenumformungen von A und E_n durch, bis links E_n steht. Dann erscheint auf der rechten Seite die zu A inverse Matrix.

Beispiel. Berechnung der zu $A = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} \in M(2 \times 2, \mathbb{R})$ inversen Matrix.

$$\left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 1 & 0 & 1 \end{array} \right) \mapsto \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & -5 & -3 & 1 \end{array} \right) \mapsto \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 1 & \frac{3}{5} & -\frac{1}{5} \end{array} \right) \mapsto \left(\begin{array}{cc|cc} 1 & 0 & -\frac{1}{5} & \frac{2}{5} \\ 0 & 1 & \frac{3}{5} & -\frac{1}{5} \end{array} \right).$$

Also ist $\begin{pmatrix} -\frac{1}{5} & \frac{2}{5} \\ \frac{3}{5} & -\frac{1}{5} \end{pmatrix}$ die gesuchte inverse Matrix zu A .

10.5 Lineare Gleichungssysteme

Nun wollen wir uns mit den in Abschnitt 8.1 (Beispiel 2) eingeführten linearen Gleichungssystemen beschäftigen. Ist $A \in M(m \times n, K)$ und $b \in K^m$, so definiert die lineare Gleichung $Ax = b$ ein System von m skalaren Gleichungen mit n Unbekannten, nämlich:

$$\begin{aligned} a_{11}x_1 + \dots + a_{1n}x_n &= b_1 \\ \vdots & \\ a_{m1}x_1 + \dots + a_{mn}x_n &= b_m. \end{aligned}$$

Nach Definition von Bild A hat $Ax = b$ genau dann eine Lösung, falls $b \in \text{Bild } A$. Ist dies der Fall, so gilt nach Satz 8.1.10

$$\mathcal{L}_{A,b} = \{x \in K^n \mid Ax = b\} = x_0 + \text{Kern } A := \{x_0 + y \mid y \in \text{Kern } A\},$$

wobei x_0 eine spezielle Lösung der Gleichung darstellt.

Wir erinnern nochmals daran, dass für $b = 0$ das Gleichungssystem homogen, andernfalls inhomogen genannt wird. Die Gesamtheit der Lösungen besteht daher aus der Summe einer speziellen Lösung der inhomogenen Gleichung und der Menge der Lösungen der homogenen Gleichung.

Bemerkung. Der Lösungsraum eines linearen Gleichungssystems ändert sich natürlich nicht, falls man das Gleichungssystem mit einer invertierbaren Matrix multipliziert. Genauer: Ist $A \in M(m \times n, K)$, $b \in K^m$, so gilt: Ist $S \in GL(m, K)$, so folgt:

$$\mathcal{L}_{A,b} = \mathcal{L}_{S^{-1}A, S^{-1}b},$$

denn die Gleichung $Ax = b$ ist äquivalent zu $S^{-1}Ax = S^{-1}b$.

Wir wollen uns jetzt damit beschäftigen, wie man die Gesamtheit der Lösungen berechnen kann. Dabei ist die Anwendung des Gaußalgorithmus und die damit verbundene Zeilenstufenform einer Matrix von entscheidender Bedeutung.

Satz 10.5.1. Sei $A \in M(m \times n, K)$ mit $\text{Rang } A = r$ und $b \in K^m$. Es sei $S \in GL(m, K)$, so dass $S^{-1}A = B$ eine Zeilenstufenform von A darstellt. Ist $b' = S^{-1}b$, so besitzt die lineare Gleichung

$$Ax = b$$

genau dann eine Lösung, falls $b'_i = 0$ für $i > r$.

Beweis. Sei also $S \in GL(m, K)$, so dass $S^{-1}A = B$ eine Zeilenstufenform von A darstellt und es sei $b' = S^{-1}b$. Da der Rang von A gleich r ist, hat die Zeilenstufenform B von A genau r von null verschiedene Zeilen. Wegen der obigen Bemerkung stimmen die Lösungsmengen der Gleichungen $Ax = b$ und $Bx = b'$ überein. Ist also x eine Lösung von $Bx = b'$, d.h.

$$B(x) = \begin{pmatrix} 0 \dots 0 & b_{1j_1} \dots & & & \dots & b_{1n} \\ \vdots & \vdots & 0 \dots 0 & b_{2j_2} \dots & & \dots & b_{2n} \\ \vdots & \vdots & \vdots & \vdots & 0 \dots 0 & b_{3j_3} \dots & \dots & b_{3n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 0 \dots 0 & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & b_{rj_r} \dots & b_{rn} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 0 & \dots & 0 \\ 0 \dots 0 & 0 \dots 0 & 0 \dots 0 & 0 \dots 0 & \dots & \vdots & \dots & \vdots \\ 0 \dots 0 & 0 \dots 0 & 0 \dots 0 & 0 \dots 0 & \dots & 0 & \dots & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b'_1 \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ b'_r \\ b'_{r+1} \\ \vdots \\ b'_m \end{pmatrix},$$

so folgt: $b'_{r+1} = \dots = b'_m = 0$.

Sei nun umgekehrt $b'_{r+1} = \dots = b'_m = 0$. Wir zeigen, dass dann eine Lösung existiert. Wir werden sogar etwas mehr zeigen, nämlich dass die Lösung eindeutig ist, wenn man als Lösungen nur Vektoren in einem gewissen Unterraum von K^n zulässt. Man betrachte dazu die Indexmenge $J = \{j_1, \dots, j_r\}$ und den Unterraum

$$\begin{aligned} U_J &:= \{x \in K^n \mid x_i = 0, \text{ falls } i \notin J\} \\ &= \{(0, \dots, x_{j_1}, 0, \dots, 0, x_{j_2}, 0, \dots, 0, x_{j_r}, 0, \dots, 0) \mid x_{j_i} \in K\}. \end{aligned}$$

Dann ist für $x \in U_J$ und $b' = \{b'_1, \dots, b'_r, 0, \dots, 0\}$ das Gleichungssystem $B(x) = b'$ äquivalent zu

$$B(x) = \begin{pmatrix} b_{1j_1}x_{j_1} + & b_{1j_2}x_{j_2} + & \dots & +b_{1j_r}x_{j_r} \\ & b_{2j_2}x_{j_2} + & \dots & +b_{2j_r}x_{j_r} \\ & & \vdots & \\ & & & b_{rj_r}x_{j_r} \\ & & & 0 \\ & & & \vdots \\ & & & 0 \end{pmatrix} = \begin{pmatrix} b'_1 \\ b'_2 \\ \vdots \\ b'_r \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Daher existiert in U_J genau eine Lösung

$$x = \{(0, \dots, x_{j_1}, 0, \dots, 0, x_{j_2}, 0, \dots, 0, x_{j_r}, 0, \dots, 0)\}$$

von $Bx = b'$. Die von null verschiedenen Komponenten x_{j_1}, \dots, x_{j_r} berechnen sich wie folgt: Aus der r -ten Zeile erhalten wir:

$$b_{rj_r}x_{j_r} = b'_r \text{ und somit } x_{j_r} = b'_r/b_{rj_r}.$$

Aus der $(r - 1)$ -ten Zeile erhalten wir:

$$b_{r-1j_{r-1}}x_{j_{r-1}} + b_{r-1j_r}x_{j_r} = b'_{r-1}.$$

Daraus berechnet sich $x_{j_{r-1}}$. Setzen wir dieses Verfahren fort, so erhalten wir schließlich alle restlichen Komponenten x_{j_1}, \dots, x_{j_r} und somit eine eindeutige Lösung $x \in U_J \subset K^n$ der Gleichung $Bx = b'$. Diese ist dann eine Lösung der Gleichung $Ax = b$. \square

Nun können wir leicht ein Verfahren angeben, um die Gesamtheit der Lösungen zu bestimmen. Sei B die oben beschriebene Zeilenstufenform von A und U_J der wie oben konstruierte Unterraum. Ist $K_m^r := \{(q_1, \dots, q_r, 0, \dots, 0) \in K^m\}$, so folgt aus den obigen Überlegungen die eindeutige Lösbarkeit der Gleichung $B(x) = q$ zu gegebenem $q \in K_m^r$ und gesuchtem $x \in U_J$. Dies bedeutet aber, dass die auf U_J eingeschränkte Abbildung

$$B_J := B|_{U_J} : U_J \rightarrow K_m^r$$

mit $B|_{U_J}(x) = B(x)$ bijektiv ist.

Um nun alle Lösungen des homogenes Problems zu erhalten, betrachte man den Unterraum

$$U_{\bar{J}} = \{y \in K^n \mid y_i = 0, i \in J\} \subset K^n.$$

$U_{\bar{J}}$ ist in K^n ein Komplement des r -dimensionalen Unterraumes U_J und hat somit die Dimension $n - r$. Wir bestimmen zu jedem vorgegeben $y \in U_{\bar{J}}$ ein $x \in U_J$ mit

$$B(x + y) = 0.$$

Diese Gleichung ist äquivalent zu der Gleichung

$$B(x) = -B(y).$$

Aus der Gestalt von B folgt $B(y) \in K_m^r$. Da $B_J : U_J \rightarrow K_m^r$ bijektiv ist, hat die Gleichung $B(x) = -B(y)$ die eindeutig bestimmte Lösung

$$x = -B_J^{-1}B(y) \in U_J.$$

Analog zur Vorgehensweise bei der Berechnung der speziellen Lösung, berechnet man die von null verschiedenen Komponenten x_{j_1}, \dots, x_{j_r} von x , indem man zu jedem $y \in U_{\bar{J}}$ das Gleichungssystem

$$B(x) = \begin{pmatrix} b_{1j_1}x_{j_1} + b_{1j_2}x_{j_2} + \dots + b_{1j_r}x_{j_r} \\ b_{2j_2}x_{j_2} + \dots + b_{2j_r}x_{j_r} \\ \vdots \\ b_{rj_r}x_{j_r} \\ 0 \\ \vdots \\ 0 \end{pmatrix} = -B(y)$$

löst. Dann definiert $G : U_{\bar{J}} \rightarrow \text{Kern } B$ mit

$$G(y) = y + x = y - B_{\bar{J}}^{-1}B(y)$$

eine lineare Abbildung. Sie ist injektiv, denn ist $G(y) = y - B_{\bar{J}}^{-1}B(y) = 0$, so ist $y = 0$ und $B_{\bar{J}}^{-1}B(y) = 0$. Dies folgt, da die Summe $U_J + U_{\bar{J}}$ direkt ist und $B_{\bar{J}}^{-1}B(y) \in U_J$ bzw. $y \in U_{\bar{J}}$. Wegen $\dim \text{Kern } B = \dim U_{\bar{J}} = n - r$ ist die Abbildung auch bijektiv. Also sind die Lösungen des homogenen Problems gegeben durch

$$\text{Kern } A = \text{Kern } B = \{y - B_{\bar{J}}^{-1}B(y) \mid y \in U_{\bar{J}}\}.$$

Alle Lösungen sind somit durch

$$\text{Kern } B + B_{\bar{J}}^{-1}(b') = \{y - B_{\bar{J}}^{-1}B(y) + B_{\bar{J}}^{-1}(b') \mid y \in U_{\bar{J}}\} = \{y + B_{\bar{J}}^{-1}(-B(y) + b') \mid y \in U_{\bar{J}}\}$$

gegeben.

In der Praxis berechnet man die allgemeine Lösung direkt und nicht getrennt nach spezieller und allgemeiner Lösung, denn in beiden Fällen sind die Rechenschritte sehr ähnlich. Man gehe wie folgt vor. Nachdem man das Gleichungssystem auf Zeilenstufenform gebracht hat, überprüfe man, wie oben beschrieben, ob eine Lösung existiert. Dann bestimmt man aus der Zeilenstufenform die Indexmenge $J = \{j_1, \dots, j_r\}$ und ihr Komplement $\bar{J} = \{1, \dots, n\} \setminus J$. Nun müssen wir zu gegeben $y \in U_{\bar{J}}$ den eindeutig bestimmten Vektor $x \in U_J$ bestimmen mit

$$x = x(y) := B_J^{-1}(-B(y) + b').$$

Dies ist natürlich zur Lösung des Gleichungssystems

$$B(x) = -B(y) + b'$$

mit $x \in U_J$ äquivalent. Die Lösungsmenge $\mathcal{L}_{A,b} = \mathcal{L}_{B,b'}$ ist dann von der Form

$$\{y + x(y) \mid y \in U_{\bar{J}}\}.$$

Beispiel. Man betrachte das reelle Gleichungssystem

$$\begin{pmatrix} 1 & 0 & 6 & 1 \\ 0 & 1 & -8 & 0 \\ 0 & 4 & -16 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 3 \\ -5 \\ -7 \end{pmatrix}.$$

Multiplizieren wir die zweite Zeile mit -4 und addieren wir das Ergebnis zur dritten Zeile, so ist die Matrix auf Zeilenstufenform gebracht und wir erhalten das äquivalente Gleichungssystem

$$\begin{pmatrix} 1 & 0 & 6 & 1 \\ 0 & 1 & -8 & 0 \\ 0 & 0 & 16 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 3 \\ -5 \\ 13 \end{pmatrix}.$$

Damit ist $J = \{1, 2, 3\}$, $\bar{J} = \{4\}$ und somit

$$U_J = \{(x_1, x_2, x_3, 0) \mid x_j \in \mathbb{R}\} \text{ und } U_{\bar{J}} = \{(0, 0, 0, y) \mid y \in \mathbb{R}\}.$$

Zu gegebenem $\tilde{y} = (0, 0, 0, y) \in U_{\bar{J}}$ und $b' = (3, -5, 13)$ müssen wir nun das Gleichungssystem

$$B(x) = -B(\tilde{y}) + b'$$

mit $x = (x_1, x_2, x_3, 0)$ lösen. Wir erhalten die Gleichungen

$$\begin{aligned} x_1 + 6x_3 &= 3 - y \\ x_2 - 8x_3 &= -5 \\ 16x_3 &= 13 + 3y. \end{aligned}$$

Daraus ergibt sich:

$$\begin{aligned} x_3 &= \frac{13}{16} + \frac{3}{16}y \\ x_2 &= -5 + \frac{13}{2} + \frac{3}{2}y \\ &= \frac{3}{2} + \frac{3}{2}y \\ x_1 &= 3 - \frac{6 \cdot 13}{16} - \frac{18}{16}y - y = \frac{24 - 39}{8} - \frac{17}{8}y \\ &= -\frac{15}{8} - \frac{17}{8}y. \end{aligned}$$

Also ist die allgemeine Lösung von der Form

$$\left\{ \left(-\frac{15}{8} - \frac{17}{8}y, \frac{3}{2} + \frac{3}{2}y, \frac{13}{16} + \frac{3}{16}y, y \right) \mid y \in \mathbb{R} \right\} = \left\{ \left(-\frac{15}{8}, \frac{3}{2}, \frac{13}{16}, 0 \right) + y \left(-\frac{17}{8}, \frac{3}{2}, \frac{3}{16}, 1 \right) \mid y \in \mathbb{R} \right\}.$$

Dabei ist

$$\left(-\frac{15}{8}, \frac{3}{2}, \frac{13}{16}, 0 \right)$$

eine spezielle Lösung der inhomogenen Gleichung und

$$\left\{ y \left(-\frac{17}{8}, \frac{3}{2}, \frac{3}{16}, 1 \right) \mid y \in \mathbb{R} \right\}$$

die Menge aller Lösungen der homogenen Gleichung.

Kapitel 11

Determinanten

11.1 Die definierenden Eigenschaften der Determinante

Definition 11.1.1. Sei K ein Körper. Die *Determinante* ist eine Abbildung

$$\det : \underbrace{K^n \times \dots \times K^n}_{n\text{-mal}} \rightarrow K$$

mit folgenden Eigenschaften.

1. \det ist *multilinear*, d.h. linear in jeder Komponente. Ist $j \in \{1, \dots, n\}$ und halten wir die von der j -ten Komponente verschiedenen Komponenten fest, so gilt:

$$\det(\dots, \lambda a_j + \mu b_j, \dots) = \lambda \det(\dots, a_j, \dots) + \mu \det(\dots, b_j, \dots)$$

für $a_j, b_j \in K^n$ und $\lambda, \mu \in K$.

2. \det ist *alternierend*, d.h.

$$\det(a_1, \dots, a_n) = 0,$$

falls zwei Vektoren übereinstimmen, d.h. falls $a_i = a_j$ für $i \neq j$.

3. \det ist *normiert*, d.h.

$$\det(e_1, \dots, e_n) = 1,$$

falls e_1, \dots, e_n die kanonische Basis von K^n ist.

Bemerkung. Aus 1. und 2. folgt:

$$\det(\dots, a_i, \dots, a_j, \dots) = -\det(\dots, a_j, \dots, a_i, \dots),$$

denn

$$\begin{aligned} 0 &= \det(\dots, a_i + a_j, \dots, a_i + a_j, \dots) \\ &= \det(\dots, a_i, \dots, a_i + a_j, \dots) + \det(\dots, a_j, \dots, a_i + a_j, \dots) \\ &= \det(\dots, a_i, \dots, a_j, \dots) + \det(\dots, a_j, \dots, a_i, \dots). \end{aligned}$$

Wir werden sehen, dass es genau eine Abbildung mit den obigen Eigenschaften gibt. Um dies zu beweisen, ist die *Permutationsgruppe* S_n von großer Bedeutung. Diese besteht aus den bijektiven Abbildungen:

$$S_n = \{\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\} \mid \sigma \text{ bijektiv}\}.$$

Die Gruppenstruktur ist durch die Verknüpfung von Abbildungen gegeben. Man schreibt auch $\sigma \in S_n$ oft in Form einer Doppelzeile:

$$\sigma = \begin{pmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{pmatrix}.$$

Satz 11.1.2. Die Gruppe S_n ist endlich mit $|S_n| = n! = n \cdot (n-1) \cdot \dots \cdot 1$.

Beweis. Wir beweisen dies durch Induktion über n :

Ist $n = 1$, so ist $|S(1)| = 1$.

Der Satz sei für $n-1$ bewiesen, d.h. es gelte $|S_{(n-1)}| = (n-1)!$.

Sei $\sigma \in S_n$ und $\sigma(1) = i \in \{1, \dots, n\}$. Dann gibt es nach Induktionsvoraussetzung genau $(n-1)!$ bijektive Abbildungen

$$\sigma : \{2, \dots, n\} \rightarrow \{1, \dots, n\} \setminus \{i\},$$

d.h. genau $(n-1)!$ Möglichkeiten, σ zu einer bijektiven Abbildung auf $\{1, \dots, n\}$ fortzusetzen. Daraus folgt: $|S_n| = n(n-1)! = n!$. \square

Die einfachsten Permutationen sind solche, die nur zwei Elemente miteinander vertauschen.

Bemerkung. Für $n \geq 3$ ist die Gruppe S_n nicht abelsch.

Definition 11.1.3. Eine Permutation $\tau_{i,j}$

$$\tau_{i,j} = \begin{pmatrix} 1 & \dots & i & \dots & j & \dots & n \\ 1 & \dots & j & \dots & i & \dots & n \end{pmatrix},$$

die nur i mit j vertauscht und alle anderen Elemente festhält, heißt *Transposition*.

Satz 11.1.4. Die Gruppe S_n wird von Transpositionen erzeugt. Genauer gilt: Jedes $\sigma \in S_n$ lässt sich als Produkt von höchstens $(n-1)$ Transpositionen schreiben.

Beweis. Sei $\sigma \in S_n$. Wir zeigen: Es gibt $(n-1)$ Transpositionen $\tau_1, \dots, \tau_{n-1}$ mit

$$\tau_{n-1} \circ \dots \circ \tau_2 \circ \tau_1 \circ \sigma = \text{id}.$$

Da die Inverse einer Transposition wieder eine Transposition ist, folgt damit die Behauptung. Sei $\tau_1 = \tau_{\sigma(1),1}$ die Transposition, die $\sigma(1)$ mit 1 vertauscht. Dann gilt für $\sigma_1 = \tau_1 \circ \sigma$

$$\sigma_1(1) = \tau_1(\sigma(1)) = 1.$$

Sei nun $\tau_2 = \tau_{\sigma_1(2),2}$ die Transposition, die $\sigma_1(2)$ mit 2 vertauscht, so gilt für $\sigma_2 = \tau_2 \circ \sigma_1 = \tau_2 \circ \tau_1 \circ \sigma$

$$\sigma_2(1) = 1 \quad \text{und} \quad \sigma_2(2) = 2.$$

Seien nun $\tau_1, \dots, \tau_{k-1}$ schon definiert und σ_{k-1} gegeben durch $\sigma_{k-1} = \tau_{k-1} \circ \dots \circ \tau_1 \circ \sigma$, so dass

$$\sigma_{k-1}(j) = j \quad \text{für} \quad 1 \leq j \leq k-1.$$

Dann definiere man $\tau_k = \tau_{\sigma_{k-1}(k),k}$ und es gilt für $\sigma_k = \tau_k \circ \sigma_{k-1}$

$$\sigma_k(j) = j \quad \text{für} \quad 1 \leq j \leq k.$$

Nach $(n - 1)$ Schritten erhalten wir:

$$\sigma_{n-1} = \tau_{n-1} \circ \dots \circ \tau_1 \circ \sigma$$

und es gilt

$$\sigma_{n-1}(j) = j \quad \text{für } 1 \leq j \leq n - 1.$$

Dann ist auch $\sigma_{n-1}(n) = n$ und somit $\sigma_{n-1} = \text{id}$.

Natürlich können einige der konstruierten Transpositionen τ_i gleich der Identität sein. In dem Falle ist die Anzahl der benötigten Transpositionen kleiner als $n - 1$. \square

Definition 11.1.5. Sei $\sigma \in S_n$ und $\sigma = \tau_k \circ \dots \circ \tau_1$ mit τ_i Transposition. Dann definiere man das Signum von σ durch

$$\text{sign}(\sigma) = (-1)^k .$$

Bemerkungen.

1. Die Darstellung einer Permutation durch Transpositionen ist natürlich nicht eindeutig. Man kann aber zeigen: Hat man zwei Darstellungen mit k und k' Transpositionen, so ist die Differenz $k - k'$ eine gerade Zahl, also von der Form $2m$ mit $m \in \mathbb{Z}$. Damit ist das Signum einer Permutation wohldefiniert.
2. Eine Permutation heißt *gerade*, wenn das Signum der Permutation 1 ist. Ist es -1 , so heißt die Permutation *ungerade*. Die Menge der geraden Permutationen in S_n wird mit A_n bezeichnet. Sie bilden eine Untergruppe von S_n . Hingegen bilden die ungeraden Permutationen keine Untergruppe von S_n .

Lemma 11.1.6. Ist $\det : K^n \times \dots \times K^n \rightarrow K$ eine Determinantenabbildung und e_1, \dots, e_n kanonische Basis von K^n , so gilt:

$$\text{sign}(\sigma) = \det(e_{\sigma(1)}, \dots, e_{\sigma(n)}) .$$

Beweis. Ist $\sigma = \tau_k \circ \dots \circ \tau_1 = \tau_k \circ \sigma_{k-1}$ mit Transpositionen τ_i , so gilt:

$$\begin{aligned} \det(e_{\sigma(1)}, \dots, e_{\sigma(n)}) &= \det(e_{\tau_k \sigma_{k-1}(1)}, \dots, e_{\tau_k \sigma_{k-1}(n)}) \\ &= -\det(e_{\sigma_{k-1}(1)}, \dots, e_{\sigma_{k-1}(n)}) \\ &\quad \vdots \qquad \qquad \qquad \vdots \\ &= (-1)^k \det(e_1, \dots, e_n) = \text{sign}(\sigma) . \end{aligned}$$

\square

Satz 11.1.7. (Eindeutigkeit und Existenz der Determinante)

Die in Definition 11.1.1 beschriebenen Eigenschaften charakterisieren die Determinante eindeutig. Sind insbesondere n Vektoren $a_i = (a_{i1}, \dots, a_{in}) \in K^n$ gegeben, so folgt

$$\det(a_1, \dots, a_n) = \sum_{\sigma \in S_n} \text{sign}(\sigma) a_{1\sigma(1)} \cdot \dots \cdot a_{n\sigma(n)} .$$

Diese Formel liefert auch eine Definition der Determinante.

Beweis. Ist $a_i = a_{i1}e_1 + \dots + a_{in}e_n$, so folgt:

$$\begin{aligned}
 \det(a_1, \dots, a_n) &= \sum_{j_1=1}^n a_{1j_1} \det(e_{j_1}, a_2, \dots, a_n) \\
 &= \sum_{j_1=1}^n a_{1j_1} \left(\sum_{j_2=1}^n a_{2j_2} \det(e_{j_1}, e_{j_2}, a_3, \dots, a_n) \right) \\
 &= \sum_{j_1=1}^n a_{1j_1} \left(\sum_{j_2=1}^n a_{2j_2} \left(\sum_{j_3=1}^n a_{3j_3} \dots \sum_{j_n=1}^n a_{nj_n} \det(e_{j_1}, \dots, e_{j_n}) \right) \dots \right) \\
 &= \sum_{j_1, \dots, j_n=1}^n a_{1j_1} a_{2j_2} \dots a_{nj_n} \det(e_{j_1}, \dots, e_{j_n}) \\
 &= \sum_{\sigma \in S_n} a_{1\sigma(1)} \dots a_{n\sigma(n)} \det(e_{\sigma(1)}, \dots, e_{\sigma(n)}) \\
 &= \sum_{\sigma \in S_n} \text{sign}(\sigma) a_{1\sigma(1)} \dots a_{n\sigma(n)}.
 \end{aligned}$$

Um zu zeigen, dass diese Formel auch eine Determinante darstellt, bleibt noch zu zeigen, dass sie die charakterisierenden Eigenschaften einer Determinante aus Definition 11.1.1 auch wirklich erfüllt. Es ist also nachzuprüfen, dass diese Formel eine multilineare, alternierende, sowie normierte Abbildung definiert.

1. Multilinearität:

$$\begin{aligned}
 &\det(a_1, \dots, a_{j-1}, \lambda a_j + \mu a'_j, a_{j+1}, \dots, a_n) \\
 &= \sum_{\sigma \in S_n} \text{sign}(\sigma) a_{1\sigma(1)} \dots (\lambda a_{j\sigma(j)} + \mu a'_{j\sigma(j)}) \dots a_{n\sigma(n)} \\
 &= \lambda \sum_{\sigma \in S_n} \text{sign}(\sigma) a_{1\sigma(1)} \dots a_{j\sigma(j)} \dots a_{n\sigma(n)} + \mu \sum_{\sigma \in S_n} \text{sign}(\sigma) a_{1\sigma(1)} \dots a_{j-1\sigma(j-1)} a'_{j\sigma(j)} \dots a_{n\sigma(n)} \\
 &= \lambda \det(a_1, \dots, a_n) + \mu \det(a_1, \dots, a_{j-1}, a'_j, a_{j+1}, \dots, a_n).
 \end{aligned}$$

2. Alternierend:

Es sei $a_i = (a_{i1}, \dots, a_{in}) = (a_{j1}, \dots, a_{jn}) = a_j$ mit $i \neq j$ und $\tau = \tau_{ij}$ Transposition, die i mit j vertauscht. Es gilt: $S_n = A_n \cup A_n \tau$ mit $A_n \cap A_n \tau = \emptyset$, wobei $A_n \tau = \{\sigma \circ \tau \mid \sigma \in A_n\}$.

$$\begin{aligned}
 \det(a_1, \dots, a_n) &= \sum_{\sigma \in A_n} \text{sign}(\sigma) a_{1\sigma(1)} \dots a_{i\sigma(i)} \dots a_{j\sigma(j)} \dots a_{n\sigma(n)} \\
 &\quad + \sum_{\sigma \in A_n} \text{sign}(\sigma \circ \tau) a_{1\sigma\tau(1)} \dots a_{i\sigma\tau(i)} \dots a_{j\sigma\tau(j)} \dots a_{n\sigma\tau(n)} \\
 &= \sum_{\sigma \in A_n} a_{1\sigma(1)} \dots a_{i\sigma(i)} \dots a_{j\sigma(j)} \dots a_{n\sigma(n)} \\
 &\quad - \sum_{\sigma \in A_n} a_{1\sigma(1)} \dots a_{i\sigma(j)} \dots a_{j\sigma(i)} \dots a_{n\sigma(n)} = 0,
 \end{aligned}$$

denn $a_{i\sigma(j)} = a_{j\sigma(j)}$ und $a_{j\sigma(i)} = a_{i\sigma(i)}$.

3. Normiertheit:

Sei $a_i = (0, \dots, 1, \dots, 0) = e_i$, so folgt $a_{ij} = \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$ und wir erhalten

$$a_{1\sigma(1)} \cdot \dots \cdot a_{i\sigma(i)} \cdot \dots \cdot a_{n\sigma(n)} = \begin{cases} 0, & \sigma \neq \text{id}, \\ 1, & \sigma = \text{id}. \end{cases}$$

Also ist $\det(e_1, \dots, e_n) = 1$.

□

Beispiel. Die Formel ergibt für $\det : K^2 \times K^2 \rightarrow K$

$$\det(a_1, a_2) = a_{11}a_{22} - a_{12}a_{21}.$$

11.2 Determinanten von Matrizen und Anwendungen

Ist $A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \in M(n, K)$, so definiere

$$\det A := \det(a_1, \dots, a_n),$$

wobei $a_i = (a_{i1}, \dots, a_{in})$ der i -te Zeilenvektor von A ist. Statt $\det A$ schreibt man auch $|A|$. Wie wir bald sehen werden, gilt $\det A = \det A^t$ und somit erhalten wir das gleiche Ergebnis, wenn wir statt der Zeilen die Spalten von A in die Determinate einsetzen.

Wir zeigen zunächst, wie sich $\det A$ unter den elementaren Zeilenumformungen verhält.

Lemma 11.2.1. *Es seien für $i \neq j$ und $\lambda \in K$, $P^{i,j}, Q^{i,j}(\lambda) \in GL(n, K)$ und $S^i(\lambda) \in M(n, K)$ die in Abschnitt 10.4 beschriebenen Elementarmatrizen. Dann folgt für $A \in M(n, K)$:*

1. $\det(P^{i,j}A) = -\det A$,
2. $\det(Q^{i,j}(\lambda)A) = \det A$,
3. $\det(S^i(\lambda)A) = \lambda \det A$.

Beweis.

1. Beim Vertauschen der Zeilen von A ändert sich das Vorzeichen von \det , denn \det ist alternierend.
2. Bezeichnen wir (a_1, \dots, a_n) die Zeilen von A , so sind $(a_1, \dots, a_{i-1}, a_i + \lambda a_j, a_{i+1}, \dots, a_n)$ die Zeilen von $Q^{i,j}(\lambda)A$. Daraus folgt:

$$\begin{aligned} \det(Q^{i,j}(\lambda)A) &= \det(a_1, \dots, a_{i-1}, a_i, a_{i+1}, \dots, a_n) + \lambda \det(a_1, \dots, a_{i-1}, a_j, a_{i+1}, \dots, a_n) \\ &= \det A. \end{aligned}$$

3. Folgt aus der Linearität in der i -ten Zeile.

Bemerkung. Ist $B = \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & \ddots & \vdots \\ 0 & \dots & b_{nn} \end{pmatrix}$ eine obere Dreiecksmatrix, so folgt

$$\det B = \sum_{\sigma \in S_n} \text{sign}(\sigma) b_{1\sigma(1)} \cdot \dots \cdot b_{n\sigma(n)} = \prod_{i=1}^n b_{ii},$$

denn für jedes $\sigma \in S_n$ mit $i > \sigma(i)$ für ein $i \in \{1, \dots, n\}$ folgt: $0 = b_{i\sigma(i)} = b_{1\sigma(1)} \cdot \dots \cdot b_{n\sigma(n)}$. Ist hingegen $\sigma \in S_n$ mit $i \leq \sigma(i)$ für alle $i \in \{1, \dots, n\}$, so folgt $\sigma(n) = n, \sigma(n-1) = n-1, \dots, \sigma(1) = 1$. Das gleiche Ergebnis erhält man natürlich für untere Dreiecksmatrizen.

Satz 11.2.2. Sei $A \in M(n, K)$. Dann gilt

$$\text{Rang } A < n \Leftrightarrow \det A = 0.$$

Beweis. “ \Rightarrow ” Rang $A < n$. Sei $B = SA$ Zeilenstufenform von B , wobei $S = B_\ell \cdot \dots \cdot B_1$ ein Produkt von Matrizen vom Typ 1 bzw. Typ 2 ist. Dann gilt: $\det B = \pm \det A$. Da Rang $A < n$, ist die n -te Zeile von B gleich null und somit folgt aus der Linearität in der n -ten Zeile $\det B = 0$.

“ \Leftarrow ” Ist $\det A = 0$, so ist Rang $A < n$, denn sonst wäre die Zeilenstufenform von der Gestalt

$$B = \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & & \vdots \\ 0 & \dots & b_{nn} \end{pmatrix}$$

mit $b_{ii} \neq 0$. Dann wäre aber wegen obiger Bemerkung und Lemma 11.2.1

$$\pm \det A = \det B = \prod_{i=1}^n b_{ii} \neq 0.$$

□

Bemerkung. Der Satz ist natürlich äquivalent zu : Rang $A = n \Leftrightarrow \det A \neq 0$. Also ist $A \in M(n, K)$ genau dann invertierbar, falls $\det A \neq 0$.

Satz 11.2.3. (Determinantenmultiplikation)

Sind $A, B \in M(n, K)$, so gilt:

$$\det(A \cdot B) = \det A \cdot \det B.$$

Beweis. Ist Rang $B < n$, so ist auch Rang $(AB) \leq$ Rang $B < n$ und somit folgt aus Satz 11.2.2:

$$0 = \det(A \cdot B) = \det A \cdot \det B.$$

Es sei Rang $B = n$, so ist $\det B \neq 0$. Betrachte die Abbildung

$$\det_B : M(n, K) \rightarrow K \text{ mit } \det_B(A) := \frac{\det(AB)}{\det B}.$$

Wir zeigen, dass \det_B die 3 charakterisierenden Eigenschaften einer Determinate besitzt. Dann folgt aus der Eindeutigkeit:

$$\det_B(A) = \det A$$

und der Satz ist bewiesen.

Sind also $a_i = (a_{i1}, \dots, a_{in})$ die Zeilen von A , so sind $(a_1 \cdot B, \dots, a_n \cdot B)$ die Zeilen von AB . Dann gilt:

1. \det_B ist linear in den Zeilen, denn

$$\begin{aligned} \det_B(A) &= \det_B(a_1, \dots, a_{i-1}, \lambda a_i + \mu a'_i, a_{i+1}, \dots, a_n) \\ &= \frac{1}{\det B} \det(a_1 B, \dots, a_{i-1} B, \lambda a_i B + \mu a'_i B, a_{i+1} B, \dots, a_n B) \\ &= \frac{1}{\det B} \det(a_1 B, \dots, \lambda a_i B, \dots, a_n B) + \frac{1}{\det B} \det(a_1 B, \dots, \mu a'_i B, \dots, a_n B) \\ &= \lambda \det_B(a_1, \dots, a_i, \dots, a_n) + \mu \det_B(a_1, \dots, a'_i, \dots, a_n) \end{aligned}$$

2. Stimmen zwei verschiedene Zeilen a_i und a_j von A überein, so stimmen die Zeilen $a_i B$ und $a_j B$ in AB überein: also ist $\det_B(A) = 0$, d.h. \det_B ist alternierend.
3. \det_B ist normiert, denn $\det_B(E_n) = \frac{\det B}{\det B} = 1$.

□

Bemerkungen. Aus Satz 11.2.2 und Satz 11.2.3 folgt:

- a) Ist $A \in GL(n, K)$, so folgt aus

$$1 = \det(E_n) = \det(AA^{-1}) = \det(A) \det(A^{-1})$$

die Gleichung $\det A^{-1} = (\det A)^{-1}$.

- b) Ist $A \in M(n, K)$ und $S \in GL(n, K)$, so gilt: $\det(SAS^{-1}) = \det S \cdot \det A \cdot (\det S)^{-1} = \det A$, d.h. $\det A$ ändert sich nicht, wenn wir A mit invertierbaren Matrizen konjugieren.

- c) Warnung: im Allgemeinen gilt: $\det(A + B) \neq \det A + \det B$.

Korollar 11.2.4. Es seien $A \in M(m, K)$, $B \in M(n, K)$ und $C \in M(m \times n, K)$. Dann ist

$\begin{pmatrix} A & C \\ 0 & B \end{pmatrix} \in M(m+n, K)$ und es gilt:

$$\det \begin{pmatrix} A & C \\ 0 & B \end{pmatrix} = \det A \cdot \det B.$$

Beweis. Ist $\det A = 0$, so ist $\text{Rang } A < m$ und daher ist

$$\text{Rang} \begin{pmatrix} A & C \\ 0 & B \end{pmatrix} < (m+n).$$

Aus Satz 11.2.2 folgt dann

$$\det \begin{pmatrix} A & C \\ 0 & B \end{pmatrix} = 0.$$

Ist $\det A \neq 0$, so ist A invertierbar und es gilt:

$$\begin{aligned} \begin{pmatrix} A & C \\ 0 & B \end{pmatrix} &= \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \begin{pmatrix} E_m & A^{-1}C \\ 0 & E_n \end{pmatrix} \\ &= \begin{pmatrix} A & 0 \\ 0 & E_n \end{pmatrix} \begin{pmatrix} E_m & 0 \\ 0 & B \end{pmatrix} \begin{pmatrix} E_m & A^{-1}C \\ 0 & E_n \end{pmatrix}. \end{aligned}$$

Dann gilt wegen Bemerkung nach Lemma 11.2.1:

$$\det \begin{pmatrix} A & C \\ 0 & B \end{pmatrix} = \det \begin{pmatrix} A & 0 \\ 0 & E_n \end{pmatrix} \cdot \det \begin{pmatrix} E_m & 0 \\ 0 & B \end{pmatrix} \cdot 1.$$

Außerdem sind $\det A = \det \begin{pmatrix} A & 0 \\ 0 & E_n \end{pmatrix}$ und $\det B = \det \begin{pmatrix} E_m & 0 \\ 0 & B \end{pmatrix}$, denn die jeweils rechten Seiten sind multilinear und alternierend in Zeilen von A bzw. B . Die rechten Seiten sind auch normiert, denn $\det \begin{pmatrix} E_m & 0 \\ 0 & E_n \end{pmatrix} = 1$. \square

Definition 11.2.5. Die Menge

$$SL(n, K) := \{A \in GL(n, K) \mid \det A = 1\}$$

ist eine Untergruppe von $GL(n, K)$. Sie heißt *spezielle lineare Gruppe*.

Satz 11.2.6. Sei $A \in M(n, K)$, so gilt: $\det A = \det A^t$.

Beweis. Sei $(a_{ij}) = A \in M(n, K) \Rightarrow A^t = (a_{ji})$. Dann folgt:
 $\det A^t = \sum_{\sigma \in S_n} \text{sign}(\sigma) a_{\sigma(1)1} \cdot \dots \cdot a_{\sigma(n)n}$. Nun gilt:

$$a_{\sigma(1)1} \cdot \dots \cdot a_{\sigma(n)n} = a_{1\sigma^{-1}(1)} \cdot \dots \cdot a_{n\sigma^{-1}(n)},$$

denn ist $\sigma(k) = i_k$, so ist $k = \sigma^{-1}(i_k)$ und damit $a_{\sigma(k)k} = a_{i_k\sigma^{-1}(i_k)}$, d.h. jeder Faktor auf der linken Seite kommt genau einmal als Faktor auf der rechten Seite vor. Daraus folgt:

$$\begin{aligned} \det A^t &= \sum_{\sigma \in S_n} \text{sign}(\sigma) a_{1\sigma^{-1}(1)}, \dots, a_{n\sigma^{-1}(n)} \\ &= \sum_{\tau \in S_n} \text{sign}(\tau^{-1}) a_{1\tau(1)}, \dots, a_{n\tau(n)}. \end{aligned}$$

Da $1 = \text{sign}(\tau^{-1}\tau) = (\text{sign}\tau^{-1})(\text{sign}\tau)$, stimmen die Vorzeichen von τ und τ^{-1} überein und $\det A^t = \sum_{\tau \in S_n} \text{sign}(\tau) a_{1\tau(1)}, \dots, a_{n\tau(n)} = \det A$. \square

Bemerkung. Dieser Satz zeigt insbesondere, dass \det multilinear und alternierend auch bezüglich der Spalten ist.

Ist $A \in GL(n, K)$, so können wir die Determinante verwenden um Gleichungssysteme zu lösen. Allerdings erfordert für große Matrizen die Berechnung der Lösungen mittels Determinaten wesentlich mehr Rechenoperationen als die Berechnung mit Hilfe des Gauss-Algorithmus.

Satz 11.2.7. (Cramersche Regel)

Sei $A \in GL(n, K)$ und $b \in K^n$. Es seien (a_1, \dots, a_n) die Spaltenvektoren von A und

$$A_i^b = (a_1, \dots, a_{i-1}, b, a_{i+1}, \dots, a_n)$$

(d.h. wir ersetzen die i -te Spalte von A durch b). Dann ist die eindeutig bestimmte Lösung $x = (x_1, \dots, x_n)$ von $Ax = b$ durch

$$x_i = \frac{\det A_i^b}{\det A}$$

gegeben.

Beweis. Da $A \in GL(n, K)$, ist $\det A \neq 0$ und es existiert genau eine Lösung $x = (x_1, \dots, x_n)$ der Gleichung

$$\sum_{j=1}^n a_j x_j = Ax = b.$$

Da \det auch linear und alternierend in den Spalten ist, gilt:

$$\det(A_i^b) = \sum_{j=1}^n x_j \det(a_1, \dots, a_{i-1}, a_j, a_{i+1}, \dots, a_n) = x_i \det A.$$

□

Mit Methoden analog zum Beweis der Cramerschen Regel erhalten wir eine Formel für die Inverse einer Matrix.

Satz 11.2.8. Sei die Matrix $A \in M(n, K)$ mit Zeilenvektoren a_1, \dots, a_n gegeben. Mit $\tilde{A} \in M(n, K)$ sei die Matrix mit den Koeffizienten

$$\tilde{a}_{ki} = \det(a_1, \dots, a_{i-1}, e_k, a_{i+1}, \dots, a_n)$$

bezeichnet. Dann folgt:

$$(\det A)E_n = A \cdot \tilde{A}.$$

Insbesondere gilt, falls $\det A \neq 0$:

$$A^{-1} = \frac{1}{\det A} \cdot \tilde{A}.$$

Beweis. Da

$$\det(a_1, \dots, a_{i-1}, a_j, a_{i+1}, \dots, a_n) = \begin{cases} \det A, & \text{für } j = i \\ 0, & \text{sonst} \end{cases} = (\det A)\delta_{ji},$$

folgt aus $a_j = \sum_{k=1}^n a_{jk} e_k$

$$\begin{aligned} (\det A) \delta_{ji} &= \sum_{k=1}^n a_{jk} \det(a_1, \dots, a_{i-1}, e_k, a_{i+1}, \dots, a_n) \\ &= \sum_{k=1}^n a_{jk} \tilde{a}_{ki}. \end{aligned}$$

Dies ist aber äquivalent zu $(\det A) E_n = A \cdot \tilde{A}$. □

Bemerkung. Ist $A \in M(n, \mathbb{Q})$ mit ganzzahligen Koeffizienten, d.h. $a_{ij} \in \mathbb{Z}$. Dann gilt: $\det A = \pm 1 \Leftrightarrow$ Koeffizienten von A^{-1} sind auch ganzzahlig.

Die Koeffizienten von \tilde{A} lassen sich wie folgt berechnen.

Lemma 11.2.9. Sei $S_{ik}(A) \in M(n-1, K)$ die Matrix, die aus A durch Streichen der i -ten Zeile und k -ten Spalte entsteht, d.h.

$$S_{ik}(A) = \begin{pmatrix} a_{11} & \dots & a_{1,k-1} & a_{1,k+1} & \dots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \dots & a_{i-1,k-1} & a_{i-1,k+1} & \dots & a_{i-1,n} \\ a_{i+1,1} & \dots & a_{i+1,k-1} & a_{i+1,k+1} & \dots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \dots & a_{n,k-1} & a_{n,k+1} & \dots & a_{nn} \end{pmatrix}.$$

Dann gilt:

$$\tilde{a}_{ki} = (-1)^{k+i} \det S_{ik}(A).$$

Beweis.

$$\tilde{a}_{ki} = \det \begin{pmatrix} a_{11} & \dots & a_{1k} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i-1,1} & \dots & a_{i-1,k} & \dots & a_{i-1,n} \\ 0 & \dots & 1 & \dots & 0 \\ a_{i+1,1} & \dots & a_{i+1,k} & \dots & a_{i+1,n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & a_{nk} & \dots & a_{nn} \end{pmatrix} = \det \begin{pmatrix} a_{11} & \dots & 0 & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i-1,1} & \dots & 0 & \dots & a_{i-1,n} \\ 0 & \dots & 1 & \dots & 0 \\ a_{i+1,1} & \dots & 0 & \dots & a_{i+1,n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & 0 & \dots & a_{nn} \end{pmatrix}$$

k -te Spalte

Dies erhalten wir, indem wir für $i \neq j$ das $-a_{jk}$ -fache der i -ten Zeile zur j -ten Zeile addieren. Durch k -maliges Vertauschen der Spalten und i -faches Vertauschen der Zeilen erhalten wir:

$$\tilde{a}_{ki} = (-1)^{i+k} \det \begin{pmatrix} 1 & \dots & 0 \\ \vdots & S_{ik}(A) & \\ 0 & & \end{pmatrix} = (-1)^{i+k} \det S_{ik}(A),$$

wobei das letzte Gleichheitszeichen aus Korollar 11.2.4 folgt. \square

Beispiel. Sei $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M(2, K)$. Dann gilt

$$(\tilde{a}_{ki}) = ((-1)^{i+k} \det S_{ik}(A)) = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Falls $\det A \neq 0$, folgt somit aus Satz 11.2.8 und $\det A = ad - bc$:

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Eine praktische Berechnungsmethode für Determinanten liefert der folgende Satz.

Satz 11.2.10. (Laplacescher Entwicklungssatz)

Sei $A \in M(n, K)$, so gilt:

$$\begin{aligned} \det A &= \sum_{k=1}^n (-1)^{i+k} a_{ik} \det S_{ik}(A) \\ &\quad (\text{Entwicklung nach der } i\text{-ten Zeile}) \\ &= \sum_{k=1}^n (-1)^{i+k} a_{ki} \det S_{ki}(A) \\ &\quad (\text{Entwicklung nach der } i\text{-ten Spalte}). \end{aligned}$$

Beweis. Aus Satz 11.2.8 und Lemma 11.2.9 folgt:

$$\det A = \sum_{k=1}^n a_{ik} \tilde{a}_{ki} = \sum_{k=1}^n (-1)^{i+k} a_{ik} \det S_{ik}(A).$$

Da $\det A = \det A^t$, gilt:

$$\det A = \det A^t = \sum_{k=1}^n (-1)^{k+i} a_{ki} \det S_{ik}(A^t)$$

und wegen $S_{ik}(A^t) = S_{ki}(A)$ folgt die Behauptung. \square

Beispiel. Für

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

ergibt z.B. die Entwicklung nach der zweiten Zeile:

$$\det A = (-1)^3 a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + (-1)^4 a_{22} \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} + (-1)^5 a_{23} \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix}$$

\square

Ist V ein K -Vektorraum der Dimension $n \in \mathbb{N}$ und $f : V \rightarrow V$ eine lineare Abbildung, so definieren wir:

Definition 11.2.11. Sei B eine Basis von V und $M_B(f) := M_{BB}(f)$ Matrixdarstellung von f bezüglich B , so setze $\det f := \det M_B(f)$.

Bemerkung. Diese Definition hängt nicht von der Wahl der Basis ab. Denn ist C eine weitere Basis, so gilt:

$$M_C(f) = M_{CC}(f) = S_{B,C} \cdot M_{BB}(f) \cdot S_{B,C}^{-1},$$

wobei $S_{B,C} = \phi_C \circ \phi_B^{-1} : K^n \rightarrow K^n \in GL(n, K)$ die zugehörige Koordinatentransformation bezeichnet (Satz 10.2.4).

Aus dem Determinantenmultiplikationssatz folgt:

$$\det M_C(f) = \det M_B(f).$$

Kapitel 12

Eigenwerte und Klassifikation von Endomorphismen

Im Satz 10.3.4 hatten wir gezeigt: Ist $f : V \rightarrow W$ eine lineare Abbildung zwischen endlich dimensionalen Vektorräumen, so können wir Basen $B \subset V$ und $C \subset W$ finden, mit

$$M_{B,C}(f) = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix},$$

wobei E_r die r -te Einheitsmatrix ist und $r = \text{Rang } f$ den Rang von f bezeichnet.

Sei nun f ein Endomorphismus, d.h. eine lineare Abbildung $f : V \rightarrow V$. Ziel ist es, eine Basis B anzugeben, so dass die quadratische Matrix $M_B(f) = M_{B,B}(f)$ eine möglichst einfache Gestalt besitzt. Dieses Problem erfordert einen wesentlich größeren algebraischen Aufwand, als die Lösung des obigen Problems. Die Matrixdarstellungen von f bezüglich zweier Basen $B, C \subset V$ sind zueinander konjugiert, d.h.

$$M_C(f) = S \cdot M_B(f) \cdot S^{-1}$$

mit $S = \phi_C \circ \phi_B^{-1} \in GL(n, K)$.

12.1 Diagonalisierbarkeit und Eigenwerte

Definition 12.1.1. Sei $f : V \rightarrow V$ ein Endomorphismus und $\dim V = n$. Dann heißt f *diagonalisierbar*, falls eine Basis $B = (b_1, \dots, b_n)$ von V existiert mit

$$M_B(f) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

d.h. $f(b_i) = \lambda_i b_i$.

Dies führt zu den Begriffen des Eigenwertes und Eigenvektors.

Definition 12.1.2. Sei $f : V \rightarrow V$ ein Endomorphismus und $\lambda \in K$. Dann heißt

$$E_\lambda(f) = \{x \in V \mid f(x) = \lambda x\}$$

der *Eigenraum* von f bezüglich λ . Ist $E_\lambda(f) \neq \{0\}$, so heißt λ *Eigenwert*. Die Vektoren $x \in E_\lambda(f)$ mit $x \neq 0$, heißen *Eigenvektoren* von f zum Eigenwert λ .

Bemerkung. Wegen

$$E_\lambda(f) = \text{Kern}(f - \lambda \text{id}_V)$$

ist $E_\lambda(f)$ ein Unterraum von V . Es ist λ genau dann Eigenwert, wenn dieser Unterraum von null verschiedene Vektoren enthält. Insbesondere ist 0 ein Eigenwert genau dann, falls $\text{Kern}(f) \neq \{0\}$. Hingegen ist der Nullvektor niemals Eigenvektor.

Satz 12.1.3. *Ein Endomorphismus $f : V \rightarrow V$ ist diagonalisierbar genau dann, wenn V eine Basis aus Eigenvektoren von f besitzt.*

Definition 12.1.4. Es seien $A, B \in M(n, K)$. Dann heißen A, B *konjugiert*, falls $S \in GL(n, K)$ existiert, mit $B = SAS^{-1}$.

Satz 12.1.5. *Eine lineare Abbildung $A : K^n \rightarrow K^n$ ist genau dann diagonalisierbar, wenn die Matrix $A \in M(n, K)$ konjugiert zu einer Diagonalmatrix ist.*

Beweis. Denn ist $b_1, \dots, b_n \in K^n$ eine Basis aus Eigenvektoren von A , so sei $S^{-1} \in GL(n, K)$ die Matrix mit den Spaltenvektoren $= (b_1, \dots, b_n)$. Somit ist $S^{-1}(e_i) = b_i$, und es folgt:

$$SAS^{-1}e_i = SAb_i = S(\lambda_i b_i) = \lambda_i e_i .$$

Ist umgekehrt $SAS^{-1} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$, mit $S \in GL(n, K)$, so sind die Spalten von S^{-1}

Eigenvektoren von A . □

Es gilt nun, dass die Summe der Eigenräume zu paarweise verschiedenen Eigenwerten direkt ist.

Lemma 12.1.6. *Sei $f : V \rightarrow V$ ein Endomorphismus mit paarweise verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_m \in K$. Dann bilden die zugehörigen Eigenräume $E_{\lambda_i}(f)$ eine direkte Summe*

$$E_{\lambda_1}(f) \oplus \dots \oplus E_{\lambda_m}(f).$$

Beweis. Zu zeigen ist: Sind $x_i \in E_{\lambda_i}(f)$ und

$$x_1 + \dots + x_m = 0,$$

so folgt: $x_1 = \dots = x_m = 0$.

Der Beweis wird durch Induktion über m geführt. Für $m = 1$ ist die Aussage offenbar trivial. Nehmen wir nun an, dass für $m - 1$ Vektoren die Aussage bewiesen ist.

Man betrachte die Gleichung

$$x_1 + \dots + x_m = 0$$

mit $x_i \in E_{\lambda_i}(f)$. Durch Anwenden von f auf beide Seiten erhält man:

$$\lambda_1 x_1 + \dots + \lambda_m x_m = 0.$$

Multiplizieren wir die erste Gleichung mit λ_m und subtrahieren wir diese von der zweiten Gleichung, so folgt

$$(\lambda_1 - \lambda_m)x_1 + \dots + (\lambda_{m-1} - \lambda_m)x_{m-1} = 0.$$

Nach Induktionsannahme gilt somit

$$(\lambda_1 - \lambda_m)x_1 = \dots = (\lambda_{m-1} - \lambda_m)x_{m-1} = 0,$$

und wegen $(\lambda_i - \lambda_m) \neq 0$ folgt

$$x_1 = \dots = x_{m-1} = 0.$$

Dann ist aber auch $x_m = 0$. □

Bemerkung. Aus dem Lemma folgt insbesondere: Ist $f : V \rightarrow V$ ein Endomorphismus mit Eigenvektoren $b_1, \dots, b_m \in V$ und paarweise verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_m \in K$, so sind b_1, \dots, b_m linear unabhängig.

Mit Hilfe dieses Lemmas erhalten wir die folgende Charakterisierung von Diagonalisierbarkeit:

Satz 12.1.7. *Sei $f : V \rightarrow V$ ein Endomorphismus, so sind folgende Aussagen äquivalent:*

1. f ist diagonalisierbar.
2. Es gibt paarweise verschiedene Eigenwerte $\lambda_1, \dots, \lambda_m$ von f , so dass gilt

$$V = E_{\lambda_1}(f) \oplus \dots \oplus E_{\lambda_m}(f).$$

3. Es existieren paarweise verschiedene Eigenwerte $\lambda_1, \dots, \lambda_m$ von f mit

$$\dim V = \dim E_{\lambda_1}(f) + \dots + \dim E_{\lambda_m}(f).$$

Beweis. Die Aussagen 1. und 2. sind offensichtlich äquivalent. Außerdem folgt aus 2. wegen Korollar 7.3.6 die Dimensionsformel in 3. Da die Summe

$$E = E_{\lambda_1}(f) \oplus \dots \oplus E_{\lambda_m}(f)$$

wegen obigen Lemmas für paarweise verschiedene Eigenwerte direkt ist, folgt wieder mit Korollar 7.3.6

$$\dim E = \dim E_{\lambda_1}(f) + \dots + \dim E_{\lambda_m}(f).$$

Ist $\dim V = \dim E$, so $V = E$, d.h. aus Aussage 3. folgt Aussage 2. □

Bemerkungen.

1. Ist insbesondere $\dim V = n$ und $f : V \rightarrow V$ ein Endomorphismus mit n paarweise verschiedenen Eigenwerten, so ist f diagonalisierbar.
2. Ist λ ein Eigenwert eines Endomorphismus $f : V \rightarrow V$, so heißt die Dimension

$$\dim E_{\lambda}(f)$$

des zugehörigen Eigenraumes auch die *geometrische Multiplizität des Eigenwertes*. Die Summe der geometrischen Multiplizitäten der Eigenwerte ist also immer kleiner oder gleich der Dimension von V . Die Gleichheit ist genau dann gegeben, falls $f : V \rightarrow V$ diagonalisierbar ist.

Nun stellt sich die Frage nach der Existenz von Eigenwerte. Ein einfaches Kriterium, mit dem wir uns genauer in den folgenden Abschnitten beschäftigen werden, liefert der folgende Satz.

Satz 12.1.8. Sei $f \in \text{End}(V)$ und $\dim V = n$. Dann ist $\lambda \in K$ genau dann Eigenwert, falls

$$\det(f - \lambda \text{id}) = 0.$$

Beweis. Da $\text{Rang}(f - \lambda \text{id}) < n \Leftrightarrow \text{Kern}(f - \lambda \text{id}) \neq \{0\}$, folgt die Behauptung aus Satz 11.2.2. \square

Bemerkung. Ist B eine Basis von V , so gilt:

$$\det(f - \lambda \text{id}) = \det(M_B(f) - \lambda E_n).$$

Wie wir im nächsten Abschnitt sehen werden, definiert

$$p(\lambda) = \det(M_B(f) - \lambda E_n)$$

ein Polynom $p : K \rightarrow K$ von der Form

$$p(\lambda) = (-1)^n \lambda^n + \dots + a_1 \lambda + a_0$$

mit $a_i \in K$. Den Nullstellen dieses Polynoms entsprechen also die Eigenwerte von f .

Um diesen Sachverhalt nutzbringend einsetzen zu können, benötigen wir einige Eigenschaften von Polynomen.

12.2 Polynome und der euklidische Algorithmus

Definition 12.2.1. Es sei K ein Körper. Ein Polynom ist eine Abbildung $p : K \rightarrow K$ mit

$$p(\tau) = \sum_{k \geq 0} a_k \tau^k,$$

wobei $a_k \neq 0$ für höchstens endlich viele $k \in \mathbb{N}_0$.

Die Zahl

$$\text{Grad } p = \max\{k \in \mathbb{N}_0 \mid a_k \neq 0\}$$

heißt der *Grad* von p . Wir setzen $\text{Grad } p = -\infty$, falls p das Nullpolynom ist. Mit $P(K)$ bezeichnen wir die Menge der Polynome auf K .

Bemerkungen.

- a) Für endliche Körper bestimmt ein Polynom nicht seine Koeffizienten. Ein Beispiel ist $K = \mathbb{Z}_2$ und $p(\lambda) = \lambda^2 + \lambda$. Dann ist $p(0) = p(1) = 0$, d.h. p ist die Nullabbildung, aber ihre Koeffizienten sind von null verschieden. Wir werden daher im Folgenden immer unendliche Körper betrachten. Dort bestimmt jedes Polynom auch seine Koeffizienten.

- b) Auf $P(K)$ sind zwei Operationen erklärt, nämlich die Addition

$$p + q = \sum_{k \geq 0} a_k \tau^k + \sum_{k \geq 0} b_k \tau^k = \sum_{k \geq 0} (a_k + b_k) \tau^k$$

und die Multiplikation

$$p \cdot q = \sum_{k \geq 0} \left(\sum_{n+m=k} a_n \cdot b_m \right) \tau^k .$$

Diese beiden Strukturen erfüllen alle Eigenschaften eines Körpers, außer die Existenz eines bezüglich der Multiplikation inversen Elementes zu $p \neq 0$. Eine solche Struktur nennt man auch *kommutativer Ring* mit 1. Die ganzen Zahlen sind z.B. ein wichtiges Beispiel eines kommutativen Ringes.

c) Offensichtlich gilt:

$$\text{Grad}(p \cdot q) = \text{Grad}(p) + \text{Grad}(q).$$

Von großer Wichtigkeit ist die Polynomdivision.

Satz 12.2.2. (Polynomdivision)

Es sei K ein Körper und P, Q Polynome in $P(K)$ mit $Q \neq 0$. Dann existieren eindeutig bestimmte Polynome $q, r \in P(K)$ mit

$$P = q \cdot Q + r ,$$

wobei $\text{Grad } r < \text{Grad } Q$.

Beweis. Existenz: Ist $P = 0$, so setze $q = r = 0$. Für $P \neq 0$ führen wir den Existenzbeweis mittels Induktion über $n = \text{Grad } P$.

Sei $n = 0 = \text{Grad } P$, d.h. $P = a\tau^0$, $a \neq 0$. Ist $\text{Grad } Q > 0$, so setze $q = 0, r = P$. Ist $\text{Grad } Q = 0$, d.h. $Q = b\tau^0$, $b \neq 0$, so setze $q = ab^{-1}\tau^0$ und $r = 0$.

Der Satz sei für $\text{Grad } P = n - 1$ bewiesen.

Sei nun der $\text{Grad } P = n$, d.h. das Polynom P ist von der Form $P = a\tau^n + \dots$ mit $a \neq 0$. Ist $\text{Grad } Q > n$, so setze $q = 0, r = P$. Falls aber $m = \text{Grad } Q \leq n$, so ist Q von der Form $Q = b\tau^m + \dots$ mit $b \neq 0$. Ist $q' = ab^{-1}\tau^{n-m}$, so gilt:

$$\text{Grad}(P - q'Q) < n.$$

Nach Induktionsvoraussetzung liefert die Division von $P - q'Q$ mit Q

$$P - q'Q = q''Q + r,$$

wobei $\text{Grad } r < \text{Grad } Q$ gilt. Somit folgt $P = (q' + q'')Q + r$, d.h. mit $q = q' + q''$ ist die Aussage auch für $\text{Grad } P = n$ bewiesen.

Eindeutigkeit: Es sei

$$P = qQ + r = q'Q + r'$$

mit $\text{Grad } r$ und $\text{Grad } r' < \text{Grad } Q$. Dann folgt:

$$(q' - q)Q = r - r'.$$

Ist $r' = r$, so ist $q' = q$, da $Q \neq 0$. Ist $r' \neq r$, so gilt:

$$\text{Grad}((q' - q)Q) = \text{Grad}(q' - q) + \text{Grad } Q = \text{Grad}(r - r') .$$

Aber dann folgt $\text{Grad}(r - r') \geq \text{Grad } Q$ im Widerspruch zur Voraussetzung. □

Bemerkungen.

1. Ist $P = q \cdot Q$, so heißt Q *Teiler* von P . Man schreibt dann auch $Q|P$. Es gilt dann $\text{Grad } Q \leq \text{Grad } P$.
2. Sind P, Q Polynome mit $Q|P$ und $P|Q$, so folgt: $\text{Grad } P = \text{Grad } Q$. Ist eines der beiden Polynome gleich null, so auch das andere Polynom. Ist ein Polynom ungleich null, so ist $P = cQ$, wobei $c \in K$ eine von null verschiedene Konstante ist.

Aus der Polynomdivision ergibt sich

Korollar 12.2.3. (Euklidischer Algorithmus)

Gegeben seien $p_1, p_2 \in P(K)$ mit $p_1 \neq 0, p_2 \neq 0$. Dann gilt:

1. Das Folgende, auf der Polynomdivision beruhende, Verfahren liefert nach endlich vielen Schritten einen Rest 0. Dieses Verfahren heißt auch *Euklidischer Algorithmus*.

$$\begin{array}{rclcl}
 p_1 & = & q_1 p_2 + p_3, & \text{Grad } p_3 & < & \text{Grad } p_2 \\
 p_2 & = & q_2 p_3 + p_4, & \text{Grad } p_4 & < & \text{Grad } p_3 \\
 p_3 & = & q_3 p_4 + p_5, & \text{Grad } p_5 & < & \text{Grad } p_4 \\
 & & \vdots & & & \vdots \\
 p_{k-2} & = & q_{k-2} p_{k-1} + p_k, & \text{Grad } p_k & < & \text{Grad } p_{k-1} \\
 p_{k-1} & = & q_{k-1} p_k & & &
 \end{array}$$

für ein $k > 0$.

2. p_k ist dann ein *größter gemeinsamer Teiler* von p_1 und p_2 (d.h. p_k teilt p_1 und p_2 und jeder gemeinsame Teiler von p_1 und p_2 teilt auch p_k). Der *größte gemeinsame Teiler* ist eindeutig bestimmt, falls wir *Normiertheit* verlangen, d.h. wenn wir verlangen, dass der *führende Koeffizient* 1 ist. Wir setzen daher

$$\text{ggT}(p_1, p_2) := \frac{1}{\alpha_k} p_k,$$

falls $p_k = \alpha_k \tau^k + \dots$ und $\alpha_k \neq 0$.

3. Es gibt Polynome $Q_1, Q_2 \in P(K)$ mit

$$\text{ggT}(p_1, p_2) = Q_1 p_1 + Q_2 p_2.$$

Sind insbesondere p_1 und p_2 *teilerfremd* (d.h. der *größte gemeinsame Teiler* ist 1), so folgt:

$$1 = Q_1 p_1 + Q_2 p_2,$$

für geeignete $Q_1, Q_2 \in P(K)$.

Bemerkung. Wir setzen $\text{ggT}(0, p) = \text{ggT}(p, 0) = p$ für $p \in P(K)$.

Beweis.

zu 1. folgt wegen Satz 12.2.2 und der Endlichkeit des Grades von p_2 .

zu 2. Aus der letzten Gleichung in 1. folgt: p_k teilt p_{k-1} . Dann folgt aus der $(k-2)$ -ten Gleichung: p_k teilt p_{k-2} . Nach endlich vielen Schritten sehen wir, dass p_k auch p_1 und p_2 teilen muss.

Ist p ein Teiler von p_1 und p_2 , so folgt aus der ersten Gleichung in 1., dass p_3 geteilt wird. Dann impliziert die zweite Gleichung, dass p auch p_4 teilt. Nach endlich vielen Schritten sehen wir, dass p auch p_k teilt.

zu 3. Die $(k-2)$ -te Gleichung liefert:

$$p_k = Q_{k-1}^1 p_{k-1} + Q_{k-2}^1 p_{k-2},$$

für $Q_{k-1}^1 := -q_{k-2}$ und $Q_k^1 := 1$. Da $p_{k-3} = q_{k-3} p_{k-2} + p_{k-1}$, folgt

$$\begin{aligned} p_k &= Q_{k-1}^1 (p_{k-3} - q_{k-3} p_{k-2}) + Q_k^1 p_{k-2} \\ &= (-Q_{k-1}^1 q_{k-3} + Q_k^1) p_{k-2} + Q_{k-1}^1 p_{k-3} \\ &= Q_{k-2}^2 p_{k-2} + Q_{k-3}^2 p_{k-3}. \end{aligned}$$

Nach i Schritten erhalten wir

$$p_k = Q_{k-i}^i p_{k-i} + Q_{k-i-1}^i p_{k-i-1}.$$

Betrachten wir diese Gleichung nach $i = (k-2)$ -Schritten und normieren wir beide Seiten, so folgt die Behauptung. \square

\square

Lemma 12.2.4. Sei $p \in P(K)$ ein von null verschiedenes Polynom. Dann ist λ genau dann eine Nullstelle von p , falls ein Polynom q existiert mit $\text{Grad } q = \text{Grad } p - 1$ und

$$p(\tau) = q(\tau)(\tau - \lambda).$$

Beweis. Sei λ eine Nullstelle von p , so folgt aus dem Euklidischen Algorithmus

$$p(\tau) = q(\tau)(\tau - \lambda) + r(\tau)$$

mit $\text{Grad } r < 1$, d.h. $r = a\tau^0 = a$ ist eine Konstante. Da $0 = p(\lambda) = a\lambda^0 = a$, folgt die Behauptung. Die Umkehrung der Aussage ist trivial. \square

Bemerkung. Ist $p \in P(K)$ mit $p(\tau) = (\tau - \lambda)^m q(\tau)$, aber λ keine Nullstelle von q , so heißt $m = m_\lambda(p)$ die *Multiplizität* (*Vielfachheit*) der Nullstelle λ von p .

Satz 12.2.5. Sei $p \in P(K)$ ein Polynom mit k verschiedenen Nullstellen $\lambda_1, \dots, \lambda_k \in K$. Dann gilt: $k \leq \text{Grad } p$.

Beweis. Folgt mittels Induktion aus obigem Lemma. \square

Definition 12.2.6. Ein Körper K heißt *algebraisch abgeschlossen*, wenn jedes Polynom $p \in P(K)$ mit $\text{Grad } p \geq 1$ eine Nullstelle in K hat.

Satz 12.2.7. Sei K algebraisch abgeschlossen und $p \in P(K)$ sowie $p \neq 0$. Dann zerfällt p in Linearfaktoren, d.h.

$$p = a_0(\tau - \lambda_1)^{m_1} \dots (\tau - \lambda_k)^{m_k}$$

mit $\text{Grad } p = m_1 + \dots + m_k$, wobei $a_0 \neq 0$.

Beweis. Induktion über $\text{Grad } p = n$. Sei $n = 0$, so ist $p = a_0\tau^0$ und $m_1, \dots, m_k = 0$. Die Aussage sei für $n - 1$ bewiesen. Sei $\text{Grad } p = n$. Da K algebraisch abgeschlossen, hat p eine Nullstelle λ und damit ist $p = (\tau - \lambda)q$ mit $\text{Grad } q = n - 1$. Dann folgt aus der Induktionsannahme die Aussage für $\text{Grad } p = n$. \square

Aus dem Fundamentalsatz 3.3.7 der Algebra folgt:

Satz 12.2.8. Der Körper der komplexen Zahlen \mathbb{C} ist algebraisch abgeschlossen.

12.3 Das charakteristische Polynom

Seien $A \in M(n, K)$ und τ eine Unbestimmte, so betrachte

$$\begin{aligned} \chi_A &= \det(A - \tau E_n) = \det \begin{pmatrix} a_{11} - \tau & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \tau & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \tau \end{pmatrix} \\ &= \sum_{\sigma \in S_n} \text{sign}(\sigma) (a_{1\sigma(1)} - \delta_{1\sigma(1)}\tau) \cdot \dots \cdot (a_{n\sigma(n)} - \delta_{n\sigma(n)}\tau) \\ &= \alpha_n \tau^n + \alpha_{n-1} \tau^{n-1} + \dots + \alpha_0 \\ &= (a_{11} - \tau) \cdot \dots \cdot (a_{nn} - \tau) + q \end{aligned}$$

mit $\text{Grad } q \leq n - 2$, denn aus $\sigma \neq \text{id}$ folgt $\delta_{j\sigma(j)} = 0$ für wenigstens zwei Elemente in $\{1, \dots, n\}$. Da

$$(a_{11} - \tau) \cdot \dots \cdot (a_{nn} - \tau) = (-1)^n \tau^n + (-1)^{n-1} (a_{11} + \dots + a_{nn}) \tau^{n-1} + q'$$

mit $\text{Grad } q' \leq n - 2$ und $\chi_A(0) = \det(A - 0E_n) = \det A = \alpha_0$, gilt insbesondere:

$$\alpha_n = (-1)^n, \alpha_{n-1} = a_{11} + \dots + a_{nn} = \text{Spur } A \text{ und } \alpha_0 = \det A.$$

Satz 12.3.1. Seien $A, B \in M(n, K)$ zueinander konjugiert, d.h. $B = SAS^{-1}$ mit $S \in GL(n, K)$, so folgt:

$$\chi_A = \chi_B \in P(K).$$

Bemerkung. Das bedeutet also: Alle Koeffizienten von χ_A und χ_B stimmen überein, d.h. die Koeffizienten des charakteristischen Polynoms sind Konjugationsinvarianten.

Beweis.

$$\begin{aligned}\chi_B(\tau) &= \det(SAS^{-1} - \tau \cdot SE_nS^{-1}) \\ &= \det(S(A - \tau \cdot E_n)S^{-1}) \\ &= \det(A - \tau E_n)\end{aligned}$$

wegen Bemerkung b) nach Satz 11.2.3. □

Definition 12.3.2. Sei $f \in \text{End}(V)$ und B eine Basis von V . Definiere

$$\chi_f(\tau) = \chi_{M_B(f)}(\tau) \in P(K).$$

Bemerkung. Diese Definition ist wegen des obigen Satzes unabhängig von der Wahl der Basis.

Wir hatten in 12.1.8 gesehen:

$$\chi_f(\lambda) = 0 \Leftrightarrow \dim E_\lambda(f) > 0.$$

Jedoch muss die Multiplizität der Nullstelle von χ_f keineswegs mit der Dimension des zugehörigen Eigenraumes übereinstimmen. Es gilt aber:

Lemma 12.3.3. Ist $f \in \text{End}(V)$, so ist

$$\dim E_\lambda(f) \leq m_\lambda(\chi_f).$$

Beweis. Sei (b_1, \dots, b_k) eine Basis von $E_\lambda(f)$, so ergänzen wir diese zu einer Basis $(b_1, \dots, b_k, b_{k+1}, \dots, b_n)$ von V . Dann gilt:

$$M_B(f) = \left(\begin{array}{cc|c} \lambda & 0 & \\ & \ddots & * \\ 0 & \lambda & \\ \hline & 0 & D \end{array} \right)$$

und damit folgt aus Korollar 11.2.4

$$\chi_f(\tau) = (\lambda - \tau)^k \cdot \det(D - \tau E_{n-k})$$

und $k \leq m_\lambda(\chi_f)$. □

Beispiel. Betrachte $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in M(2, K)$, so folgt: $\chi_A(\tau) = (1 - \tau)^2$, und daher ist

$m_1(\chi_A) = 2$, aber $\text{Kern}(A - 1E_2) = \text{Kern} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \text{Span}\{(1, 0)\}$, d.h. $\dim E_1(A) = 1$.

Dies zeigt insbesondere: A ist nicht diagonalisierbar, da K^2 keine Basis aus Eigenvektoren von A besitzt.

Nun wollen wir ein Kriterium für die Diagonalisierbarkeit eines Endomorphismus angeben.

Satz 12.3.4. Sei V ein n -dimensionaler K -Vektorraum und $f \in \text{End}(V)$. Dann sind folgende Aussagen äquivalent:

1. f ist diagonalisierbar.
2. Das charakteristische Polynom χ_f zerfällt in Linearfaktoren, d.h.

$$\chi_f = (\lambda_1 - \tau)^{m_1} \cdot \dots \cdot (\lambda_k - \tau)^{m_k}$$

(der führende Koeffizient von χ_f ist $(-1)^n$) und es gilt

$$m_{\lambda_i}(\chi_f) = m_i = \dim E_{\lambda_i}(f).$$

Bemerkung. Ist insbesondere K der Körper der komplexen Zahlen, so ist K algebraisch abgeschlossen. Dann ist also f genau dann diagonalisierbar, falls die Multiplizitäten der Nullstellen des charakteristischen Polynoms mit den Dimensionen der zugehörigen Eigenräume übereinstimmen.

Beweis. Ist f diagonalisierbar, so existieren wegen Satz 12.1.7 paarweise verschiedene Eigenwerte $\lambda_1, \dots, \lambda_k$ mit

$$V = E_{\lambda_1}(f) \oplus \dots \oplus E_{\lambda_k}(f).$$

Sind nun B_i Basen von $E_{\lambda_i}(f)$, so ist $B = B_1 \cup \dots \cup B_m$ eine Basis von V . Ist $m_i = \dim E_{\lambda_i}(f)$, so hat f bezüglich B die Matrixdarstellung

$$M_B(f) = \left(\begin{array}{cccc} \lambda_1 & & & 0 \\ & \ddots & & \\ & & \lambda_1 & \\ \vdots & & \ddots & \vdots \\ & & & \lambda_k \\ & & & \ddots \\ 0 & & & & \lambda_k \end{array} \right) \left. \begin{array}{l} \left. \vphantom{\begin{matrix} \lambda_1 \\ \vdots \\ \lambda_k \end{matrix}} \right\} m_1 \\ \vdots \\ \left. \vphantom{\begin{matrix} \lambda_k \\ \vdots \\ \lambda_k \end{matrix}} \right\} m_k \end{array} \right\}$$

Damit gilt:

$$\chi_f(\tau) = \det(M_B(f) - \tau \text{id}) = (\lambda_1 - \tau)^{m_1} \cdot \dots \cdot (\lambda_k - \tau)^{m_k},$$

und $m_{\lambda_i}(f) = m_i = \dim E_{\lambda_i}(f)$.

Ist umgekehrt 2. erfüllt, so gilt:

$$\dim E_{\lambda_1}(f) + \dots + \dim E_{\lambda_k}(f) = m_1 + \dots + m_k = n = \dim V,$$

und die Aussage folgt aus Satz 12.1.7. □

12.4 Der Satz von Cayley-Hamilton

Es wird im Folgenden wichtig sein, auch Endomorphismen als Variable von Polynomen zuzulassen.

Satz 12.4.1. *Sei V ein Vektorraum über K und $f : V \rightarrow V$ eine lineare Abbildung. Betrachte die Abbildung $\phi_f : P(K) \rightarrow \text{End}(V)$ mit*

$$\phi_f(p) = p(f) = a_0 f^0 + a_1 f + \dots + a_m f^m,$$

falls $p(\tau) = a_0 \tau^0 + a_1 \tau + \dots + a_m \tau^m$. Dann gilt

$$\phi_f(p + q) = \phi_f(p) + \phi_f(q) \quad \text{und} \quad \phi_f(p \cdot q) = \phi_f(p) \circ \phi_f(q).$$

Beweis. Die Linearitätseigenschaft $\phi_f(p + q) = \phi_f(p) + \phi_f(q)$ folgt unmittelbar aus der Definition. Wegen der Linearität genügt es, die zweite Eigenschaft für Monome, also Polynome der Form $p(\tau) = \tau^k$ nachzuweisen. Seien also $p(\tau) = \tau^k$ und $q(\tau) = \tau^m$, so folgt

$$\phi_f(p \cdot q) = f^{k+l} = f^k \circ f^l = \phi_f(p) \circ \phi_f(q).$$

□

Nun werden wir sehen, dass jede Matrix $A \in M(n, K)$ Nullstelle ihres charakteristischen Polynoms ist, d.h.

$$\phi_A(\chi_A(\tau)) = \phi_A\left(\sum_{k=0}^n \alpha_k \tau^k\right) = \sum_{k=0}^n \alpha_k A^k = 0 \in M(n, K).$$

Dies ist der Inhalt des Satzes von Cayley-Hamilton. Die Tatsache, dass jede Matrix $A \in M(n, K)$ Nullstelle einer Polynomgleichung ist, folgt, da $\dim M(n, K) = n^2$. Dann sind die Vektoren A^0, A^1, \dots, A^{n^2} linear abhängig, d.h. es existieren Koeffizienten $\alpha_0, \dots, \alpha_{n^2}$ mit

$$\alpha_0 A^0 + \alpha_1 A^1 + \dots + \alpha_{n^2} A^{n^2} = 0.$$

Der Grad dieses Polynoms ist jedoch n^2 .

Satz 12.4.2. (Cayley-Hamilton)

Sei $A \in M(n, K)$ und

$$\chi_A(\tau) = (-1)^n \tau^n + \alpha_1 \tau^{n-1} + \dots + \alpha_0 = \det(A - \tau \text{id})$$

das charakteristische Polynom. Dann gilt

$$\phi_A(\chi_A) = (-1)^n A^n + \alpha_1 A^{n-1} + \dots + \alpha_0 E_n = 0.$$

Beweis. Man betrachte zu $A - \tau \text{id}$ wie in Satz 11.2.8 die Matrix

$$\widetilde{A - \tau \text{id}} = (\tilde{a}_{ki}(\tau)) = \tilde{A}(\tau)$$

mit

$$\tilde{a}_{ki}(\tau) = \det(a_1 - \tau e_1, \dots, a_{i-1} - \tau e_{i-1}, e_k, a_{i+1} - \tau e_{i+1}, \dots, a_n - \tau e_n) \in P(K).$$

Dies impliziert: Grad $\tilde{a}_{ki}(\tau) = n - 1$. Dann gilt nach 11.2.8

$$\det(A - \tau \text{id})E_n = (A - \tau \text{id}) \cdot \tilde{A}(\tau) .$$

Da $\tilde{a}_{ki}(\tau) = \sum_{\ell=0}^{n-1} C_\ell^{ki} \tau^\ell$, so können wir die Matrix $\tilde{A}(\tau)$ darstellen als:

$$\tilde{A}(\tau) = \sum_{\ell=0}^{n-1} C_\ell \tau^\ell$$

mit $C_\ell \in M(n, K)$ und $(C_\ell)_{ki} = C_\ell^{ki}$. Damit gilt:

$$\begin{aligned} (A - \tau \text{id}) \left(\sum_{\ell=0}^{n-1} C_\ell \tau^\ell \right) &= \sum_{\ell=0}^{n-1} AC_\ell \tau^\ell - C_\ell \tau^{\ell+1} \\ &= AC_0 + (AC_1 - C_0)\tau + (AC_2 - C_1)\tau^2 \\ &\quad + \dots + (AC_{n-1} - C_{n-2})\tau^{n-1} - C_{n-1}\tau^n \\ &= (\alpha_0 + \alpha_1\tau + \dots + \alpha_{n-1}\tau^{n-1} + (-1)^n\tau^n)E_n . \end{aligned}$$

Koeffizientenvergleich liefert die $n + 1$ Gleichungen:

$$\begin{aligned} AC_0 &= \alpha_0 E_n \\ AC_1 - C_0 &= \alpha_1 E_n \\ &\vdots \\ AC_{n-1} - C_{n-2} &= \alpha_{n-1} E_n \\ -C_{n-1} &= (-1)^n E_n . \end{aligned}$$

Nach Multiplikation der i -ten Zeile mit A^{i-1} und Addition der Zeilen der linken und rechten Seite, erhält man:

$$\begin{aligned} &AC_0 + (A^2C_1 - AC_0) + (A^3C_2 - A^2C_1) + \dots + \\ &(A^n C_{n-1} - A^{n-1} C_{n-2}) - A^n C_{n-1} = 0 \\ &= \alpha_0 E_n + \alpha_1 A + \dots + \alpha_{n-1} A^{n-1} + (-1)^n A^n \\ &= \phi_A(\chi_A) . \end{aligned}$$

□

Bemerkung. Ist $A \in M(2, K)$, so gilt insbesondere:

$$A^2 - (\text{Spur } A)A + (\det A)E_2 = 0.$$

Das charakteristische Polynom von A ist im Allgemeinen keineswegs das Polynom mit minimalem Grad, das A als "Nullstelle" besitzt.

Definition 12.4.3. Sei $A \in M(n, K)$ und

$$N_A = \{p \in P(K) \mid p \neq 0, \phi_A(p) = 0\}.$$

Dann heißt $\mu \in N_A$ *Minimalpolynom*, falls

$$\text{Grad } \mu = \min\{\text{Grad } p \mid p \in N_A\}$$

und μ normiert ist, d.h. der höchste Koeffizient 1 ist.

Satz 12.4.4. Sei $A \in M(n, K)$. Dann gibt es genau ein Minimalpolynom $\mu = \mu_A$. Ist $p \in N_A$, so teilt μ das Polynom p , d.h. $p = q \cdot \mu$.

Beweis. Wegen des Satzes 12.4.2 von Cayley-Hamilton ist $\chi_A \in N_A$ und somit $N_A \neq \emptyset$. Daher existiert ein Minimalpolynom $\mu \neq 0$. Sei $p \in N_A$, so liefert der euklidische Algorithmus Polynome q, r mit

$$p = q \cdot \mu + r,$$

wobei $\text{Grad } r < \text{Grad } \mu$, falls $r \neq 0$. Ist $r \neq 0$, so gilt:

$$0 = \phi_A(p) = \phi_A(q \cdot \mu) + \phi_A(r) = (\phi_A(q)) \cdot (\phi_A(\mu)) + \phi_A(r) = \phi_A(r),$$

im Widerspruch zur Minimalität von μ . Daher gilt: $r = 0$.

Ist μ' ein weiteres Minimalpolynom, so folgt: $\mu' = q\mu$ mit $\text{Grad } q = 0$, d.h. $q = a\tau^0$. Da μ, μ' normiert sind, ist $a = 1$. \square

Bemerkung. Insbesondere gilt: Das Minimalpolynom teilt das charakteristische Polynom.

Satz 12.4.5. Sei $A \in M(n, K)$, so haben das charakteristische Polynom χ_A sowie das Minimalpolynom μ_A die gleichen Nullstellen. Dabei sind die Multiplizitäten der Nullstellen von μ_A nicht größer, als die Multiplizitäten von χ_A . Zerfällt insbesondere das charakteristische Polynom in Linearfaktoren, d.h.

$$\chi_A(\tau) = (\lambda_1 - \tau)^{m_1} \cdot \dots \cdot (\lambda_k - \tau)^{m_k},$$

so ist

$$\begin{aligned} \mu_A(\tau) &= (-1)^{l_1 + \dots + l_k} (\lambda_1 - \tau)^{l_1} \cdot \dots \cdot (\lambda_k - \tau)^{l_k} \\ &= (\tau - \lambda_1)^{l_1} \cdot \dots \cdot (\tau - \lambda_k)^{l_k}, \end{aligned}$$

mit $l_i \leq m_i$.

Beweis. Sei $\lambda \in K$ eine Nullstelle von χ_A , so gibt es einen Eigenvektor $x \in K^n$ mit $Ax = \lambda x$. Sei

$$\mu_A(\tau) = \tau^m + a_{m-1}\tau^{m-1} + \dots + a_0$$

das Minimalpolynom von A . Dann folgt:

$$\begin{aligned} 0 &= (A^m + a_{m-1}A^{m-1} + \dots + a_0E_n)x \\ &= A^m x + a_{m-1}A^{m-1}x + \dots + a_0E_n x \\ &= \lambda^m x + a_{m-1}\lambda^{m-1}x + \dots + a_0x \\ &= (\lambda^m + a_{m-1}\lambda^{m-1} + \dots + a_0)x. \end{aligned}$$

Da $x \neq 0$, folgt $\mu_A(\lambda) = 0$. Da das Minimalpolynom das charakteristische Polynom teilt, sind die Multiplizitäten der Nullstellen von μ_A nicht größer als die von χ_A . Ist insbesondere

$$\chi_A(\tau) = (\lambda_1 - \tau)^{m_1} \cdot \dots \cdot (\lambda_k - \tau)^{m_k},$$

so ist μ_A von der Form

$$\mu_A(\tau) = (\tau - \lambda_1)^{l_1} \cdot \dots \cdot (\tau - \lambda_k)^{l_k} p(\tau),$$

wobei $l_i \leq m_i$ und das Polynom p keine Nullstellen besitzt. Da μ_A das Polynom χ_A teilt, existiert ein Polynom q mit

$$(\lambda_1 - \tau)^{m_1 - l_1} \cdot \dots \cdot (\lambda_k - \tau)^{m_k - l_k} = p(\tau)q(\tau).$$

Da $p(\lambda_i) \neq 0$, ist λ_i Nullstelle von q mit Multiplizität $m_i - l_i$. Daraus folgt, dass p konstant ist und da μ_A normiert ist, ist diese Konstante eins. \square

12.5 Invariante Unterräume

Definition 12.5.1. Es sei V ein n -dimensionaler K -Vektorraum und $f : V \rightarrow V$ eine lineare Abbildung. Ein Unterraum $U \subset V$ heißt *invariant* unter f (*f-invariant*), falls $f(U) \subseteq U$.

Eigenräume sind Beispiele invarianter Unterräume. Wir wollen nun Kriterien herleiten, die eine Aufspaltung von V in f -invariante Unterräume implizieren. Dafür werden Polynome $P(K)$ von großer Bedeutung sein.

Lemma 12.5.2. Sei V ein K -Vektorraum, $f \in \text{End}(V)$ und $p \in P(K)$. Dann ist der Unterraum

$$U = \text{Kern } p(f)$$

f-invariant.

Beweis. Aus $p(f) = \alpha_k f^k + \dots + a_0 \text{id}$ folgt:

$$p(f)f = \alpha_k f^{k+1} + \dots + a_0 f = f(\alpha_k f^k + \dots + \alpha_0 \text{id}) = fp(f).$$

Also gilt für $x \in \text{Kern } p(f)$

$$p(f)f(x) = fp(f)(x) = f(0) = 0,$$

d.h. $f(x) \in \text{Kern } p(f)$. \square

Unter den Voraussetzungen des obigen Lemmas gilt:

Lemma 12.5.3. *Es seien $p_1, p_2 \in P(K)$ Polynome und $p_1 \mid p_2$, d.h. p_1 teile p_2 . Dann gilt:*

$$\text{Kern } p_1(f) \subset \text{Kern } p_2(f).$$

Beweis. Aus $p_1 \mid p_2$ folgt die Existenz eines Polynoms $q \in P(K)$ mit $p_2 = q \cdot p_1$. Dann folgt $p_2(f) = q(f) \circ p_1(f)$ und somit $\text{Kern } p_1(f) \subset \text{Kern } p_2(f)$. \square

Satz 12.5.4. *Es seien $p_1, p_2 \in P(K)$. Dann gilt:*

1. *Ist $q = \text{ggT}(p_1, p_2)$, so folgt:*

$$\text{Kern } q(f) = \text{Kern } p_1(f) \cap \text{Kern } p_2(f).$$

2. *Ist $p = p_1 p_2$ und $1 = \text{ggT}(p_1, p_2)$, so folgt:*

$$\text{Kern } p(f) = \text{Kern } p_1(f) \oplus \text{Kern } p_2(f).$$

Beweis. zu 1.: Da $q \mid p_1, q \mid p_2$, erhalten wir aus Lemma 12.5.3:

$$\text{Kern } q(f) \subset \text{Kern } p_1(f) \cap \text{Kern } p_2(f).$$

Da q größter gemeinsamer Teiler ist, gibt es nach 12.2.3 Polynome $P, Q \in P(K)$ mit $q = Pp_1 + Qp_2$. Ist $x \in \text{Kern } p_1(f) \cap \text{Kern } p_2(f)$, so folgt:

$$q(f)(x) = P(f)p_1(f)x + Q(f)p_2(f)x = 0,$$

und somit ist $x \in \text{Kern } q(f)$.

zu 2.: Wegen 1. bleibt zu zeigen:

$$\text{Kern } p = \text{Kern } p_1 + \text{Kern } p_2.$$

Da $p_1 \mid p$ und $p_2 \mid p$, gilt:

$$\text{Kern } p_1(f) \subset \text{Kern } p(f), \quad \text{Kern } p_2(f) \subset \text{Kern } p(f)$$

und somit

$$\text{Kern } p_1(f) + \text{Kern } p_2(f) \subset \text{Kern } p(f).$$

Zu zeigen bleibt die umgekehrte Inklusion. Da $1 = \text{ggT}(p_1, p_2)$, existieren Polynome $q_1, q_2 \in P(K)$ mit

$$1 = q_1 p_1 + q_2 p_2.$$

Somit erhält man für jedes $x \in V$ die Zerlegung

$$x = 1(f)x = q_1(f)p_1(f)x + q_2(f)p_2(f)x = x_2 + x_1$$

mit $x_2 = q_1(f)p_1(f)x$ und $x_1 = q_2(f)p_2(f)x$. Ist $x \in \text{Kern } p(f)$, so ist $x_2 \in \text{Kern } p_2(f)$, denn

$$p_2(f)x_2 = p_2(f)q_1(f)p_1(f)x = q_1(f)p_1(f)p_2(f)x = q_1(f)(p_1 p_2)(f)x = 0.$$

Genauso folgt: $x_1 \in \text{Kern } p_1(f)$. \square

Wir wollen den zweiten Teil dieses Satzes auf beliebig viele Faktoren ausdehnen. Dazu benötigen wir noch das folgende Lemma:

Lemma 12.5.5. *Es seien $p_1, \dots, p_n \in P(K)$ Polynome mit $\text{ggT}(p_i, p_j) = 1$ für $i \neq j$. Dann folgt*

$$\text{ggT}(p_1 \cdot \dots \cdot p_{n-1}, p_n) = 1.$$

Beweis. Induktion über n . Für $n = 2$ ist nichts zu zeigen.

Die Aussage sei für $2 \leq k \leq n$ bewiesen.

Seien $p_1, \dots, p_{n+1} \in P(K)$ mit $\text{ggT}(p_i, p_j) = 1$ für $i \neq j$. Es sei

$$p = \text{ggT}(p_1 \cdot \dots \cdot p_n, p_{n+1}).$$

Aus der Induktionsvoraussetzung folgt: $\text{ggT}(p_1 \cdot \dots \cdot p_{n-1}, p_{n+1}) = 1 \Rightarrow \exists q_1, q_2 \in P(K)$

$$1 = q_1 p_1 \cdot \dots \cdot p_{n-1} + q_2 p_{n+1}$$

und daher $p_n = q_1 p_1 \cdot \dots \cdot p_n + q_2 p_{n+1} p_n$. Da $p|p_{n+1}$ und $p|p_1 \cdot \dots \cdot p_n \Rightarrow p|p_n$. Wegen $\text{ggT}(p_n, p_{n+1}) = 1$, folgt $p = 1$. \square

Satz 12.5.6. *Es sei $f \in \text{End}(V)$ und $p = p_1 \cdot \dots \cdot p_l$ mit $\text{ggT}(p_i, p_j) = 1$ für $i \neq j$. Dann gilt:*

$$\text{Kern } p(f) = \text{Kern } p_1(f) \oplus \dots \oplus \text{Kern } p_l(f).$$

Beweis. Induktion über $l \in \mathbb{N}$.

$l = 1$: trivial.

Die Aussage sei für $l \in \mathbb{N}$ bewiesen.

Sei $p = p_1 \cdot \dots \cdot p_l \cdot p_{l+1}$ mit $\text{ggT}(p_i, p_j) = 1$ für $i \neq j$. Dann folgt aus Lemma 12.5.5:

$$\text{ggT}(q, p_{l+1}) = 1, \text{ wobei } q = p_1 \cdot \dots \cdot p_l.$$

Mit Satz 12.5.5 (2) erhalten wir:

$$\text{Kern } p(f) = \text{Kern } q(f) \oplus \text{Kern } p_{l+1}(f) = \text{Kern } p_1(f) \oplus \dots \oplus \text{Kern } p_l(f) \oplus \text{Kern } p_{l+1}(f)$$

nach Induktionsvoraussetzung. \square

Bemerkung. Insbesondere lässt sich dieser Satz auf das charakteristische bzw. Minimalpolynom anwenden. In diesem Fall ist $p(f) = 0$ und daher $\text{Kern } p(f) = V$. Zerfällt das charakteristische Polynom in Linearfaktoren, so erhalten wir:

Satz 12.5.7. *Sei $f \in \text{End}(V)$, wobei das charakteristische Polynom in Linearfaktoren zerfalle, d.h.*

$$\chi_f(\tau) = (\lambda_1 - \tau)^{m_1} \dots (\lambda_k - \tau)^{m_k},$$

mit $\lambda_i \neq \lambda_j$ für $i \neq j$. Dann existiert eine Zerlegung von V

$$V = V_1 \oplus \dots \oplus V_k$$

in f -invariante Unterräume $V_i = \text{Kern}(f - \lambda_i \text{id})^{m_i}$. Ferner ist $(\lambda_i - \tau)^{m_i}$ das charakteristische Polynom von $f|_{V_i} : V_i \rightarrow V_i$ und daher folgt:

$$\dim V_i = m_i.$$

Da das Minimalpolynom μ_f von der Form

$$\mu_f(\tau) = (\tau - \lambda_1)^{l_1} \dots (\tau - \lambda_k)^{l_k}$$

ist, mit $l_i \leq m_i$, gilt sogar: $V_i = \text{Kern}(f - \lambda_i \text{id})^{l_i}$.

Beweis. Da $p_i(\tau) = (\tau - \lambda_i)^{m_i}, i \in \{1, \dots, k\}$ paarweise teilerfremd sind, folgt aus Satz 12.5.6 angewandt auf das charakteristische Polynom die Existenz der Zerlegung $V = V_1 \oplus \dots \oplus V_k$ in f -invariante Unterräume $V_i = \text{Kern}(f - \lambda_i \text{id})^{m_i}$. Wenden wir Satz 12.5.6 auf das Minimalpolynom an, so erhalten wir eine weitere Zerlegung

$$V = V'_1 \oplus \dots \oplus V'_k$$

mit $V'_i = \text{Kern}(f - \lambda_i \text{id})^{l_i}$. Da $l_i \leq m_i$, folgt

$$\text{Kern}(f - \lambda_i \text{id})^{l_i} \subset \text{Kern}(f - \lambda_i \text{id})^{m_i}$$

und somit $V'_i \subset V_i$. Dann muss aber $\dim V'_i = \dim V_i$ gelten und daher ist $V'_i = V_i$. □

Wir geben nun ein weiteres Kriterium für die Diagonalisierbarkeit eines Endomorphismus an.

Korollar 12.5.8. Sei $f \in \text{End}(V)$ und V ein n -dimensionaler K -Vektorraum. Der Endomorphismus f ist genau dann diagonalisierbar, falls das Minimalpolynom μ_f in Linearfaktoren der Multiplizität eins zerfällt, d.h.

$$\mu_f(\tau) = (\tau - \lambda_1) \cdot \dots \cdot (\tau - \lambda_k)$$

mit $\lambda_i \neq \lambda_j$ für $i \neq j$.

Beweis. Ist f diagonalisierbar, so existiert eine Zerlegung

$$V = E_{\lambda_1} \oplus \dots \oplus E_{\lambda_k}$$

in Eigenräume mit $\lambda_i \neq \lambda_j$. Man betrachte das Polynom $m(\tau) = (\tau - \lambda_1) \cdot \dots \cdot (\tau - \lambda_k)$. Dann gilt:

$$m(f) = (f - \lambda_1 \text{id}) \circ \dots \circ (f - \lambda_k \text{id}) = 0,$$

denn da

$$(f - \lambda_j \text{id}) \circ (f - \lambda_i \text{id}) = (f - \lambda_i \text{id}) \circ (f - \lambda_j \text{id}),$$

folgt für alle $j \in \{1, \dots, k\}$ und $x \in E_{\lambda_j}$:

$$m(f)x = \prod_{\substack{i=1 \\ i \neq j}}^k (f - \lambda_i \text{id}) \circ (f - \lambda_j \text{id})x = 0.$$

Aus der Zerlegung von V in Eigenräume folgt $m(f) = 0$. Damit teilt μ_f wegen Satz 12.4.3 das Polynom m und da μ_f die Eigenwerte λ_i auch als Nullstellen besitzt (Satz 12.4.4) und μ_f normiert ist, folgt $m = \mu_f$.

Hat umgekehrt das Minimalpolynom μ_f die obige Form, so folgt nach Satz 12.5.7 die Zerlegung von V in f -invariante Unterräume $V_i = \text{Kern}(f - \lambda_i \text{id}) = E_{\lambda_i}$. Dies ist aber gleichbedeutend mit der Diagonalisierbarkeit von f . □

□

Definition 12.5.9. Sei V ein endlich dimensionaler K -Vektorraum.

$f \in \text{End}(V)$ heißt *nilpotent*, falls $f^m = 0$ für ein $m \in \mathbb{N}$. Ist $m \in \mathbb{N}$ die minimale Zahl mit dieser Eigenschaft, d.h. ist $f^m = 0, f^{m-1} \neq 0$, so heißt m der *Nilpotenzgrad* von f .

Wie wir im Satz 12.5.7 gesehen haben, entspricht dem Zerfall des charakteristischen Polynoms in Linearfaktoren eine Aufspaltung des Vektorraumes in invariante Unterräume der Form $V_i = \text{Kern}(f - \lambda_i \text{id})^{m_i}$. Aus der f -Invarianz von V_i folgt: $(f - \lambda_i \text{id})(V_i) \subset V_i$ und nach Definition von V_i ist diese Abbildung nilpotent, denn $\left((f - \lambda_i \text{id})|_{V_i}\right)^{m_i} = 0$. Wir werden uns deshalb im folgenden genauer mit nilpotenten Endomorphismen beschäftigen.

Lemma 12.5.10. *Sei $f \in \text{End}(V)$ und $\dim V = n$. Dann sind folgende Aussagen äquivalent:*

1. f ist nilpotent.
2. Es ist $f^n = 0$, d.h. der Nilpotenzgrad von f ist $\leq n$.
3. Es existieren Unterräume V_0, \dots, V_n mit $\dim V_i = i$ und

$$V_0 = \{0\} \subset V_1 \subset V_2 \subset \dots \subset V_{n-1} \subset V_n = V,$$

so dass $f(V_i) \subset V_{i-1}$ gilt.

4. Es existiert eine Basis $B = (b_1, \dots, b_n)$, so dass $M_B(f)$ eine obere Dreiecksmatrix darstellt, deren Diagonalelemente alle gleich null sind.

Beweis. Wir zeigen die Aussage durch Induktion über $n \in \mathbb{N}$. Aus 1. folgen die Aussagen 2. und 3. Sei also f nilpotent und $n = 1$. Dann ist f die Nullabbildung und somit sind die Aussagen 2. und 3. trivialerweise erfüllt.

Nehmen wir also an, dass für $n - 1$ aus 1. die Aussagen 2. und 3. folgen.

Sei nun $f \in \text{End}(V)$ nilpotent und $\dim V = n > 0$. Definiere $V_n = V$. Offensichtlich ist f nicht bijektiv, denn sonst ist auch f^k bijektiv für alle $k \in \mathbb{N}$. Somit ist f auch nicht surjektiv und es existiert ein Unterraum V_{n-1} mit $\dim V_{n-1} = n - 1$ und $\text{Bild}(f) \subset V_{n-1}$. Dann ist $f : V_{n-1} \rightarrow V_{n-1}$ ein nilpotenter Endomorphismus und nach Induktionsvoraussetzung folgt $f^{n-1}(x) = 0$ für alle $x \in V_{n-1}$. Nach Induktionsvoraussetzung existieren somit Unterräume

$$0 = V_0 \subset V_1 \subset \dots \subset V_{n-1}$$

mit $f(V_i) \subset V_{i-1}$.

Aus 3. folgt 4. Sei $b_1 \in V_1$ und $b_1 \neq 0$ so existiert ein Vektor $b_2 \in V_2$, so dass (b_1, b_2) eine Basis von V_2 ist. Mit einem weiteren Vektor $b_3 \in V_3$ können wir (b_1, b_2) zu einer Basis von (b_1, b_2, b_3) von V_3 ergänzen. Nach n Schritten erhalten wir somit eine Basis (b_1, \dots, b_n) von V mit der Eigenschaft, dass (b_1, \dots, b_i) eine Basis von V_i bildet. Da $f(V_j) \subset V_{j-1}$ ist

$$f(b_j) = \sum_{i=1}^{j-1} a_{ij} b_i,$$

und somit hat $M_B(f)$ die in 4. beschriebene Form. Aus 4. folgt 3. mit $V_i := (b_1, \dots, b_i)$. Aus 3. folgt 2. denn $f^i(V) \subset V_{n-i}$. Aus 2. folgt 1. trivialerweise. \square

Satz 12.5.11. *Sei $f \in \text{End}(V)$, wobei das charakteristische Polynom in Linearfaktoren zerfalle, d.h.*

$$\chi_f(\tau) = (\tau - \lambda_1)^{m_1} \dots (\tau - \lambda_k)^{m_k}.$$

Dann existiert eine Basis $B \subset V$, so dass

$$M_B(f) = \begin{pmatrix} A_{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & A_{\lambda_k} \end{pmatrix},$$

wobei die Matrizen A_{λ_i} obere Dreiecksmatrizen der Form

$$A_{\lambda_i} \begin{pmatrix} \lambda_i & * & \dots & * \\ 0 & \lambda_i & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_i \end{pmatrix} \in M(m_j, K)$$

darstellen.

Beispiel. Sei $f \in \text{End}(\mathbb{R}^3)$ mit $f(x_1, x_2, x_3) = (5x_1 + 4x_2 + 3x_3, -x_1 - 3x_3, x_1 - 2x_2 + x_3)$. f entspricht bezüglich der Basis $e_1 = (1, 0, 0), e_2 = (0, 1, 0), e_3 = (0, 0, 1)$ die Matrix

$$A = (f(e_1), f(e_2), f(e_3)) = \begin{pmatrix} 5 & 4 & 3 \\ -1 & 0 & -3 \\ 1 & -2 & 1 \end{pmatrix}.$$

Dann gilt:

$$\begin{aligned} \chi_A(\tau) &= \det \begin{pmatrix} 5 - \tau & 4 & 3 \\ -1 & -\tau & -3 \\ 1 & -2 & 1 - \tau \end{pmatrix} = \det \begin{pmatrix} 5 - \tau & 4 & 3 \\ -1 & -\tau & -3 \\ 0 & -2 - \tau & -2 - \tau \end{pmatrix} \\ &= (5 - \tau)(\tau(2 + \tau) + 3(-2 - \tau)) + 1(-8 - 4\tau + 6 + 3\tau) \\ &= (2 + \tau)((5 - \tau)(\tau - 3)) + (-2 - \tau) \\ &= (2 + \tau)(-\tau^2 + 8\tau - 15 - 1) = -(2 + \tau)(\tau - 4)^2, \end{aligned}$$

d.h. das charakteristische Polynom zerfällt in Linearfaktoren mit $\lambda_1 = -2, m_1 = 1$ und $\lambda_2 = +4, m_2 = 2$. Daraus folgt zunächst die Existenz einer f -invarianten Zerlegung

$$\mathbb{R}^3 = V_{\lambda_1} \oplus V_{\lambda_2},$$

wobei $V_{\lambda_1} = \text{Kern}(A - \lambda_1 \text{id}) = E_{\lambda_1}$ mit $\dim V_{\lambda_1} = 1$ und $V_{\lambda_2} = \text{Kern}(A - \lambda_2 \text{id})^2$ mit $\dim V_{\lambda_2} = 2$. Wegen Satz 12.5.11 existiert somit eine Basis $B = (b_1, b_2, b_3)$ von \mathbb{R}^3 mit $b_1 \in V_{\lambda_1}$ und $b_2, b_3 \in V_{\lambda_2}$, so dass $M_B(f)$ von der Gestalt

$$M_B(f) = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 4 & * \\ 0 & 0 & 4 \end{pmatrix}$$

ist. Eine Basis $B = (b_1, b_2, b_3)$ mit dieser Eigenschaft erhalten wir wie folgt:

$$V_{\lambda_1} = \text{Kern}(A - \lambda_1 \text{id}) = \text{Kern} \begin{pmatrix} 7 & 4 & 3 \\ -1 & 2 & -3 \\ 1 & -2 & 3 \end{pmatrix} = \text{Span}\{(-1, 1, 1)\}.$$

Man setze also $b_1 = (-1, 1, 1)$ und betrachte nun

$$A - \lambda_2 \text{id} = \begin{pmatrix} 1 & 4 & 3 \\ -1 & -4 & -3 \\ 1 & -2 & -3 \end{pmatrix}, \quad \text{und} \quad (A - \lambda_2 \text{id})^2 = \begin{pmatrix} 0 & -18 & -18 \\ 0 & 18 & 18 \\ 0 & 18 & 18 \end{pmatrix}.$$

Dann folgt:

$$W := V_{\lambda_2} = \text{Kern}(A - \lambda_2 \text{id})^2 = \{(x_1, x_2, x_3) \mid x_2 + x_3 = 0\} = \text{Span}\{(1, -1, 1), (0, -1, 1)\}.$$

Man betrachte nun den nilpotenten Endomorphismus $g := A - \lambda_2 \text{id} : W \rightarrow W$. Es gilt:

$$\text{Bild}(g) = \text{Span}\{g(1, -1, 1), g(0, -1, 1)\} = \text{Span}\{(-1, 1, -1)\} = W_1.$$

Man setze $b_2 = (-1, 1, -1)$ und $b_3 = (0, -1, 1)$. Da $g(b_2) = 0$ und $g(b_3) = (-1, 1, -1) = b_2$, erhalten wir für die Matrixdarstellung von $g := A - \lambda_2 \text{id} : W \rightarrow W$ bezüglich $(b_2, b_3) =: B_2$

$$M_{B_2}(g) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \text{und somit} \quad M_{B_2}(f) = \begin{pmatrix} 4 & 1 \\ 0 & 4 \end{pmatrix}.$$

Für die Basis $B = (b_1, b_2, b_3)$ von \mathbb{R}^3 ergibt sich somit:

$$M_B(A) = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 4 & 1 \\ 0 & 0 & 4 \end{pmatrix}.$$

Kapitel 13

Euklidische und unitäre Vektorräume

In diesem Kapitel werden wir für Vektorräume über \mathbb{R} und \mathbb{C} eine zusätzliche Struktur einführen, die es z.B. ermöglicht, die Länge sowie den Winkel zwischen Vektoren zu definieren. Diese Struktur wird durch das *skalare Produkt* gegeben.

13.1 Das skalare Produkt

Definition 13.1.1. Sei V ein Vektorraum über K mit $K = \mathbb{R}$ oder $K = \mathbb{C}$. Unter einem *Skalarprodukt* auf V verstehen wir eine Abbildung

$$\langle \cdot, \cdot \rangle : V \times V \rightarrow K$$

mit

(i) $\langle \cdot, \cdot \rangle$ ist linear im ersten Argument, d.h. für alle $x, x', y \in V$ und $\alpha, \beta \in K$ gilt

$$\langle \alpha x + \beta x', y \rangle = \alpha \langle x, y \rangle + \beta \langle x', y \rangle.$$

(ii) $\langle x, y \rangle = \overline{\langle y, x \rangle}$.

(iii) $\langle x, x \rangle > 0$ für $x \neq 0$.

Bemerkungen.

(a) Damit gilt für das zweite Argument:

$$\langle x, y + y' \rangle = \overline{\langle y + y', x \rangle} = \overline{\langle y, x \rangle} + \overline{\langle y', x \rangle} = \langle x, y \rangle + \langle x, y' \rangle$$

und

$$\langle x, \alpha y \rangle = \overline{\langle \alpha y, x \rangle} = \overline{\alpha \langle y, x \rangle} = \bar{\alpha} \overline{\langle y, x \rangle} = \bar{\alpha} \langle x, y \rangle.$$

Im Falle $K = \mathbb{R}$ ist $\bar{\alpha} = \alpha$ und somit ist das Skalarprodukt auch linear im zweiten Argument. Im Falle $K = \mathbb{C}$ ist das Skalarprodukt *semilinear* (*konjugiert linear*) im

zweiten Argument. Allgemein definiert man: sind V, W komplexe Vektorräume, so heißt eine Abbildung $f : V \rightarrow W$ *semilinear* (*konjugiert linear*), falls

$$f(x + y) = f(x) + f(y) \quad \text{und} \quad f(\alpha x) = \bar{\alpha} f(x)$$

für alle $x, y \in V$ und $\alpha \in \mathbb{C}$ gilt.

(b) Aus $\langle x, x \rangle = \overline{\langle x, x \rangle}$ folgt $\langle x, x \rangle \in \mathbb{R}$.

(c) Im Falle $K = \mathbb{R}$ ist wegen (ii) das Skalarprodukt $\langle \cdot, \cdot \rangle$ symmetrisch, d.h.

$$\langle x, y \rangle = \langle y, x \rangle.$$

(d) Die reelle Zahl

$$\|x\| = \sqrt{\langle x, x \rangle}$$

heißt die *Länge* von $x \in V$.

Definition 13.1.2. Sei V ein Vektorraum über K mit einem Skalarprodukt. Im Falle $K = \mathbb{R}$ heißt dann V auch *euklidischer Vektorraum*, im Falle $K = \mathbb{C}$ heißt V auch *unitärer Vektorraum*.

Beispiele.

(a) $V = K^n$. Dann heißt

$$\langle x, y \rangle = \sum_{i=1}^n x_i \bar{y}_i$$

mit $x = (x_1, \dots, x_n) \in K^n$ und $y = (y_1, \dots, y_n) \in K^n$ das *kanonische Skalarprodukt* (*Standardskalarprodukt*) auf K^n .

(b) Für $n = 2$ definiert auch

$$\langle x, y \rangle = 4x_1 \bar{y}_1 - 2x_1 \bar{y}_2 - 2x_2 \bar{y}_1 + 3x_2 \bar{y}_2$$

ein Skalarprodukt auf K^2 .

(c) Sei

$$\mathbb{C}^0([a, b], K) := \{f : [a, b] \rightarrow K \mid f \text{ stetig}\}.$$

Dann ist

$$\langle f, g \rangle = \int_a^b f(s) \cdot \overline{g(s)} ds = \int_a^b \operatorname{Re}(f(s) \overline{g(s)}) ds + i \int_a^b \operatorname{Im}(f(s) \overline{g(s)}) ds$$

ein Skalarprodukt.

13.2 Eigenschaften des Skalarproduktes

Viele Eigenschaften des Skalarproduktes gelten sowohl für euklidische als auch für unitäre Vektorräume.

Im Folgenden betrachten wir daher Vektorräume V über K für $K = \mathbb{R}$ oder \mathbb{C} und $\langle \cdot, \cdot \rangle$ sei ein festgewähltes Skalarprodukt auf V . Die erste wichtige Eigenschaft ist die *Cauchy-Schwarzsche Ungleichung*.

Satz 13.2.1. (Cauchy-Schwarz)

Sei V ein K -Vektorraum und $\langle \cdot, \cdot \rangle$ ein Skalarprodukt. Dann gilt:

$$|\langle x, y \rangle|^2 \leq \langle x, x \rangle \langle y, y \rangle.$$

Dabei gilt Gleichheit genau dann, falls $x = \lambda y$ für ein $\lambda \in K$.

Beweis. Im Falle $y = 0$ ist die Ungleichung erfüllt. Daher sei $x, y \in V$ und $y \neq 0$. Dann gilt für alle $\lambda \in K$:

$$\begin{aligned} 0 &\leq \langle x - \lambda y, x - \lambda y \rangle = \langle x, x \rangle + \langle x, -\lambda y \rangle + \langle -\lambda y, x \rangle + \langle -\lambda y, -\lambda y \rangle \\ &= \langle x, x \rangle - \bar{\lambda} \langle x, y \rangle - \lambda \langle y, x \rangle + \lambda \bar{\lambda} \langle y, y \rangle. \end{aligned}$$

Insbesondere folgt damit für $\lambda = \frac{\langle x, y \rangle}{\langle y, y \rangle}$:

$$\begin{aligned} 0 &\leq \langle x, x \rangle - \frac{\langle x, y \rangle}{\langle y, y \rangle} \langle x, y \rangle - \frac{\langle x, y \rangle}{\langle y, y \rangle} \overline{\langle x, y \rangle} + \frac{\langle x, y \rangle}{\langle y, y \rangle} \frac{\overline{\langle x, y \rangle}}{\langle y, y \rangle} \langle y, y \rangle \\ &= \langle x, x \rangle - \frac{\langle x, y \rangle}{\langle y, y \rangle} \langle x, y \rangle. \end{aligned}$$

Daraus folgt:

$$\overline{\langle x, y \rangle} \langle x, y \rangle \leq \langle x, x \rangle \langle y, y \rangle.$$

Dabei gilt Gleichheit dann und nur dann, falls $x = \lambda \cdot y$. □

Ist $(V, \langle \cdot, \cdot \rangle)$ ein K -Vektorraum mit Skalarprodukt, so definiere $\|x\| = \sqrt{\langle x, x \rangle}$.

Korollar 13.2.2. Sei $(V, \langle \cdot, \cdot \rangle)$ ein K -Vektorraum mit Skalarprodukt. Dann gilt für alle $x, y \in V$ die Dreiecksungleichung, d.h.

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

Ist $v \in V$ ein weiterer Vektor, so folgt:

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

Beweis. Sei $x, y \in V$. Dann gilt:

$$\begin{aligned} \|x + y\|^2 &= \langle x + y, x + y \rangle = \langle x, x \rangle + \langle y, x \rangle + \langle x, y \rangle + \langle y, y \rangle \\ &= \|x\|^2 + \langle x, y \rangle + \langle x, y \rangle + \|y\|^2 = \|x\|^2 + 2 \operatorname{Re} \langle x, y \rangle + \|y\|^2 \\ &\leq \|x\|^2 + 2|\langle x, y \rangle| + \|y\|^2 \leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \\ &= (\|x\| + \|y\|)^2. \end{aligned}$$

Sei v ein weiterer Vektor, so gilt:

$$\|x - y\| = \|x - v + v - y\| \leq \|x - v\| + \|v - y\|.$$

□

Die Abbildung $\|\cdot\| : V \rightarrow \mathbb{R}$ definiert eine Norm auf V .

Allgemein definiert man:

Definition 13.2.3. Sei V ein Vektorraum über K . Eine Norm $\|\cdot\|$ auf V ist eine Abbildung $V \rightarrow \mathbb{R}$ mit folgenden Eigenschaften. Für alle $x, y \in V$ und $\lambda \in K$ gilt:

- (i) $\|x\| \geq 0$ und $\|x\| = 0 \Leftrightarrow x = 0$.
- (ii) $\|\lambda x\| = |\lambda| \|x\|$.
- (iii) $\|x + y\| \leq \|x\| + \|y\|$.

Vektorräume $(V, \|\cdot\|)$ mit einer Norm $\|\cdot\|$ nennt man auch *normierte Vektorräume*.

Bemerkungen.

- (a) Ist $(V, \langle \cdot, \cdot \rangle)$ ein K -Vektorraum mit Skalarprodukt, so definiert $\|x\| = \sqrt{\langle x, x \rangle}$ eine Norm auf V . Die Eigenschaft (ii) gilt, da

$$\|\lambda x\| = \sqrt{\lambda \bar{\lambda} \langle x, x \rangle} = |\lambda| \sqrt{\langle x, x \rangle}.$$

- (b) Nicht jede Norm auf einem K -Vektorraum kommt von einem Skalarprodukt.

- (c) Ist V ein normierter Vektorraum, so nennt man

$$d(x, y) = \|x - y\|$$

den *Abstand* von x und y . Der Abstand (die *Metrik*) hat dann folgende Eigenschaften

- (i) $d(x, y) \geq 0$ und $d(x, y) = 0 \Leftrightarrow x = y$.
- (ii) $d(x, y) = d(y, x)$.
- (iii) $d(x, y) \leq d(x, z) + d(z, y)$. Räume mit dieser Eigenschaft nennen wir *metrische Räume*. Dieser Begriff benötigt nicht die Vektorraumstruktur. Z.B. ist jede Teilmenge M eines normierten Vektorraumes ein metrischer Raum, indem man d auf $M \times M$ einschränkt.

Ein Skalarprodukt ist schon durch die von ihm induzierte Norm bestimmt. Dies ist eine Konsequenz der folgenden Formeln (*Polarisationsformeln*).

Lemma 13.2.4. Sei $(V, \langle \cdot, \cdot \rangle)$ ein K -Vektorraum mit Skalarprodukt. Dann gilt im Euklidischen Falle, d.h. für $K = \mathbb{R}$:

$$\langle x, y \rangle = \frac{1}{4} (\|x + y\|^2 - \|x - y\|^2).$$

Ist $K = \mathbb{C}$, so gilt:

$$\langle x, y \rangle = \frac{1}{4} (\|x + y\|^2 - \|x - y\|^2 + i\|x + iy\|^2 - i\|x - iy\|^2).$$

Beweis. Sei $K = \mathbb{R}$ und $x, y \in V$. Dann gilt:

$$\|x + y\|^2 = \langle x + y, x + y \rangle = \langle x, x \rangle + 2\langle x, y \rangle + \langle y, y \rangle$$

und

$$\|x - y\|^2 = \langle x - y, x - y \rangle = \langle x, x \rangle - 2\langle x, y \rangle + \langle y, y \rangle.$$

Damit folgt die Behauptung aus der Subtraktion der zweiten von der ersten Gleichung.

Der Fall $K = \mathbb{C}$ sei als Übung überlassen. □

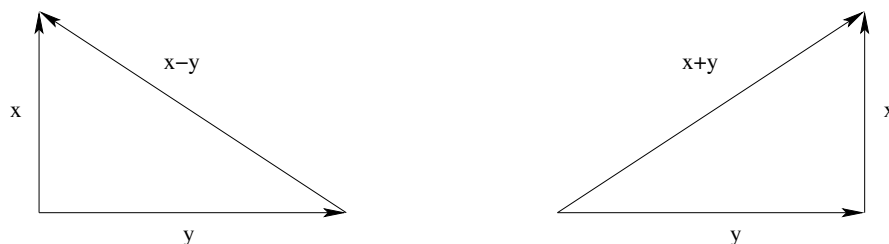


Abbildung 13.1:

Ist (V, \langle, \rangle) ein Euklidischer Vektorraum, so lässt sich der *Winkel* zwischen Vektoren wie folgt erklären.

Definition 13.2.5. Sei $x, y \in V$ mit $x, y \neq 0$, so sei $\angle(x, y) := \alpha \in [0, \pi]$ die eindeutig bestimmte reelle Zahl mit

$$\cos \alpha = \frac{\langle x, y \rangle}{\|x\| \|y\|}.$$

Bemerkung. Dies folgt aus der Cauchy-Schwarzschen Ungleichung und da $\cos : [0, \pi] \rightarrow [-1, +1]$ bijektiv ist.

Damit erhalten wir den *Cosinussatz*.

Satz 13.2.6. Es sei (V, \langle, \rangle) ein Euklidischer Vektorraum und $x, y \in V$ mit $x, y \neq 0$. Dann gilt:

$$\|x - y\|^2 = \|x\|^2 + \|y\|^2 - 2\|x\| \|y\| \cos \angle(x, y).$$

Ist insbesondere $\angle(x, y) = \frac{\pi}{2}$ ($\Leftrightarrow \langle x, y \rangle = 0$), so gilt der Satz des Pythagoras, d.h.

$$\|x - y\|^2 = \|x + y\|^2 = \|x\|^2 + \|y\|^2.$$

Beweis.

$$\begin{aligned} \|x - y\|^2 &= \langle x - y, x - y \rangle = \langle x, x \rangle - 2\langle x, y \rangle + \langle y, y \rangle \\ &= \|x\|^2 + \|y\|^2 - 2\|x\| \|y\| \frac{\langle x, y \rangle}{\|x\| \|y\|} \\ &= \|x\|^2 + \|y\|^2 - 2\|x\| \|y\| \cos \angle(x, y). \end{aligned}$$

□

Bemerkung. Der Satz des Pythagoras gilt auch für komplexe Vektorräume, d.h. aus $\langle x, y \rangle = 0$ folgt

$$\|x - y\|^2 = \|x + y\|^2 = \|x\|^2 + \|y\|^2.$$

Definition 13.2.7. Sei (V, \langle, \rangle) ein Euklidischer oder unitärer Vektorraum.

(a) Dann heißen $x, y \in V$ *orthogonal zueinander*, falls

$$\langle x, y \rangle = 0$$

gilt. (Wegen $\bar{0} = 0$ ist dies äquivalent zu $\langle y, x \rangle = 0$). Wir schreiben $x \perp y$.

(b) Es seien $U_1, U_2 \subset V$ Unterräume von V . U_1, U_2 heißen *orthogonal zueinander* ($U_1 \perp U_2$), falls $\langle x, y \rangle = 0$ für alle $x \in U_1$ und alle $y \in U_2$.

13.3 Orthonormalsysteme

Definition 13.3.1. Sei (V, \langle, \rangle) ein Vektorraum mit Skalarprodukt und $E \subset V$. E heißt

(i) *Orthogonalsystem (OG-System)*, falls

$$\langle x, y \rangle = 0$$

für $x, y \in E$ und $x \neq y$.

(ii) *Orthonormalsystem (ON-System)*, falls

$$\langle x, y \rangle = \delta_{xy} := \begin{cases} 0 & x \neq y, \\ 1 & x = y \end{cases}$$

für alle $x, y \in E$.

Satz 13.3.2. Sei V ein Vektorraum mit Skalarprodukt und $E \subset V$ ein OG-System mit $0 \notin E$. Dann ist E linear unabhängig. Insbesondere ist jedes ON-System linear unabhängig.

Beweis. Sei $\{x_1, \dots, x_n\} \subset E$ eine endliche Teilmenge von E und $\sum_{i=1}^n \alpha_i x_i = 0$ für $\alpha_i \in K$.

Dann gilt:

$$\begin{aligned} 0 &= \left\langle \sum_{i=1}^n \alpha_i x_i, \sum_{j=1}^n \alpha_j x_j \right\rangle = \sum_{i=1}^n \alpha_i \left\langle x_i, \sum_{j=1}^n \alpha_j x_j \right\rangle = \sum_{i,j=1}^n \alpha_i \bar{\alpha}_j \langle x_i, x_j \rangle \\ &= \sum_{i=1}^n \alpha_i \bar{\alpha}_i \|x_i\|^2 = \sum_{i=1}^n |\alpha_i|^2 \|x_i\|^2, \end{aligned}$$

da $\|x_i\| > 0$ für $i \in \{1, \dots, n\}$ ist $\alpha_1 = \alpha_2 = \dots = \alpha_n = 0$. □

Bemerkung. Ist also $\dim V = n$, so gibt es kein OG-System (ON-System) mit mehr als n von null verschiedenen Elementen. Jedes OG-System (ON-System) mit n von null verschiedenen Elementen ist eine Basis. Bildet eine Basis ein OG-System bzw. ON-System, so sprechen wir von einer *Orthogonalbasis (OG-Basis)* bzw. *Orthonormalbasis (ON-Basis)*.

Satz 13.3.3. Sei $E = \{v_1, \dots, v_k\} \subset V$ ein ON-System und $v \in V$. Dann folgt:

(i) Für alle $j \in \{1, \dots, k\}$ gilt die Äquivalenz:

$$v - \sum_{i=1}^k \alpha_i v_i \perp v_j \Leftrightarrow \alpha_j = \langle v, v_j \rangle.$$

$$(ii) \left\| v - \sum_{i=1}^k \alpha_i v_i \right\|^2 = \left\| v - \sum_{i=1}^k \langle v, v_i \rangle v_i \right\|^2 + \sum_{i=1}^k |\langle v, v_i \rangle - \alpha_i|^2.$$

Bemerkungen.

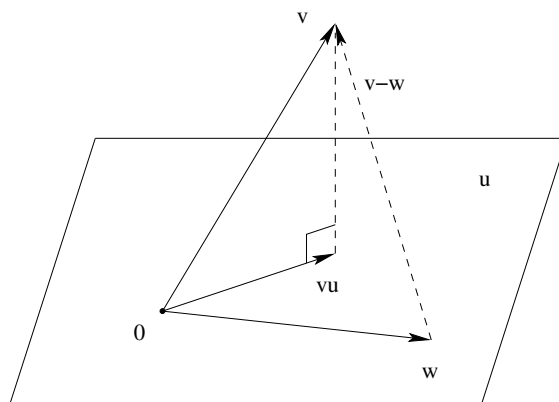


Abbildung 13.2:

(a) Aus (i) folgt: Sei $w \in \text{Span}(E) = U$, d.h. $w = \sum_{i=1}^k \alpha_i v_i$. Dann gilt: $v - w \perp E$ genau dann, falls $\alpha_i = \langle v, v_i \rangle$.

(b) Aus (ii) folgt:

$$\left\| v - \sum_{i=1}^k \langle v, v_i \rangle v_i \right\|^2 \leq \left\| v - \sum_{i=1}^k \alpha_i v_i \right\|^2.$$

Dabei gilt Gleichheit genau dann, falls $\alpha_i = \langle v, v_i \rangle$ für alle $i \in \{1, \dots, k\}$ gilt. Dies hat folgende wichtige geometrische Interpretation:

$$\min\{\|v - w\| \mid w \in \text{Span}(E)\} = \|v - v_U\|,$$

wobei $v_U = \sum_{i=1}^k \langle v, v_i \rangle v_i$. Dies bedeutet, dass der Vektor v_U hat unter allen Vektoren des Unterraumes $U = \text{Span } E$ den kleinsten Abstand von v . Er ist durch diese Eigenschaft eindeutig bestimmt.

Beweis.

(i) Es sei $E = \{v_1, \dots, v_k\} \subset V$ ein ON-System und $v \in V$. Dann gilt für jedes $j \in \{1, \dots, k\}$:

$$0 = \left\langle v - \sum_{i=1}^k \alpha_i v_i, v_j \right\rangle = \langle v, v_j \rangle - \alpha_j \Leftrightarrow \langle v, v_j \rangle = \alpha_j.$$

(ii)

$$\left\| v - \sum_{i=1}^k \alpha_i v_i \right\|^2 = \left\| v - \sum_{i=1}^k \langle v, v_i \rangle v_i + \sum_{i=1}^k \langle v, v_i \rangle v_i - \sum_{i=1}^k \alpha_i v_i \right\|^2.$$

Da $v - \sum_{i=1}^k \langle v, v_i \rangle v_i \perp \text{Span}(E)$ (wegen (i)), folgt aus dem Satz des Pythagoras

$$\begin{aligned} \left\| v - \sum_{i=1}^k \alpha_i v_i \right\|^2 &= \left\| v - \sum_{i=1}^k \langle v, v_i \rangle v_i \right\|^2 + \left\| \sum_{i=1}^k (\langle v, v_i \rangle - \alpha_i) v_i \right\|^2 \\ &= \left\| v - \sum_{i=1}^k \langle v, v_i \rangle v_i \right\|^2 + \sum_{i=1}^k |\langle v, v_i \rangle - \alpha_i|^2. \end{aligned}$$

Das letzte Gleichheitszeichen folgt, da E ein ON-System ist. □

Korollar 13.3.4. Sei $E = \{v_1, \dots, v_k\} \subset V$ ein ON-System und $v \in V$. Dann gilt:

$$\sum_{i=1}^k |\langle v, v_i \rangle|^2 \leq \|v\|^2 \quad (\text{Besselsche Ungleichung}).$$

Gleichheit gilt dann und nur dann, falls $v = \sum_{i=1}^k \langle v, v_i \rangle v_i$.

Beweis. Setzen wir in Satz 13.3.3 (ii) $\alpha_1 = \alpha_2 = \dots = \alpha_k = 0$. Dann folgt:

$$\|v\|^2 = \left\| v - \sum_{i=1}^k \langle v, v_i \rangle v_i \right\|^2 + \sum_{i=1}^k |\langle v, v_i \rangle|^2.$$

□

Bemerkung. E ist genau dann eine Basis, falls $\sum_{i=1}^k |\langle v, v_i \rangle|^2 = \|v\|^2$ für alle $v \in V$, denn dann gilt:

$$v = \sum_{i=1}^k \langle v, v_i \rangle v_i$$

für alle $v \in V$, d.h. $v \in \text{Span}(E)$.

Beispiele.

(a) Sei $V = K^n$, mit kanonischem Skalarprodukt $\langle \cdot, \cdot \rangle$. Ist $e_i = (0, \dots, 1, \dots, 0) \in K^n$, so ist

$$E = \{e_1, \dots, e_k\}$$

ein ON-System in K^n . Ist $k = n$, so ist E eine ON-Basis.

(b) Betrachte $C^0([0, 2\pi], \mathbb{C}) = V$ mit $\langle f, g \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(t) \cdot \overline{g(t)} dt$. Dann ist

$$E = \{e^{ikt} = \cos(kt) + i \sin(kt) \mid k \in \mathbb{Z}\}$$

ein ON-System in $C^0([0, 2\pi], \mathbb{C})$, denn

$$\begin{aligned}\langle e^{ikt}, e^{i\ell t} \rangle &= \frac{1}{2\pi} \int_0^{2\pi} e^{ikt} e^{-i\ell t} dt \\ &= \frac{1}{2\pi} \int_0^{2\pi} e^{i(k-\ell)t} dt = \begin{cases} 1 & \text{für } k = \ell, \\ 0 & \text{für } k \neq \ell, \end{cases}\end{aligned}$$

da $\int_0^{2\pi} e^{imt} dt := \int_0^{2\pi} \cos(mt) dt + i \int_0^{2\pi} \sin(mt) dt = 0$ für $m \neq 0$ und $m \in \mathbb{Z}$.

Betrachte den Unterraum

$$E_m = \text{Span}\{e^{ikt} \mid |k| \leq m, t \in [0, 2\pi]\}$$

Sei $f : [0, 2\pi] \rightarrow \mathbb{C}$ eine stetige Funktion, so ist

$$f_m(t) := f_{E_m}(t) = \sum_{k=-m}^m c_k e^{ikt}$$

mit $c_k = \langle f, e^{ikt} \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(t) e^{-ikt} dt$, die orthogonale Projektion von f auf den Unterraum E_m . Die Funktion $f_m(t)$ ist nach Satz 13.3.3 bezüglich der Norm die durch obiges Skalarprodukt definiert wird, unter allen $h \in \text{Span}(E_m)$ die beste Approximation von f , d.h.

$$\|f_m - f\| \leq \|h - f\|$$

für alle $h \in \text{Span}(E_m)$. Dabei gilt die Gleichheit nur im Falle $h = f_m$. Die Koeffizienten c_k nennt man auch *Fourierkoeffizienten* von f . Man kann zeigen: Ist $f : [0, 2\pi] \rightarrow \mathbb{C}$ stetig und stückweise differenzierbar und ist $f(0) = f(2\pi)$, so konvergiert die Folge f_m gleichmäßig gegen f . Wir werden diesen Satz bald beweisen.

Funktionen der Form $\sum_{k=-m}^m c_k e^{ikt}$ heißen auch *trigonometrische Polynome*.

- (c) Sei $V = C^0([0, 2\pi], \mathbb{R})$ der reelle Vektorraum der stetigen Funktionen auf $[0, 2\pi]$ mit Skalarprodukt $\langle f, g \rangle = \frac{1}{\pi} \int_0^{2\pi} f(t) \cdot g(t) dt$. Dann bilden die Funktionen

$$E = \left\{ \frac{1}{2}, \cos(kt), \sin(kt) \mid k \in \mathbb{N} \right\}$$

ein ON-System in $C^0([0, 2\pi], \mathbb{R})$.

Beweis von (c): Für $k, \ell \in \mathbb{N}$ betrachte

$$\cos(kt) = \frac{1}{2}(e^{ikt} + e^{-ikt}) \quad \text{und} \quad \sin \ell t = \frac{1}{2i}(e^{i\ell t} - e^{-i\ell t}).$$

Dann gilt:

$$\begin{aligned} \frac{1}{\pi} \int_0^{2\pi} \cos(kt) \cdot \sin(\ell t) dt &= \frac{2}{2\pi} \int_0^{2\pi} \cos(kt) \cdot \overline{\sin(\ell t)} dt \\ &= \frac{1}{2i} \frac{1}{2\pi} \int_0^{2\pi} (e^{ikt} + e^{-ikt})(e^{i\ell t} - e^{-i\ell t}) dt \\ &= \frac{1}{2i} (\delta_{k,-\ell} - \delta_{k,\ell} + \delta_{-k,-\ell} - \delta_{-k,\ell}) = 0 \end{aligned}$$

denn $\delta_{-k,-\ell} = \delta_{k,\ell}$ und $\delta_{-k,\ell} = \delta_{k,-\ell}$. Ist $k \neq \ell$, so folgt:

$$\frac{1}{\pi} \int_0^{2\pi} \cos(kt) \cos(\ell t) dt = \frac{1}{2} (\delta_{k,-\ell} + \delta_{k,\ell} + \delta_{-k,-\ell} + \delta_{-k,\ell}) = 0$$

und

$$\langle \cos(kt), \cos(kt) \rangle = \frac{1}{\pi} \int_0^{2\pi} \cos^2(kt) dt = \frac{1}{2} (\delta_{k,-k} + \delta_{k,k} + \delta_{-k,-k} + \delta_{-k,k}) = 1$$

Genauso sieht man: $\sin(kt) \perp \sin(\ell t)$ für $k \neq \ell$ und $k, \ell \in \mathbb{N}$, sowie $\langle \sin(kt), \sin(kt) \rangle = 1$.

Wir halten nochmals fest. Will man einen Vektor in einem Vektorraum mit Skalarprodukt V durch Vektoren eines Unterraumes U approximieren, so lässt sich diese Aufgabe leicht lösen, falls man in U eine ON-Basis gegeben ist. Es gibt nun ein einfaches Verfahren aus jeder Basis eines endlich dimensionalen Vektorraumes (oder abzählbar unendlich dimensionalen) eine ON-Basis zu konstruieren.

Satz 13.3.5. (Orthonormalisierungsverfahren von Gram-Schmidt)

Es sei $(V, \langle \cdot, \cdot \rangle)$ ein K -Vektorraum mit Skalarprodukt und $\{w_1, \dots, w_n\}$ eine linear unabhängige (geordnete) Teilmenge von V . Dann gibt es genau ein geordnetes ON-System $\{v_1, \dots, v_n\}$ mit

$$v_k = \sum_{j=1}^k a_{jk} w_j$$

und $a_{kk} \in \mathbb{R}, a_{kk} > 0$. Insbesondere gilt:

$$\text{Span}\{v_1, \dots, v_k\} = \text{Span}\{w_1, \dots, w_k\}.$$

Beweis. Induktion über k .

$k = 1$: setze $v_1 = \frac{w_1}{\|w_1\|}$. Damit erfüllt v_1 die verlangten Eigenschaften und wird dadurch auch eindeutig bestimmt.

Für $k - 1$ sei das ON-System $\{v_1, \dots, v_{k-1}\}$ mit den geforderten Eigenschaften gegeben und dadurch eindeutig bestimmt. Dann ist

$$\tilde{v}_k = w_k - \sum_{i=1}^{k-1} \langle w_k, v_i \rangle v_i \perp v_j$$

für $1 \leq j \leq k-1$. Dann erfüllt der normierte Vektor

$$v_k = \frac{\tilde{v}_k}{\|\tilde{v}_k\|}$$

die verlangten Eigenschaften. Der Vektor v_k ist aber dadurch auch eindeutig bestimmt, denn wäre b_k ein weiterer Vektor, so könnten wir ihn folgendermaßen darstellen:

$$b_k = \alpha_k w_k + \sum_{i=1}^{k-1} \beta_i v_i \text{ und } \alpha_k > 0.$$

Da $\langle b_k, v_j \rangle = 0$ für $j \in \{1, \dots, k-1\}$, folgt

$$0 = \alpha_k \langle w_k, v_j \rangle + \beta_j \langle v_j, v_j \rangle$$

und damit $\beta_j = -\alpha_k \langle w_k, v_j \rangle$. Also gilt:

$$b_k = \alpha_k \left(w_k - \sum_{i=1}^{k-1} \langle w_k, v_i \rangle v_i \right) = \alpha_k \tilde{v}_k.$$

Da $1 = \langle b_k, b_k \rangle = \alpha_k^2 \langle \tilde{v}_k, \tilde{v}_k \rangle$, folgt aus $\alpha_k > 0$: $\alpha_k = \frac{1}{\|\tilde{v}_k\|}$, d.h. $b_k = v_k$. □

Definition 13.3.6. Sei $(V, \langle \cdot, \cdot \rangle)$ ein K -Vektorraum mit Skalarprodukt und $M \subset V$ eine beliebige nichtleere Menge. Dann heißt

$$M^\perp := \{v \in V \mid \langle v, x \rangle = 0 \text{ für alle } x \in M\}$$

der *Orthogonalraum* zu M .

Lemma 13.3.7.

- (a) M^\perp ist ein Unterraum von V .
- (b) $M \cap M^\perp \subset \{0\}$.
- (c) $\{0\}^\perp = V$ und $V^\perp = \{0\}$.

Bemerkung. Ist M ebenfalls ein Unterraum, so gilt: $M \cap M^\perp = \{0\}$ denn der Nullvektor ist in jedem Unterraum enthalten.

Beweis.

- (a) Sei $v_1, v_2 \in M^\perp$ und $\alpha, \beta \in K$. Dann gilt:

$$\langle \alpha v_1 + \beta v_2, x \rangle = \alpha \langle v_1, x \rangle + \beta \langle v_2, x \rangle = 0$$

für alle $x \in M$, d.h. $\alpha v_1 + \beta v_2 \in M^\perp$.

- (b) Wir können annehmen, dass $M \cap M^\perp \neq \emptyset$. Sei $v \in M \cap M^\perp$. Dann folgt: $\langle v, x \rangle = 0$ für alle $x \in M$. Da $v \in M$, gilt insbesondere: $\langle v, v \rangle = 0$ und deshalb ist $v = 0$.
- (c) $\{0\}^\perp = V$, denn $\langle v, 0 \rangle = 0$ für alle $v \in V$. Da wegen (b) $V \cap V^\perp = \{0\}$, ist $V^\perp = \{0\}$.

□

Satz 13.3.8. Sei $(V, \langle \cdot, \cdot \rangle)$ Vektorraum mit Skalarprodukt und $U \subset V$ Unterraum mit $\dim U < \infty$. Dann gilt:

$$V = U \oplus U^\perp.$$

Beweis. Die Summe ist direkt, denn wegen Lemma 13.3.7 ist $U \cap U^\perp = \{0\}$. Ist $x \in V$, so betrachte die Zerlegung

$$x = x_U + (x - x_U),$$

wobei x_U die "beste Approximation" des Vektors x bezüglich U ist ($x_U = \sum_{i=1}^k \langle x, v_i \rangle v_i$, falls v_1, \dots, v_k eine ON-Basis von U ist). Wegen Satz 13.3.3 ist $x - x_U \in U^\perp$. \square

Bemerkung. Allgemein gilt: Sind $U_1, U_2 \subset V$ Unterräume mit $U_1 \perp U_2$, so ist die Summe $U_1 + U_2$ direkt, d.h. $U_1 \cap U_2 = \{0\}$.

Ein Skalarprodukt induziert eine lineare (semilineare, im Fall $K = \mathbb{C}$) Abbildung $\sigma : V \rightarrow V^*$, wobei $V^* = \{\varphi : V \rightarrow K \mid \varphi \text{ linear}\}$ den *Dualraum* von V bezeichnet.

Satz 13.3.9. Sei $(V, \langle \cdot, \cdot \rangle)$ ein K -Vektorraum mit Skalarprodukt und $\sigma : V \rightarrow V^*$ die Abbildung mit $\sigma(y) = \langle \cdot, y \rangle$, d.h. $\sigma(y)(x) = \langle x, y \rangle$. Dann ist σ eine lineare (im Falle $K = \mathbb{R}$) bzw. semilineare (im Falle $K = \mathbb{C}$) injektive Abbildung.

Beweis. $\sigma(y) \in V^*$, denn $\langle \cdot, y \rangle$ ist linear im ersten Argument. Außerdem gilt:

$$\sigma(\alpha x + \beta y) = \langle \cdot, \alpha x + \beta y \rangle = \bar{\alpha} \langle \cdot, x \rangle + \bar{\beta} \langle \cdot, y \rangle = \bar{\alpha} \sigma(x) + \bar{\beta} \sigma(y).$$

Im Falle $K = \mathbb{R}$ ist somit σ linear. Im Falle $K = \mathbb{C}$ ist σ semilinear. In jedem Fall ist σ injektiv, denn $\sigma(y_1) = \sigma(y_2)$ impliziert

$$\sigma(y_1)(x) = \langle x, y_1 \rangle = \langle x, y_2 \rangle = \sigma(y_2)(x)$$

für alle $x \in V$. Damit gilt:

$$\langle x, y_1 \rangle - \langle x, y_2 \rangle = \langle x, y_1 - y_2 \rangle = 0,$$

für alle $x \in V$. Insbesondere gilt: $\langle y_1 - y_2, y_1 - y_2 \rangle = 0$ und damit folgt $y_1 = y_2$. \square

Bemerkung. Sind insbesondere $y_1, y_2 \in V$ und gilt

$$\langle x, y_1 \rangle = \langle x, y_2 \rangle$$

für alle $x \in V$, so folgt $y_1 = y_2$. Man sagt auch: das Skalarprodukt ist *nicht ausgeartet*.

Ist V endlich dimensional, so ist σ bijektiv, d.h. jede Linearform kann durch das Skalarprodukt dargestellt werden. Genauer gilt:

Korollar 13.3.10 (Rieszscher Darstellungssatz).

Ist $\dim V < \infty$, so ist für $K = \mathbb{R}$ die Abbildung σ ein (linearer) Isomorphismus. Ist $K = \mathbb{C}$, so ist σ ein semilinearer Isomorphismus, d.h. σ ist eine bijektive semilineare Abbildung.

Beweis. Da wegen Satz 8.3.4 $\dim V = \dim V^* < \infty$, ist die injektive lineare Abbildung ein Isomorphismus (dies gilt auch für semilineare Abbildungen, denn Basen werden auf Basen abgebildet). \square

Bemerkungen.

(a) Der Satz besagt also: für alle $\varphi \in V^*$ existiert ein Vektor $R(\varphi) \in V$ mit

$$\varphi(x) = \langle x, R(\varphi) \rangle.$$

$R : V^* \rightarrow V$ ist also die Umkehrung von σ . $R(\varphi)$ kann man auch explizit angeben. Ist v_1, \dots, v_n eine ON-Basis von V , so ist

$$R(\varphi) = \sum_{i=1}^n \overline{\varphi(v_i)} v_i.$$

Dies folgt, denn für alle $x \in V$ gilt:

$$\langle x, R(\varphi) \rangle = \sum_{i=1}^n \langle x, \overline{\varphi(v_i)} v_i \rangle = \sum_{i=1}^n \varphi(v_i) \langle x, v_i \rangle = \varphi \left(\sum_{i=1}^n \langle x, v_i \rangle v_i \right) = \varphi(x).$$

Hier haben wir benutzt, dass die Koeffizienten in der Basisdarstellung $x = \sum_{i=1}^n \alpha_i v_i$ durch $\alpha_i = \langle x, v_i \rangle$ gegeben sind.

(b) Das Korollar ist für unendlich dimensionale Vektorräume falsch. Betrachte den Vektorraum $V = C^0([0, 1], \mathbb{R})$ mit dem Skalarprodukt $\langle f, g \rangle = \int_0^1 f(t)g(t)dt$. Sei $\delta : V \rightarrow \mathbb{R}$ die lineare Abbildung mit $\delta(f) = f(0)$ (δ heißt auch δ -Funktional oder *Dirac Funktional*). Dann kann δ nicht bezüglich des Skalarproduktes dargestellt werden, denn würde ein $g \in V$ existieren mit

$$\langle f, g \rangle = \delta(f) = f(0)$$

für alle $f \in V$, so wäre $\int_0^1 f(t)g(t)dt = f(0)$ für alle $f \in V$. Setze $f(t) = tg(t)$. Dann folgt: $\int_0^1 tg^2(t)dt = 0$. Aber dann ist $g \equiv 0$ und damit wäre $f(0) = 0$ für alle auf $[0, 1]$ stetigen Funktionen. Dies ist offensichtlich ein Widerspruch.

13.4 Die adjungierte Abbildung

Definition 13.4.1. Es seien $(V, \langle \cdot, \cdot \rangle_V)$ und $(W, \langle \cdot, \cdot \rangle_W)$ Vektorräume über K ($K = \mathbb{R}$ oder \mathbb{C}) mit Skalarprodukten und $f : V \rightarrow W$ eine lineare Abbildung. Dann heißt eine Abbildung $f^* : W \rightarrow V$ *adjungiert* zu f , falls

$$\langle f(x), y \rangle_W = \langle x, f^*(y) \rangle_V$$

für alle $x \in V$ und $y \in W$.

Lemma 13.4.2. *Existiert eine adjungierte Abbildung, so ist sie eindeutig bestimmt und linear.*

Beweis. Es seien $f_1^*, f_2^* : W \rightarrow V$ zwei Abbildungen adjungiert zu f und $y \in W$. Dann gilt:

$$\langle x, f_1^*(y) \rangle_V = \langle x, f_2^*(y) \rangle_V$$

für alle $x \in V$ und aus Satz 13.3.9 folgt $f_1^*(y) = f_2^*(y)$. Existiert die adjungierte Abbildung $f^* : W \rightarrow V$, so ist f^* automatisch linear, denn für alle $x \in V$ gilt:

$$\begin{aligned} \langle x, f^*(\alpha y_1 + \beta y_2) \rangle_V &= \langle f(x), \alpha y_1 + \beta y_2 \rangle_V = \bar{\alpha} \langle f(x), y_1 \rangle + \bar{\beta} \langle f(x), y_2 \rangle_V \\ &= \bar{\alpha} \langle x, f^*(y_1) \rangle_V + \bar{\beta} \langle x, f^*(y_2) \rangle_V \\ &= \langle x, \alpha f^*(y_1) + \beta f^*(y_2) \rangle_V \end{aligned}$$

Also folgt wieder aus Satz 13.3.9

$$f^*(\alpha y_1 + \beta y_2) = \alpha f^*(y_1) + \beta f^*(y_2)$$

für alle $\alpha, \beta \in K$ und $y_1, y_2 \in W$. □

Ist V endlich dimensional, so existiert die adjungierte Abbildung. Sie hängt eng mit der transponierten Abbildung zusammen (s. Definition 8.3.5).

Satz 13.4.3. *Es seien $(V, \langle \cdot, \cdot \rangle_V)$ und $(W, \langle \cdot, \cdot \rangle_W)$ K -Vektorräume mit Skalarprodukten und $f : V \rightarrow W$ eine lineare Abbildung. Ist $\dim V < \infty$, so existiert die adjungierte Abbildung $f^* : W \rightarrow V$.*

Beweis. Es sei $f^t : W^* \rightarrow V^*$ mit $f^t(\varphi) = \varphi \circ f : V \rightarrow K$ die transponierte Abbildung und $\sigma_W : W \rightarrow W^*$ sowie $\sigma_V : V \rightarrow V^*$ die Abbildungen mit $\sigma_W(y) = \langle \cdot, y \rangle_W$ und $\sigma_V(y) = \langle \cdot, y \rangle_V$. Wir zeigen:

$$f^* = g := \sigma_V^{-1} \circ f^t \circ \sigma_W.$$

Insbesondere ist g die Abbildung von $W \rightarrow V$ ist, so dass das folgende Diagramm kommutiert.

$$\begin{array}{ccc} W^* & \xrightarrow{f^t} & V^* \\ \sigma_W \uparrow & & \uparrow \sigma_V \\ W & \xrightarrow{f^*} & V \end{array}$$

Zunächst bemerken wir: Ist $\varphi \in V^*$, so ist $\sigma_V^{-1}(\varphi) = y$ gleichbedeutend mit $\varphi = \sigma_V(y) = \langle \cdot, y \rangle_V$, d.h. für alle $x \in V$ gilt

$$\varphi(x) = \langle x, y \rangle_V = \langle x, \sigma_V^{-1}(\varphi) \rangle_V.$$

Dann folgt

$$\langle x, g(y) \rangle_V = \langle x, \sigma_V^{-1}(\underbrace{f^t \circ \sigma_W(y)}_{\varphi}) \rangle_V = f^t \circ \sigma_W(y)(x) = \sigma_W(y)(f(x)) = \langle f(x), y \rangle_W$$

und somit ist $g = f^*$. □

Bemerkung. Man kann den Existenzbeweis auch ohne die transponierte Abbildung führen. Sei v_1, \dots, v_n eine ON-Basis von V , so definieren wir

$$f^*(y) = \sum_{i=1}^n \langle y, f(v_i) \rangle_W v_i.$$

Dann gilt:

$$\langle f^*(y), x \rangle_V = \sum_{i=1}^n \langle y, f(v_i) \rangle_W \langle v_i, x \rangle_V.$$

Für $x = \sum_{i=1}^n \langle x, v_i \rangle_V v_i$ ist $f(x) = \sum_{i=1}^n \langle x, v_i \rangle_V f(v_i)$ und damit folgt:

$$\langle y, f(x) \rangle_W = \sum_{i=1}^n \langle y, \langle x, v_i \rangle_V f(v_i) \rangle_W = \sum_{i=1}^n \overline{\langle x, v_i \rangle_V} \cdot \langle y, f(v_i) \rangle_W.$$

Die Abbildung $*$ besitzt die folgenden Eigenschaften:

Satz 13.4.4. *Es seien $(V, \langle \cdot, \cdot \rangle_V)$, $(W, \langle \cdot, \cdot \rangle_W)$ und $(Z, \langle \cdot, \cdot \rangle_Z)$ endlich dimensionale K -Vektorräume mit Skalarprodukten. Dann gilt:*

- (a) $(f^*)^* = f$.
- (b) $(\alpha f)^* = \bar{\alpha} f^*$ für $\alpha \in K$.
- (c) $(f + g)^* = f^* + g^*$.
- (d) Sind $f \in L(V, W)$, $g \in L(W, Z)$, so gilt: $(g \circ f)^* = f^* \circ g^*$.
- (e) Für $\text{id}_V \in L(V, V)$ ist $\text{id}_V^* = \text{id}_V$. Aus (d) folgt damit:
- (f) Sei $f \in L(V, W)$ bijektiv, dann ist auch $f^* \in L(W, V)$ bijektiv mit $(f^*)^{-1} = (f^{-1})^*$.

Bemerkung. (b) und (c) besagen: $*$: $L(V, W) \rightarrow L(W, V)$ mit $f \mapsto f^*$ ist semilinear. Im Falle $K = \mathbb{R}$ ist $*$ eine lineare Abbildung.

Beweis.

- (a) $\langle (f^*)^*(x), y \rangle_W = \langle x, f^*(y) \rangle_V = \langle f(x), y \rangle_W$.
- (b) $\langle (\alpha f)^*(y), x \rangle_V = \langle y, \alpha f(x) \rangle_W = \bar{\alpha} \langle y, f(x) \rangle_W = \langle \bar{\alpha} f^*(y), x \rangle_V$.

Der Rest sei als Übung überlassen. □

Satz 13.4.5. *Es seien $(V, \langle \cdot, \cdot \rangle_V)$ und $(W, \langle \cdot, \cdot \rangle_W)$ endlich dimensionale Vektorräume mit Skalarprodukten und $f : V \rightarrow W$ linear. Dann gilt*

$$\text{Kern } f = (\text{Bild } f^*)^\perp.$$

Insbesondere sind die Summen

$$V = \text{Kern } f \oplus \text{Bild } f^* \quad \text{und} \quad W = \text{Kern } f^* \oplus \text{Bild } f.$$

direkt und orthogonal.

Beweis. Es gilt : $\text{Kern } f = (\text{Bild } f^*)^\perp$, denn

$$\begin{aligned} x \in \text{Kern } f &\Leftrightarrow f(x) = 0 \Leftrightarrow \langle f(x), y \rangle_W = 0 \text{ für alle } y \in W \\ &\Leftrightarrow \langle x, f^*(y) \rangle_V = 0 \text{ für alle } y \in W \Leftrightarrow x \in (\text{Bild } f^*)^\perp. \end{aligned}$$

Damit folgt aus Satz 13.3.8

$$V = (\text{Bild } f^*)^\perp \oplus \text{Bild } f^* = \text{Kern } f \oplus \text{Bild } f^*.$$

Setzen wir für f die adjungierte Abbildung f^* , so erhalten wir wegen $(f^*)^* = f$ auch die zweite Zerlegung. □

Bemerkung. Es folgt insbesondere: (f ist injektiv $\Leftrightarrow f^*$ surjektiv) und (f^* injektiv $\Leftrightarrow f$ surjektiv).

Korollar 13.4.6. *Es gilt:*

$$\text{Rang } f = \text{Rang } f^*.$$

Beweis. $\text{Rang } f^* = \dim \text{Bild } f^* = \dim V - \dim \text{Kern } f = \dim \text{Bild } f = \text{Rang } f.$ \square

Wie sehen die Matrixdarstellungen der zu f adjungierten Abbildung aus? Verwendet man ON-Basen, so gilt:

Satz 13.4.7. *Es seien $(V, \langle \cdot, \cdot \rangle_V)$ und $(W, \langle \cdot, \cdot \rangle_W)$ Vektorräume mit Skalarprodukten und $B \subset V$ und $C \subset W$ ON-Basen. Ist $f : V \rightarrow W$ eine lineare Abbildung und $A = M_{B,C}(f)$ die Matrixdarstellung von f bezüglich B und C , so gilt:*

$$M_{C,B}(f^*) = A^*$$

mit $A^* = \bar{A}^t$, wobei $(\bar{A})_{ij} = \bar{a}_{ij}$ die komplex konjugierte Matrix bezeichnet.

Bemerkung. Im reellen Fall kann die komplexe Konjugation natürlich wegfallen, d.h. A^* ist nichts anderes als die transponierte Matrix. Wir nennen A^* die zu A adjungierte Matrix.

Beweis. Es seien $B = (b_1, \dots, b_n)$ und $C = (c_1, \dots, c_m)$ die geordneten ON-Basen bezüglich B und C . Es sei $A = M_{B,C}(f)$, d.h. (s. Satz 10.2.2)

$$f(b_j) = \sum_{i=1}^m a_{ij} c_i$$

für $j \in \{1, \dots, n\}$. Da C eine ON-Basis von W ist, folgt:

$$a_{kj} = \langle f(b_j), c_k \rangle$$

und somit

$$a_{kj} = \langle b_j, f^*(c_k) \rangle = \overline{\langle f^*(c_k), b_j \rangle}.$$

Ist $D = M_{C,B}(f^*)$ Matrixdarstellung von f^* bezüglich C und B , d.h.

$$f^*(c_k) = \sum_{i=1}^n d_{ik} b_i,$$

so folgt aus

$$d_{jk} = \langle f^*(c_k), b_j \rangle = \bar{a}_{kj}$$

die Behauptung. \square

Bemerkung. Ist insbesondere $V = K^n$ und $W = K^m$ mit $K = \mathbb{R}$ oder \mathbb{C} und $\langle x, y \rangle_{K^n} = \sum_{i=1}^n x_i \bar{y}_i$ das kanonische Skalarprodukt und $A : K^n \rightarrow K^m$ ($A \in M(m \times n, K)$) eine lineare Abbildung, so gilt:

$$\langle Ax, y \rangle_{K^m} = \langle x, A^* y \rangle_{K^n}$$

mit $A^* = \bar{A}^t$, denn die kanonische Basis $E_n = (e_1, \dots, e_n)$ ist eine ON-Basis von K^n (E_m ist eine ON-Basis von K^m). Daher gilt:

$$A^* = M_{E_m, E_n}(A^*) = (M_{E_n, E_m}(A))^* = \bar{A}^t.$$

Zum Schluß dieses Abschnittes wollen wir noch eine wichtige Klasse von Abbildungen zwischen Vektorräumen mit Skalarprodukten betrachten, nämlich solche, die das Skalarprodukt invariant lassen.

Definition 13.4.8. Es seien $(V, \langle \cdot, \cdot \rangle_V)$ und $(W, \langle \cdot, \cdot \rangle_W)$ K -Vektorräume mit Skalarprodukten. Eine lineare Abbildung $f : V \rightarrow W$ heißt *isometrische Einbettung*, falls

$$\langle f(x), f(y) \rangle_W = \langle x, y \rangle_V$$

für alle $x, y \in V$. Ist f ein Isomorphismus, so heißt f eine *Isometrie*. Falls eine Isometrie zwischen V und W existiert, so heißen V und W *isometrisch*.

Satz 13.4.9. *Ist $f : V \rightarrow W$ eine isometrische Einbettung, so ist f injektiv. Existiert die adjungierte Abbildung $f^* : W \rightarrow V$, so gilt:*

$$f^* \circ f = \text{id}_V.$$

Ist insbesondere $\dim V = \dim W < \infty$, so ist f eine Isometrie mit $f^ = f^{-1}$.*

Beweis. f ist injektiv, denn aus $f(x) = 0$ folgt:

$$0 = \langle f(x), f(x) \rangle_W = \langle x, x \rangle_V$$

und daher ist $x = 0$.

Existiert die adjungierte Abbildung $f^* : W \rightarrow V$, so folgt für alle $x, y \in V$:

$$\langle x, f^* \circ f(y) \rangle_V = \langle f(x), f(y) \rangle_W = \langle x, y \rangle_V.$$

Somit ist $f^* \circ f = \text{id}_V$.

Ist $\dim V = \dim W < \infty$, so folgt aus Satz 13.4.3 die Existenz von f^* . Aus der Injektivität von f erhalten wir, dass f ein Isomorphismus und somit eine Isometrie ist. Insbesondere folgt aus $f^* \circ f = \text{id}_V$ auch $f^* = f^{-1}$. □

Bemerkungen.

- (a) Die Verknüpfung von Isometrien ist wieder eine Isometrie.
- (b) Ist f eine Isometrie, so ist auch f^{-1} eine Isometrie, denn

$$\langle f^{-1} \circ f(x), f^{-1} \circ f(y) \rangle = \langle x, y \rangle = \langle f(x), f(y) \rangle.$$

Satz 13.4.10. *Sei $(V, \langle \cdot, \cdot \rangle)$ ein K -Vektorraum mit Skalarprodukt und $\dim V = n$. Dann ist $(V, \langle \cdot, \cdot \rangle)$ isometrisch zu $(K^n, \langle \cdot, \cdot \rangle_{K^n})$, wobei $\langle \cdot, \cdot \rangle_{K^n}$ das Standardskalarprodukt auf K^n bezeichnet.*

Beweis. Wähle eine ON-Basis $B = (v_1, \dots, v_n)$ in V . Dann ist die Koordinatenabbildung $\phi_B : V \rightarrow K^n$ mit $\phi_B(x) = (x_1, \dots, x_n)$ und $x = \sum_{i=1}^n x_i v_i$, eine Isometrie, denn

$$\langle \phi_B(x), \phi_B(y) \rangle_{K^n} = \sum_{i=1}^n x_i \bar{y}_i$$

und

$$\langle x, y \rangle_V = \left\langle \sum_{i=1}^n x_i v_i, \sum_{j=1}^n y_j v_j \right\rangle_V = \sum_{i,j=1}^n x_i \bar{y}_j \langle v_i, v_j \rangle_V = \sum_{i=1}^n x_i \bar{y}_i.$$

Außerdem ist ϕ_B ein Isomorphismus und somit eine Isometrie. □

Korollar 13.4.11. *Sind V und W K -Vektorräume mit Skalarprodukten und ist $\dim V = \dim W < \infty$, so sind V und W isometrisch.*

Beweis. Ist $n = \dim V = \dim W$, so gibt es Isometrien $\phi : V \rightarrow K^n$ und $\psi : K^n \rightarrow W$. Dann ist die Verknüpfung $\psi \circ \phi : V \rightarrow W$ die gesuchte Isometrie. \square

13.5 Normale Endomorphismen

Im Folgenden betrachten wir endlich dimensionale Vektorräume mit Skalarprodukt. Eine wichtige Klasse von Endomorphismen solcher Vektorräume bilden die normalen Endomorphismen. In unitären Vektorräumen (d.h. für komplexe Vektorräume mit Skalarprodukt) entspricht, wie wir sehen werden, diese Klasse genau den Endomorphismen, die eine ON-Basis aus Eigenvektoren besitzen. Insbesondere sind in unitären Vektorräumen normale Endomorphismen diagonalisierbar.

Definition 13.5.1. Sei $(V, \langle \cdot, \cdot \rangle)$ endlich dimensionaler Vektorraum mit Skalarprodukt. Ein Endomorphismus $f : V \rightarrow V$ heißt *normal*, falls

$$f \circ f^* = f^* \circ f.$$

Bemerkung. f ist genau dann normal, falls

$$\langle f(x), f(y) \rangle = \langle f^*(x), f^*(y) \rangle$$

für alle $x, y \in V$, denn aus $(f^*)^* = f$ folgt:

$$\langle f(x), f(y) \rangle = \langle f^*(x), f^*(y) \rangle \Leftrightarrow \langle x, f^* \circ f(y) \rangle = \langle x, f \circ f^*(y) \rangle.$$

Die letzte Gleichung ist dazu äquivalent, dass f eine normale Abbildung ist.

Wichtige Spezialfälle von normalen Endomorphismen sind die folgenden Abbildungen:

Definition 13.5.2. Sei $(V, \langle \cdot, \cdot \rangle)$ ein K -Vektorraum mit Skalarprodukt. Ein Endomorphismus $f : V \rightarrow V$ heißt

1. *selbstadjungiert*, falls $f^* = f$.
Ist $K = \mathbb{R}$, so nennt man f auch *symmetrisch*.
Ist $K = \mathbb{C}$, so nennt man f auch *hermitesch*.
2. *schiefadjungiert*, falls $f^* = -f$.
Ist $K = \mathbb{R}$, so nennt man f auch *schiefsymmetrisch*.
Ist $K = \mathbb{C}$, so nennt man f auch *schiefhermitesch*.
3. *Isometrie*, falls $f^* = f^{-1}$.
Ist $K = \mathbb{R}$, so nennt man f auch *orthogonal*.
Ist $K = \mathbb{C}$, so nennt man f auch *unitär*.

Bemerkungen.

- (a) 1, 2, 3 definieren normale Endomorphismen.

(b) Wie wir in Satz 13.4.9 gesehen haben, ist $f^* = f^{-1}$ äquivalent zu

$$\langle f(x), f(y) \rangle = \langle x, y \rangle$$

für alle $x, y \in V$, d.h. die Definition 13.4.8 stimmt in dem Spezialfall eines Endomorphismus überein.

(c) Ist B eine ON-Basis von V , so ist wegen Satz 13.4.7 $M_B(f^*) = \overline{M_B(f)}^t$.

(i) Daher ist $f : V \rightarrow V$ genau dann selbstadjungiert ($f^* = f$), falls

$$M_B(f) = \overline{M_B(f)}^t$$

gilt. Matrizen $A \in M(n, \mathbb{R})$ bzw. $A \in M(n, \mathbb{C})$ mit $A = A^t$ bzw. $A = \bar{A}^t$ nennt man *symmetrisch* bzw. *hermitesch*.

(ii) $f : V \rightarrow V$ ist genau dann schiefadjungiert ($f^* = -f$), falls

$$-M_B(f) = \overline{M_B(f)}^t$$

gilt. Matrizen $A \in M(n, \mathbb{R})$ bzw. $A \in M(n, \mathbb{C})$ mit $-A = A^t$ bzw. $A = \bar{A}^t$ nennt man *schief-symmetrisch* bzw. *schiefhermitesch*.

(iii) $f : V \rightarrow V$ ist genau dann eine Isometrie ($f^* = f^{-1}$), falls

$$M_B(f)^{-1} = \overline{M_B(f)}^t$$

gilt. Matrizen $A \in M(n, \mathbb{R})$ bzw. $A \in M(n, \mathbb{C})$ mit $A^{-1} = A^t$ bzw. $A^{-1} = \bar{A}^t$ nennt man orthogonal bzw. unitär. Insbesondere ist $A \in M(n, \mathbb{R})$ bzw. $A \in M(n, \mathbb{C})$ genau dann *orthogonal* bzw. *unitär*, falls ihre Spaltenvektoren bezüglich der Standardskalarprodukte des \mathbb{R}^n bzw. \mathbb{C}^n eine ON-Basis bilden.

Lemma 13.5.3. *Ist $f : V \rightarrow V$ normal, so auch $f + \mu \text{id}$, für alle $\mu \in K$.*

Beweis. Es gilt: $(f + \mu \text{id})^* = f^* + \bar{\mu} \text{id}$. Dann folgt:

$$(f^* + \bar{\mu} \text{id})(f + \mu \text{id}) = f^*(f + \mu \text{id}) + \bar{\mu}(f + \mu \text{id}) =$$

$$(f + \mu \text{id})f^* + (f + \mu \text{id})\bar{\mu} \text{id} = (f + \mu \text{id})(f^* + \bar{\mu} \text{id}).$$

□

Korollar 13.5.4. *Es sei f ein normaler Endomorphismus. Dann ist v ein Eigenvektor zum Eigenwert λ von f genau dann, wenn v ein Eigenvektor zum Eigenwert $\bar{\lambda}$ von f^* ist. Insbesondere haben f und f^* die gleichen Eigenvektoren.*

Beweis. Man betrachte den normalen Endomorphismus $f + \mu \text{id}$. Dann gilt:

$$\langle (f + \mu \text{id})x, (f + \mu \text{id})x \rangle = \langle (f^* + \bar{\mu} \text{id})x, (f^* + \bar{\mu} \text{id})x \rangle.$$

Damit folgt: $v \in E_\lambda(f) \Leftrightarrow (f - \lambda \text{id})v = 0 \Leftrightarrow (f^* - \bar{\lambda} \text{id})v = 0 \Leftrightarrow v \in E_{\bar{\lambda}}(f^*)$.

□

Bemerkung. Daraus folgt unmittelbar:

- (a) Ist f selbstadjungiert ($f^* = f$), so sind alle Eigenwerte reell. Denn ist v Eigenvektor zum Eigenwert λ von f , so gilt:

$$\lambda v = f(v) = f^*(v) = \bar{\lambda}v.$$

- (b) Ist f schiefadjungiert ($f^* = -f$), so sind alle Eigenwerte rein imaginär, d.h. reelle Vielfache von i , denn ist v Eigenvektor zum Eigenwert λ von f , so gilt:

$$-\lambda v = -f(v) = f^*(v) = \bar{\lambda}v,$$

und somit ist $\lambda + \bar{\lambda} = 0$.

- (c) Ist f eine Isometrie ($f^* = f^{-1}$), so liegen alle Eigenwerte λ auf dem Einheitskreis, denn ist v Eigenvektor zum Eigenwert λ von f , so gilt:

$$f(v) = \lambda v \Rightarrow v = \lambda f^{-1}(v) = \lambda f^*(v) = \lambda \bar{\lambda}v,$$

und somit ist $\lambda \bar{\lambda} = 1$.

Das folgende Lemma beschreibt eine wichtige Eigenschaft des adjungierten Endomorphismus.

Lemma 13.5.5. Sei $(V, \langle \cdot, \cdot \rangle)$ ein K -Vektorraum mit Skalarprodukt und $f : V \rightarrow V$ ein Endomorphismus. Ist U invariant unter f , so ist U^\perp invariant unter f^* . Ist U invariant unter f^* , so ist U^\perp invariant unter f .

Beweis. Sei U invariant unter f . Ist $x \in U^\perp$, so gilt für alle $y \in U$:

$$\langle f^*(x), y \rangle = \langle x, f(y) \rangle = 0,$$

d.h. $f^*(x) \in U^\perp$. Die zweite Aussage folgt aus der ersten, denn $(f^*)^* = f$. \square

Satz 13.5.6. Sei $(V, \langle \cdot, \cdot \rangle)$ ein unitärer Vektorraum (d.h. $K = \mathbb{C}$) und $f : V \rightarrow V$ ein Endomorphismus. Dann sind äquivalent:

1. f ist normal.
2. V besitzt eine ON-Basis aus Eigenvektoren von f .

Bemerkung. Insbesondere sind normale Endomorphismen in unitären Vektorräumen diagonalisierbar.

Beweis. Der Beweis wird durch Induktion über $n = \dim V$ geführt. Sei $\dim V = 1$. Dann ist jeder Endomorphismus f normal, denn es gilt: $f = \lambda \text{id}$ für $\lambda \in \mathbb{C}$. Ist $\text{Span}\{v\} = V$ mit $\|v\| = 1$, so ist v eine ON-Basis, bestehend aus einem Eigenvektor von f und somit ist 2. erfüllt.

Es sei nun $n > 1$ und die Behauptung sei für $\dim V = n - 1$ bewiesen. Sei $\dim V = n$. Da das charakteristische Polynom über \mathbb{C} in Linearfaktoren zerfällt, existiert ein Eigenwert λ_1 und ein Eigenvektor v_1 mit $\|v_1\| = 1$. Sei $W = \text{Span}\{v_1\}$, so ist W f - und f^* -invarianter Unterraum (Korollar 13.5.4). Wegen Lemma 13.5.5 ist dann W^\perp f - und auch f^* -invariant. Damit ist für $U = W^\perp$ $f|_U : U \rightarrow U$ eine normale Abbildung und da $\dim U = n - 1$, gibt es

nach Induktionsvoraussetzung eine ON-Basis $\{v_2, \dots, v_n\}$ bestehend aus Eigenvektoren von f . Sei umgekehrt $B = \{v_1, \dots, v_n\}$ eine ON-Basis, bestehend aus Eigenvektoren von f . Dann ist

$$M_B(f) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \in M(n, \mathbb{C})$$

eine Diagonalmatrix und da B eine ON-Basis ist, gilt wegen Satz 13.4.7

$$M_B(f^*) = M_B(f)^* := \overline{M_B(f)}^t = \begin{pmatrix} \bar{\lambda}_1 & & 0 \\ & \ddots & \\ 0 & & \bar{\lambda}_n \end{pmatrix}.$$

Dann folgt:

$$M_B(f \circ f^*) = M_B(f^* \circ f) = \begin{pmatrix} \lambda_1 \bar{\lambda}_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \bar{\lambda}_n \end{pmatrix},$$

und damit gilt: $f \circ f^* = f^* \circ f$, d.h. f ist normal. \square

Nun wollen wir eine Normalform für reelle normale Endomorphismen herleiten. Da das charakteristische Polynom eines reellen Endomorphismus über \mathbb{R} im Allgemeinen nicht in Linearfaktoren zerfällt, sind die Beweismethoden von Satz 13.5.6 nicht direkt anwendbar. Die Lösung besteht darin, den Endomorphismus zu komplexifizieren.

Sei V ein reeller Vektorraum und $V_{\mathbb{C}} = \{x + iy \mid x, y \in V\}$ die *Komplexifizierung* von V . Die Multiplikation mit komplexen Skalaren $\alpha + i\beta \in \mathbb{C}$ ist durch

$$(\alpha + i\beta)(x + iy) := \alpha x - \beta y + i(\beta x + \alpha y)$$

auf $V_{\mathbb{C}}$ erklärt (s. Übungsblatt Nr. 2, Mathematik für Physiker III). Ist $f : V \rightarrow V$ ein Endomorphismus, so definiert $f_{\mathbb{C}} : V_{\mathbb{C}} \rightarrow V_{\mathbb{C}}$ mit

$$f_{\mathbb{C}}(x + iy) := f(x) + if(y)$$

einen Endomorphismus, genannt die *Komplexifizierung* von f . Ist B eine Basis von V , so ist $B \subset V_{\mathbb{C}}$ ebenfalls eine Basis von $V_{\mathbb{C}}$. Insbesondere gilt: $M_B(f) = M_B(f_{\mathbb{C}})$ und daher stimmt das charakteristische Polynom von f mit dem von $f_{\mathbb{C}}$ überein. Sind $f, g : V \rightarrow V$ Endomorphismen, so folgt:

$$(f \circ g)_{\mathbb{C}} = f_{\mathbb{C}} \circ g_{\mathbb{C}}, \quad (*)$$

denn

$$(f \circ g)_{\mathbb{C}}(x + iy) = (f \circ g)(x) + i(f \circ g)(y) = f_{\mathbb{C}}(g(x) + ig(y)) = (f_{\mathbb{C}} \circ g_{\mathbb{C}})(x + iy).$$

Bezeichnet $\bar{\cdot} : V_{\mathbb{C}} \rightarrow V_{\mathbb{C}}$ mit $\overline{x + iy} = x - iy$ die komplexe Konjugation, so gilt:

$$f_{\mathbb{C}}(\overline{x + iy}) = \overline{f_{\mathbb{C}}(x + iy)}.$$

Ist insbesondere $v = v_1 + iv_2$ ein Eigenvektor von $f_{\mathbb{C}}$ zum Eigenwert $\lambda \in \mathbb{C}$, so ist $\bar{v} = v_1 - iv_2$ Eigenvektor zum Eigenwert $\bar{\lambda} \in \mathbb{C}$, denn $f_{\mathbb{C}}(\bar{v}) = \overline{f_{\mathbb{C}}(v)} = \bar{\lambda}\bar{v}$.

Ist V ein Euklidischer Vektorraum, d.h. ein reeller Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$, so lässt es sich zu einem Skalarprodukt auf $V_{\mathbb{C}}$ fortsetzen.

Satz 13.5.7. *Sei $(V, \langle \cdot, \cdot \rangle)$ ein Euklidischer Vektorraum und $V_{\mathbb{C}}$ die Komplexifizierung von V . Dann gibt es genau ein Skalarprodukt $\langle \cdot, \cdot \rangle_{\mathbb{C}}$ auf $V_{\mathbb{C}}$ mit*

$$\langle x, y \rangle = \langle x, y \rangle_{\mathbb{C}}$$

für alle $x, y \in V$. Außerdem ist

$$\|x + iy\|_{\mathbb{C}}^2 = \langle x + iy, x + iy \rangle_{\mathbb{C}} = \langle x, x \rangle + \langle y, y \rangle,$$

für alle $x, y \in V$.

Beweis. Zunächst beweisen wir die Eindeutigkeit. Falls $\langle \cdot, \cdot \rangle_{\mathbb{C}}$ existiert und $x_1 + iy_1, x_2 + iy_2 \in V_{\mathbb{C}}$, folgt:

$$\begin{aligned} \langle x_1 + iy_1, x_2 + iy_2 \rangle_{\mathbb{C}} &= \langle x_1, x_2 \rangle_{\mathbb{C}} + \langle x_1, iy_2 \rangle_{\mathbb{C}} + \langle iy_1, x_2 \rangle_{\mathbb{C}} + \langle iy_1, iy_2 \rangle_{\mathbb{C}} \\ &= \langle x_1, x_2 \rangle - i\langle x_1, y_2 \rangle + i\langle y_1, x_2 \rangle + i(-i)\langle y_1, y_2 \rangle \\ &= \langle x_1, x_2 \rangle + \langle y_1, y_2 \rangle + i(\langle y_1, x_2 \rangle - \langle x_1, y_2 \rangle). \end{aligned}$$

Auf der anderen Seite definiert die letzte Zeile ein Skalarprodukt auf $V_{\mathbb{C}}$, denn wie man leicht nachrechnet ist $\langle \cdot, \cdot \rangle_{\mathbb{C}}$ mit

$$\langle x_1 + iy_1, x_2 + iy_2 \rangle_{\mathbb{C}} := \langle x_1, x_2 \rangle + \langle y_1, y_2 \rangle + i(\langle y_1, x_2 \rangle - \langle x_1, y_2 \rangle)$$

linear im ersten Argument und es gilt:

$$\langle x_1 + iy_1, x_2 + iy_2 \rangle_{\mathbb{C}} = \overline{\langle x_2 + iy_2, x_1 + iy_1 \rangle_{\mathbb{C}}}.$$

Außerdem ist

$$\langle x + iy, x + iy \rangle_{\mathbb{C}} = \langle x, x \rangle + \langle y, y \rangle,$$

also ist $\langle \cdot, \cdot \rangle_{\mathbb{C}}$ positiv definit. □

Lemma 13.5.8. *Ist $(V, \langle \cdot, \cdot \rangle)$ ein Euklidischer Vektorraum und $f : V \rightarrow V$ ein Endomorphismus, so gilt:*

$$(f_{\mathbb{C}})^* = (f^*)_{\mathbb{C}},$$

d.h. die adjungierte Abbildung des komplexifizierten Endomorphismus ist die Komplexifizierung des adjungierten Endomorphismus.

Beweis. Es gilt: $(f^*)_{\mathbb{C}}(x + iy) = f^*(x) + if^*(y)$. Zu zeigen ist:

$$\langle (f^*)_{\mathbb{C}}(x_1 + iy_1), x_2 + iy_2 \rangle_{\mathbb{C}} = \langle x_1 + iy_1, f_{\mathbb{C}}(x_2 + iy_2) \rangle_{\mathbb{C}}.$$

In der Tat ist

$$\begin{aligned} \langle x_1 + iy_1, f_{\mathbb{C}}(x_2 + iy_2) \rangle_{\mathbb{C}} &= \langle x_1 + iy_1, f(x_2) + if(y_2) \rangle_{\mathbb{C}} \\ &= \langle x_1, f(x_2) \rangle + \langle y_1, f(y_2) \rangle + i(\langle y_1, f(x_2) \rangle - \langle x_1, f(y_2) \rangle) \\ &= \langle f^*(x_1), x_2 \rangle + \langle f^*(y_1), y_2 \rangle + i(\langle f^*(y_1), x_2 \rangle - \langle f^*(x_1), y_2 \rangle) \\ &= \langle f^*(x_1) + if^*(y_1), x_2 + iy_2 \rangle_{\mathbb{C}}. \end{aligned}$$

□

Korollar 13.5.9. Sei $(V, \langle \cdot, \cdot \rangle)$ ein Euklidischer Vektorraum und $f : V \rightarrow V$ ein Endomorphismus. Ist f normal (bzw. selbstadjungiert, schiefadjungiert oder eine Isometrie), so gilt dies auch für $f_{\mathbb{C}}$.

Beweis. Sei $f \in \text{End}(V)$ normal, d.h. $f^* \circ f = f \circ f^*$. Dann gilt wegen der Beziehung (*) auf der Seite 257:

$$(f_{\mathbb{C}})^* \circ f_{\mathbb{C}} = (f^*)_{\mathbb{C}} \circ f_{\mathbb{C}} = (f^* \circ f)_{\mathbb{C}} = (f \circ f^*)_{\mathbb{C}} = f_{\mathbb{C}} \circ (f^*)_{\mathbb{C}} = f_{\mathbb{C}} \circ (f_{\mathbb{C}})^*.$$

Ist f selbstadjungiert, d.h. $f^* = f$, so auch $(f_{\mathbb{C}})^* = f_{\mathbb{C}}$. Die restlichen Behauptungen folgen genauso. \square

Lemma 13.5.10. Sei $(V, \langle \cdot, \cdot \rangle)$ ein Euklidischer Vektorraum. Es sei $f : V \rightarrow V$ ein normaler Endomorphismus und $f_{\mathbb{C}} : V_{\mathbb{C}} \rightarrow V_{\mathbb{C}}$ die Komplexifizierung von f . Sei $v = v_1 + iv_2$ ein Eigenvektor von $f_{\mathbb{C}}$ zum Eigenwert λ . Dann ist $\bar{v} = v_1 - iv_2$ ebenfalls Eigenvektor von $f_{\mathbb{C}}$ zum Eigenwert $\bar{\lambda}$ und es gilt $\|v\|_{\mathbb{C}} = \|\bar{v}\|_{\mathbb{C}}$. Sind die zugehörigen Eigenwerte $\lambda, \bar{\lambda}$ verschieden, d.h. $\lambda \notin \mathbb{R}$, so folgt $\langle v, \bar{v} \rangle_{\mathbb{C}} = 0$. Ist zusätzlich $\|v\|_{\mathbb{C}} = \|\bar{v}\|_{\mathbb{C}} = \sqrt{2}$, so bilden die reellen Vektoren v_1, v_2 ein ON-System in V . Ist $\lambda = a + ib$, so erhalten wir:

$$f(v_1) = av_1 - bv_2 \text{ sowie } f(v_2) = bv_1 + av_2, \text{ und } f^*(v_1) = av_1 + bv_2 \text{ sowie } f^*(v_2) = -bv_1 + av_2.$$

Insbesondere ist $U = \text{Span}\{v_1, v_2\}$ ein f -invarianter und f^* -invarianter Unterraum von V und es gilt:

$$M_{(v_2, v_1)}(f|_U) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \quad \text{und} \quad M_{(v_2, v_1)}(f^*|_U) = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

Beweis. Ist $v = v_1 + iv_2$ ein Eigenvektor von $f_{\mathbb{C}}$ zum Eigenwert λ , so ist, wie wir auf Seite 258 gesehen haben, $\bar{v} = v_1 - iv_2$ ein Eigenvektor von $f_{\mathbb{C}}$ zum Eigenwert $\bar{\lambda}$. Außerdem gilt:

$$\langle v, v \rangle_{\mathbb{C}} = \langle v_1, v_1 \rangle + \langle v_2, v_2 \rangle = \langle v_1, v_1 \rangle + \langle -v_2, -v_2 \rangle = \langle \bar{v}, \bar{v} \rangle_{\mathbb{C}},$$

und somit folgt $\|v\|_{\mathbb{C}} = \|\bar{v}\|_{\mathbb{C}}$. Wegen Korollar 13.5.4 folgt $f_{\mathbb{C}}^*(v) = \bar{\lambda}v$. Daher gilt:

$$\bar{\lambda} \langle v, \bar{v} \rangle_{\mathbb{C}} = \langle f_{\mathbb{C}}^*(v), \bar{v} \rangle = \langle v, f_{\mathbb{C}}(\bar{v}) \rangle = \langle v, \bar{\lambda} \bar{v} \rangle = \lambda \langle v, \bar{v} \rangle_{\mathbb{C}}.$$

Ist $\lambda \neq \bar{\lambda}$, so folgt somit $\langle v, \bar{v} \rangle_{\mathbb{C}} = 0$. Damit erhalten wir

$$0 = \langle v, \bar{v} \rangle_{\mathbb{C}} = \langle v_1, v_1 \rangle - \langle v_2, v_2 \rangle + i(2\langle v_1, v_2 \rangle).$$

Dies impliziert

$$\langle v_1, v_2 \rangle = 0 \text{ sowie } \langle v_1, v_1 \rangle = \langle v_2, v_2 \rangle.$$

Ist $\|v\|_{\mathbb{C}} = \sqrt{2}$, so ist

$$\langle v_1, v_1 \rangle + \langle v_2, v_2 \rangle = 2$$

und wir erhalten

$$\langle v_1, v_1 \rangle = \langle v_2, v_2 \rangle = 1.$$

Ist nun $\lambda = a + ib$, so folgt

$$f(v_1) + if(v_2) = f_{\mathbb{C}}(v) = \lambda v = (a + ib)(v_1 + iv_2) = av_1 - bv_2 + i(bv_1 + av_2)$$

sowie

$$f^*(v_1) + if^*(v_2) = f_{\mathbb{C}}^*(v) = \bar{\lambda}v = (a - ib)(v_1 + iv_2) = av_1 + bv_2 + i(-bv_1 + av_2),$$

woraus die restlichen Behauptungen des Lemmas folgen. \square

Bemerkung. Die 2×2 -Matrix $J(a, b)$ beschreibt eine Drehung gefolgt von einer Streckung, denn

$$0 \neq \begin{pmatrix} a & -b \\ b & a \end{pmatrix} = \sqrt{a^2 + b^2} \begin{pmatrix} \frac{a}{\sqrt{a^2 + b^2}} & \frac{-b}{\sqrt{a^2 + b^2}} \\ \frac{b}{\sqrt{a^2 + b^2}} & \frac{a}{\sqrt{a^2 + b^2}} \end{pmatrix}.$$

Da der erste Spaltenvektor die Länge 1 hat, gibt es genau ein $\varphi \in [0, 2\pi)$ mit

$$(\cos \varphi, \sin \varphi) = \left(\frac{a}{\sqrt{a^2 + b^2}}, \frac{b}{\sqrt{a^2 + b^2}} \right)$$

und daher ist

$$J(a, b) = \|(a, b)\| \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}.$$

Dabei beschreibt $D(\varphi) = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}$ eine Drehung um φ im Uhrzeigersinn.

Korollar 13.5.12. *Es sei f ein normaler Endomorphismus eines Euklidischen Vektorraumes V . Dann ist V direkte Summe von ein- und zweidimensionalen invarianten Unterräumen. Auf den zweidimensionalen Unterräumen bewirkt f eine "Drehstreckung".*

Bevor wir die Spezialfälle normaler Endomorphismen genauer untersuchen, wollen wir noch eine Matrixversion der Sätze 13.5.6 und 13.5.11 angeben.

Sei $(K^n, \langle \cdot, \cdot \rangle_{K^n})$ der Euklidische oder unitäre Standardvektorraum und sei $A \in \text{End}(K^n)$ identifiziert mit $A \in M(n, K)$. Dann ist wegen der Bemerkung nach Definition 13.5.2 $A^* = \overline{A}^t$ und damit übersetzt sich die Begriffsbildung in 13.5.2 auf Matrizen wie folgt:

Sei $A \in M(n, K)$, so ist

1. A normal $\Leftrightarrow A\overline{A}^t = \overline{A}^t A$,
2. A selbstadjungiert:
 - (a) A symmetrisch $\Leftrightarrow K = \mathbb{R}$ und $A^t = A$,
 - (b) A hermitesch $\Leftrightarrow K = \mathbb{C}$ und $\overline{A}^t = A$,
3. A schiefadjungiert:
 - (a) A schiefsymmetrisch $\Leftrightarrow K = \mathbb{R}$ und $A^t = -A$,
 - (b) A schiefhermitesch $\Leftrightarrow K = \mathbb{C}$ und $\overline{A}^t = -A$,
4. A Isometrie:
 - (a) A orthogonal $\Leftrightarrow K = \mathbb{R}$ und $A^t = A^{-1}$,
 - (b) A unitär $\Leftrightarrow K = \mathbb{C}$ und $\overline{A}^{-t} = A^{-1}$.

Eine wichtige Klasse normaler Abbildungen bilden die orthogonalen und unitären Endomorphismen, d.h. die Klasse der Isometrien. Diese beschreiben eine Gruppe bezüglich der Verknüpfung von Abbildung.

$$\text{wobei } J(a, b) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \in M(2, \mathbb{R}) \text{ und } \lambda_j \in \mathbb{R}.$$

Bemerkung. Die komplexen normalen Matrizen sind also genau diejenigen, die sich mit Matrizen aus $U(n)$ zu einer Diagonalmatrix konjugieren lassen.

Beweis. Sei $K = \mathbb{C}$. Dann ist wegen Satz 13.5.6 die Matrix $A \in M(n, \mathbb{C})$ genau dann normal, falls eine ON-Basis $B = (b_1, \dots, b_n)$ von \mathbb{C}^n aus Eigenvektoren von A existiert. Dies ist äquivalent zu

$$M_B(A) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = \phi_B \circ A \circ \phi_B^{-1} = CAC^{-1} = C\overline{A}^t$$

mit $C \in U(n)$. Den Fall $K = \mathbb{R}$ behandelt man analog. \square

13.6 Charakterisierung der selbstadjungierten, schiefadjungierten sowie der unitären und orthogonalen Endomorphismen

In diesem Abschnitt untersuchen wir die o.g. speziellen normalen Endomorphismen.

Satz 13.6.1. *Es sei $(V, \langle \cdot, \cdot \rangle)$ ein unitärer Vektorraum und $f \in \text{End}(V)$. Dann gilt:*

- i) *f ist hermitesch \Leftrightarrow es gibt eine ON-Basis aus Eigenvektoren und die Eigenwerte sind reell.*
- ii) *f ist schiefhermitesch \Leftrightarrow es gibt eine ON-Basis aus Eigenvektoren und die Eigenwerte sind rein imaginär.*
- iii) *f ist unitär \Leftrightarrow es gibt eine ON-Basis aus Eigenvektoren und die Eigenwerte haben Betrag 1.*

Beweis. Ist f normal, so ist wegen Korollar 13.5.4 jeder Eigenvektor von f zum Eigenwert λ auch Eigenvektor von f^* zum Eigenwert $\bar{\lambda}$. Damit gilt:

- i) Ist $f = f^*$, so existiert wegen Satz 13.5.6 eine ON-Basis B aus Eigenvektoren v_1, \dots, v_n und $f(v_j) = \lambda_j v_j = f^*(v_j) = \bar{\lambda}_j v_j$, d.h. $\lambda_j = \bar{\lambda}_j$. Ist umgekehrt B eine ON-Basis aus Eigenvektoren von f mit reellen Eigenwerten, so folgt mit Satz 13.4.7 $M_B(f^*) = \overline{M_B(f)}^t = M_B(f)$.
- ii) Ist $f^* = -f$, so existiert eine ON-Basis aus Eigenvektoren mit $\bar{\lambda}_j = -\lambda_j$, d.h. λ_j ist rein imaginär. Ist B umgekehrt eine ON-Basis aus Eigenvektoren von f zu rein imaginären Eigenwerten, so gilt:

$$M_B(f^*) = \overline{M_B(f)}^t = -M_B(f).$$

- iii) Ist $f^* = f^{-1}$, so existiert eine ON-Basis v_1, \dots, v_n aus Eigenvektoren von f mit Eigenwerten $\lambda_1, \dots, \lambda_n$. Da $\bar{\lambda}_j v_j = f^*(v_j) = f^{-1}(v_j) = \frac{1}{\lambda_j} v_j$, ist $\bar{\lambda}_j \cdot \lambda_j = 1$, d.h. $|\lambda_j| = 1$.

Ist umgekehrt B eine ON-Basis aus Eigenvektoren von f mit Eigenwerten $\lambda_1, \dots, \lambda_n$, so dass $\lambda_j \cdot \bar{\lambda}_j = 1$, so folgt:

$$M_B(f^*) = \overline{M_B(f)}^t = \begin{pmatrix} \bar{\lambda}_1 & & 0 \\ & \ddots & \\ 0 & & \bar{\lambda}_n \end{pmatrix}.$$

Dann gilt: $M_B(f)\overline{M_B(f)}^t = E_n$, d.h. $f \circ f^* = \text{id}$. \square

Dieser Satz hat die folgende Matrixversion.

Korollar 13.6.2. Sei $(\mathbb{C}^n, \langle \cdot, \cdot \rangle_{\mathbb{C}^n})$ der unitäre Vektorraum mit dem Standardskalarprodukt und $A \in M(n, \mathbb{C})$. Dann gilt:

- i) A ist hermitesch ($\overline{A}^t = A$) \Leftrightarrow es existiert eine unitäre Matrix $C \in U(n)$, so dass $CAC^{-1} = C\overline{A}^t$ eine reelle Diagonalmatrix ist.
- ii) A ist schieferhermitesch ($\overline{A}^t = -A$) \Leftrightarrow es existiert eine unitäre Matrix $C \in U(n)$, so dass $CAC^{-1} = C\overline{A}^t$ eine Diagonalmatrix ist, wobei die Diagonalelemente rein imaginär sind.
- iii) A ist unitär ($\overline{A}^t A = \text{id}$) \Leftrightarrow es existiert eine unitäre Matrix $C \in U(n)$, so dass $CAC^{-1} = C\overline{A}^t$ eine Diagonalmatrix ist, wobei die Diagonalelemente vom Betrag 1 sind.

Beweis. Die Existenz einer ON-Basis B aus Eigenvektoren von $A \in M(n, \mathbb{C})$ ist äquivalent dazu, dass es ein $C \in U(n)$ gibt, so dass CAC^{-1} eine Diagonalmatrix mit den Eigenwerten von A in der Diagonalen ist. Damit folgt das Korollar aus obigem Satz. \square

Nun wollen wir noch den zu Satz 13.6.1 analogen Satz für Euklidische Vektorräume angeben.

Satz 13.6.3. Es sei $(V, \langle \cdot, \cdot \rangle)$ ein Euklidischer Vektorraum und $f \in \text{End}(V)$. Dann gilt:

- i) f ist symmetrisch \Leftrightarrow es gibt eine ON-Basis B aus Eigenvektoren, d.h.

$$M_B(f) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \in M(n, \mathbb{R}).$$

- ii) f ist schiefsymmetrisch \Leftrightarrow es gibt eine ON-Basis B mit

$$M_B(f) = \begin{pmatrix} 0 & & & & & & & & & 0 \\ & \ddots & & & & & & & & \\ & & 0 & & & & & & & \\ & & & 0 & -b_1 & & & & & \\ & & & b_1 & 0 & & & & & \\ & & & & & \ddots & & & & \\ & & & & & & 0 & -b_\ell & & \\ 0 & & & & & & b_\ell & 0 & & \end{pmatrix},$$

Ist f zusätzlich schiefssymmetrisch, so folgt: $M_B(f) = -M_B(f)^t$. Dies ist genau dann erfüllt, falls für alle $i \in \{1, \dots, k\}$ und $j \in \{1, \dots, \ell\}$ gilt:

$$\lambda_i = -\lambda_i \quad \text{und} \quad \begin{pmatrix} a_j & -b_j \\ b_j & a_j \end{pmatrix} = \begin{pmatrix} -a_j & -b_j \\ b_j & -a_j \end{pmatrix}.$$

Also ist $\lambda_i = 0$ und $a_j = 0$. Hat $M_B(f)$ die in (ii) beschriebene Form, so folgt $M_B(f) = -M_B(f)^t$ und somit ist f schiefssymmetrisch.

Ist f zusätzlich orthogonal, so folgt: $M_B(f) \cdot M_B(f)^t = E_n$. Dies ist genau dann erfüllt, falls für alle $i \in \{1, \dots, k\}$ und $j \in \{1, \dots, \ell\}$ gilt:

$$\lambda_i^2 = 1 \quad \text{und} \quad \begin{pmatrix} a_j & -b_j \\ b_j & a_j \end{pmatrix} \begin{pmatrix} a_j & b_j \\ -b_j & a_j \end{pmatrix} = \begin{pmatrix} a_j^2 + b_j^2 & 0 \\ 0 & a_j^2 + b_j^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

d.h. $\lambda_i = \pm 1$ und $a_j^2 + b_j^2 = 1$. Da $b_j \neq 0$, existiert $\varphi_j \in (0, 2\pi) \setminus \{\pi\}$ mit $a_j = \cos \varphi_j, b_j = \sin \varphi_j$. Hat $M_B(f)$ die in (iii) beschriebene Form, so folgt: $M_B(f) \cdot M_B(f)^t = \text{id}$ und somit ist f orthogonal. \square

Diesem Satz entspricht wieder die folgende Matrixversion.

Korollar 13.6.4. Sei $(\mathbb{R}^n, \langle \cdot, \cdot \rangle_{\mathbb{R}^n})$ der Euklidische Vektorraum \mathbb{R}^n mit Standardskalarprodukt und $A \in M(n, \mathbb{R})$. Dann gilt:

i) A ist symmetrisch \Leftrightarrow es gibt eine orthogonale Matrix $C \in O(n)$ mit

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = CAC^{-1} = CAC^t.$$

ii) A ist schiefssymmetrisch \Leftrightarrow es gibt eine orthogonale Matrix $C \in O(n)$ mit

$$\begin{pmatrix} 0 & & & & & & & & & 0 \\ & \ddots & & & & & & & & \\ & & 0 & & & & & & & \\ & & & 0 & b_1 & & & & & \\ & & & -b_1 & 0 & & & & & \\ & & & & & \ddots & & & & \\ & & & & & & 0 & b_\ell & & \\ 0 & & & & & & -b_\ell & 0 & & \end{pmatrix} = CAC^{-1} = CAC^t.$$

Der Eigenvektor b_1 zum Eigenwert 1 erzeugt einen eindimensionalen Unterraum, der punktweise festgehalten wird. Ist $f \neq \text{id}$, so ist dieser Unterraum eindeutig bestimmt und heißt *Drehachse*. Die zu b_1 orthogonale, invariante Ebene, die durch (b_2, b_3) aufgespannt wird, heißt *Drehebene*. Für den *Drehwinkel* φ gilt:

$$\cos \varphi = \frac{1}{2}(\text{Spur } f - 1)$$

Wir erinnern nochmals daran (siehe Abschnitt 12.3 und Übungsblatt 11, Mathematik für Physiker II), dass die Spur eines Endomorphismus f durch $\text{Spur } f = \text{Spur } M_C(f)$ bezüglich einer beliebigen Basis C definiert. Dabei ist $\text{Spur } M_C(f)$, also die Summe der Diagonalelemente der Matrix $M_C(f)$ unabhängig von der Wahl der Basis.

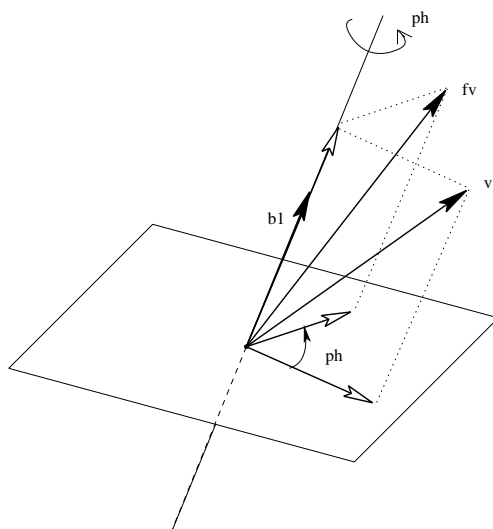


Abbildung 13.3:

13.7 Symmetrische und Hermitesche Formen

Mit diesem Abschnitt werden wir das Studium der linearen Algebra abschliessen. Er ist der Untersuchung symmetrischer und hermitescher Formen gewidmet. Diese stellen eine Verallgemeinerung der Skalarprodukte dar. Wie in allen Abschnitten dieses Kapitels betrachten wir wieder ausschließlich Vektorräume über $K = \mathbb{R}$ oder \mathbb{C} .

Definition 13.7.1. Sei V ein Vektorraum über K . Unter einer *hermiteschen Form* auf V verstehen wir eine Abbildung $s : V \times V \rightarrow K$ mit folgenden Eigenschaften:

- (i) s ist linear im ersten Argument, d.h. für alle $x, x', y \in V$ und $\alpha, \beta \in K$

$$s(\alpha x + \beta x', y) = \alpha s(x, y) + \beta s(x', y)$$

- (ii) $s(x, y) = \overline{s(y, x)}$.

Im Fall $K = \mathbb{R}$ sprechen wir auch von *symmetrischen Formen*.

Die Form s heißt *positiv definit* (*positiv semidefinit*), falls $s(x, x) > 0$ ($s(x, x) \geq 0$) für alle $x \in V, x \neq 0$.

Bemerkungen.

- (a) Aus der Definition einer hermiteschen Form s folgt, dass s semilinear (konjugiert linear) im zweiten Argument ist, d.h. es gilt

$$s(x, \alpha y + \beta y') = \bar{\alpha}s(x, y) + \bar{\beta}s(x, y')$$

für alle $x, y, y' \in V$ und $\alpha, \beta \in K$. Ist $K = \mathbb{R}$, so ist s natürlich auch im zweiten Argument linear. Dann nennt man s auch *symmetrische Bilinearform*.

- (b) Eigenschaft (ii) impliziert, dass $s(x, x)$ eine reelle Zahl ist. Ist s positiv definit, so ist s ein Skalarprodukt (siehe Definition 13.1.1).
- (c) Verzichtet man in obiger Definition auf die Symmetriebedingung (ii) und ersetzt sie durch die Forderung der Semilinearität im zweiten Argument, so spricht man auch von *Sesquilinearformen*.

In endlichdimensionalen Vektorräumen mit Skalarprodukt entsprechen den hermiteschen Formen die selbstadjungierten Endomorphismen. Es gilt:

Satz 13.7.2. *Sei $(V, \langle \cdot, \cdot \rangle)$ ein endlichdimensionaler Vektorraum mit Skalarprodukt. Dann ist $s : V \times V \rightarrow K$ genau dann hermitesch, wenn ein selbstadjungierter Endomorphismus $f : V \rightarrow V$ existiert mit*

$$s(x, y) = \langle x, f(y) \rangle.$$

Beweis. Sei $f : V \rightarrow V$ ein selbstadjungierter Endomorphismus, so ist $s(x, y) := \langle x, f(y) \rangle$ eine hermitesche Form, denn

$$s(x, y) = \langle x, f(y) \rangle = \langle f(x), y \rangle = \overline{\langle y, f(x) \rangle} = \overline{s(y, x)}.$$

Die Umkehrung folgt aus dem Rieszschen Darstellungssatz 13.3.10, der besagt: Ist V ein endlichdimensionaler Vektorraum mit Skalarprodukt, so ist $\sigma : V \rightarrow V^*$ mit $\sigma(y) = \langle \cdot, y \rangle$ eine invertierbare semilineare Abbildung. Ist nun s eine hermitesche Form, so ist $\tilde{s} : V \rightarrow V^*$ mit

$$y \rightarrow s(\cdot, y)$$

ebenfalls eine semilineare Abbildung. Man definiere $f : V \rightarrow V$ durch

$$f(y) := \sigma^{-1}(s(\cdot, y)) = \sigma^{-1}(\tilde{s}(y))$$

Dies ist äquivalent zu

$$\sigma \circ f(y) = s(\cdot, y) = \langle \cdot, f(y) \rangle \Leftrightarrow \langle x, f(y) \rangle = s(x, y)$$

für alle $x, y \in V$. Dann ist f eine lineare Abbildung (denn die Verknüpfung zweier semilinearer Abbildungen ist linear) und es gilt:

$$\langle x, f(y) \rangle = s(x, y) = \overline{s(y, x)} = \overline{\langle y, f(x) \rangle} = \langle f(x), y \rangle,$$

d.h. f ist selbstadjungiert. □

Sei s eine hermitesche Form und $B = (b_1, \dots, b_n) \subset V$ eine Basis. Dann können wir s wie folgt eine Matrix zuordnen.

Definition 13.7.3. Die Matrix $G_B(s) = (g_{ij})$ mit $g_{ij} = s(b_i, b_j)$ heißt *Fundamentalmatrix* von s bezüglich B .

Lemma 13.7.4. Sei s eine hermitesche Form auf V und $G_B(s) = (g_{ij})$ die Fundamentalmatrix von s bezüglich einer Basis B . Dann gilt:

1.

$$s(x, y) = \langle \phi_B(x), G_B(s) \circ \phi_B(y) \rangle_{K^n} = \sum_{i,j=1}^n g_{ij} x_i \bar{y}_j$$

wobei $\phi_B : V \rightarrow K^n$ die Koordinatenabbildung bezüglich B mit $\phi_B(x) = (x_1, \dots, x_n)$ und $\langle \cdot, \cdot \rangle_{K^n}$ das Standardskalarprodukt auf K^n bezeichnet.

2. Die Fundamentalmatrix $G_B(s)$ ist hermitesch.

3. Ist B' eine weitere Basis von V , so gilt:

$$G_{B'}(s) = \bar{T}^t G_B(s) T,$$

wobei $T = \phi_B \circ \phi_{B'}^{-1} : K^n \rightarrow K^n$ die Koordinatentransformationsmatrix bezeichnet.

Beweis.

1. Ist s eine hermitesche Form, so folgt:

$$\begin{aligned} s(x, y) &= s\left(\sum_i x_i b_i, \sum_j y_j b_j\right) = \sum_{i,j} x_i \bar{y}_j s(b_i, b_j) \\ &= \sum_j \left(\sum_i g_{ij} x_i\right) \bar{y}_j = \sum_i x_i \left(\sum_j g_{ij} \bar{y}_j\right) = \langle \phi_B(x), G_B(s) \circ \phi_B(y) \rangle_{K^n}. \end{aligned}$$

2. Da s hermitesch ist, folgt: $g_{ij} = s(b_j, b_i) = \overline{s(b_i, b_j)} = \overline{g_{ji}}$.

3. Ist B' eine weitere Basis, so gilt:

$$\begin{aligned} s(x, y) &= \langle \phi_B(x), G_B(s) \circ \phi_B(y) \rangle_{K^n} = \langle \phi_B \circ \phi_{B'}^{-1} \circ \phi_{B'}(x), G_B(s) \circ \phi_B \circ \phi_{B'}^{-1} \circ \phi_{B'}(y) \rangle_{K^n} \\ &= \langle T \circ \phi_{B'}(x), G_B(s) \circ T \circ \phi_{B'}(y) \rangle_{K^n} = \langle \phi_{B'}(x), \bar{T}^t \circ G_B(s) \circ T \circ \phi_{B'}(y) \rangle_{K^n} \\ &= \langle \phi_{B'}(x), G_{B'}(s) \circ \phi_{B'}(y) \rangle_{K^n}. \end{aligned}$$

□

Sei $(V, \langle \cdot, \cdot \rangle)$ ein Vektorraum mit Skalarprodukt und s eine symmetrische bzw. hermitesche Form. Von wichtiger geometrischer Bedeutung sind die *Hauptachsen* von s bezüglich $\langle \cdot, \cdot \rangle$.

Definition 13.7.5. Ist $(V, \langle \cdot, \cdot \rangle)$ ein Vektorraum mit Skalarprodukt und s eine hermitesche Form. Eine ON-Basis $B = (b_1, \dots, b_n)$ von V heißt *Hauptachsensystem* von s bezüglich $\langle \cdot, \cdot \rangle$, falls die Fundamentalmatrix $G_B(s)$ reelle Diagonalgestalt besitzt, d.h. $g_{ij} = \lambda_j \delta_{ij} = s(b_i, b_j)$ mit $\lambda_j \in \mathbb{R}$.

Satz 13.7.6. Sei V ein Vektorraum endlicher Dimension und s eine hermitesche Form. Sei \langle, \rangle ein Skalarprodukt auf V , so besitzt s ein Hauptachsensystem bezüglich \langle, \rangle .

Bemerkung. Ein Hauptachsensystem ist im Allgemeinen nicht eindeutig bestimmt. Ist z.B. $s = \langle, \rangle$, so ist jede ON-Basis Hauptachsensystem von s bezüglich \langle, \rangle .

Die Konstruktion eines Hauptachsensystems bezeichnet man auch als *Hauptachsentransformation*.

Beweis. Sei s eine hermitesche Form und \langle, \rangle ein Skalarprodukt. Dann gibt es wegen Satz 13.7.2 eine selbstadjungierte Abbildung $f : V \rightarrow V$ mit $s(x, y) = \langle x, f(y) \rangle = \langle f(x), y \rangle$. Wegen 13.6.1 i) besitzt V eine ON-Basis $B = (b_1, \dots, b_n)$ aus Eigenvektoren von f mit reellen Eigenwerten. Insbesondere gilt:

$$s(b_i, b_j) = \langle b_i, f(b_j) \rangle = \langle b_i, \lambda_j b_j \rangle = \lambda_j \langle b_i, b_j \rangle = \lambda_j \delta_{ij},$$

d.h. B ist ein Hauptachsensystem von s bezüglich \langle, \rangle . □

Diese Hauptachsen lassen sich wie folgt geometrisch kennzeichnen.

Satz 13.7.7. Sei (V, \langle, \rangle) ein Vektorraum mit Skalarprodukt und s eine hermitesche Form. Sei $B = (b_1, \dots, b_n)$ ein Hauptachsensystem von s bezüglich \langle, \rangle mit

$$\lambda_1 := s(b_1, b_1) \geq \dots \geq \lambda_n := s(b_n, b_n).$$

Dann gilt:

$$\begin{aligned} \lambda_1 &= \max\{s(x, x) \mid \langle x, x \rangle = 1\} = s(b_1, b_1). \\ \lambda_2 &= \max\{s(x, x) \mid \langle x, x \rangle = 1, x \perp b_1\} = s(b_2, b_2). \\ \lambda_3 &= \max\{s(x, x) \mid \langle x, x \rangle = 1, x \perp b_1, b_2\} = s(b_3, b_3). \\ &\vdots \\ \lambda_n &= \max\{s(x, x) \mid \langle x, x \rangle = 1, x \perp b_1, \dots, b_{n-1}\} = s(b_n, b_n). \end{aligned}$$

Beweis. Es sei $B = (b_1, \dots, b_n)$ Hauptachsensystem von s bezüglich \langle, \rangle und $x = \sum_{j=1}^n x_j b_j$.

Dann gilt:

$$s(x, x) = s\left(\sum_{i=1}^n x_i b_i, \sum_{j=1}^n x_j b_j\right) = \sum_{i,j=1}^n x_i \bar{y}_j s(b_i, b_j) = \sum_{j=1}^n x_j \bar{x}_j s(b_j, b_j) = \sum_{j=1}^n x_j \bar{x}_j \lambda_j.$$

Ist $\langle x, x \rangle = \sum_{j=1}^n x_j \bar{x}_j = 1$, so folgt:

$$s(x, x) = \lambda_1 + \sum_{j=1}^n (\lambda_j - \lambda_1) x_j \bar{x}_j \leq \lambda_1 = s(b_1, b_1).$$

Ist $x \perp b_1$, d.h. $x = \sum_{j=2}^n x_j b_j$, so ist

$$s(x, x) = \lambda_2 + \sum_{j=2}^n (\lambda_j - \lambda_2) x_j \bar{x}_j \leq \lambda_2 = s(b_2, b_2)$$

Der Beweis ergibt sich aus der Fortsetzung des Prozesses. □

Beispiele. Man betrachte $V = \mathbb{R}^2$ mit Standardskalarprodukt $\langle \cdot, \cdot \rangle_{\mathbb{R}^2}$.

- a) Ist $s(x, x) = 4x_1^2 + x_2^2$, so gilt: $\lambda_1 = 4, \lambda_2 = 1$ und die Hauptachsen sind $(1, 0)$ und $(0, 1)$.

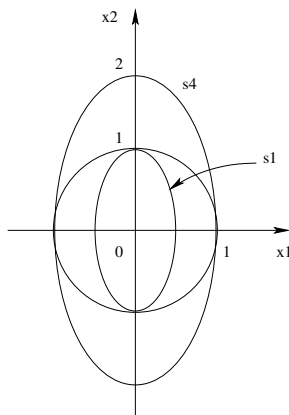


Abbildung 13.4:

Auf $S^1 := \{(x_1, x_2) \mid x_1^2 + x_2^2 = 1\}$ nimmt also die Form s ihr Maximum an der Stelle $(1, 0)$ und ihr Minimum an der Stelle $(0, 1)$ an.

- b) Man betrachte $s(x, x) = 4x_1^2 - x_2^2$, so ist $\lambda_1 = 4, \lambda_2 = -1$ und die Hauptachsen sind wieder $(1, 0)$ und $(0, 1)$

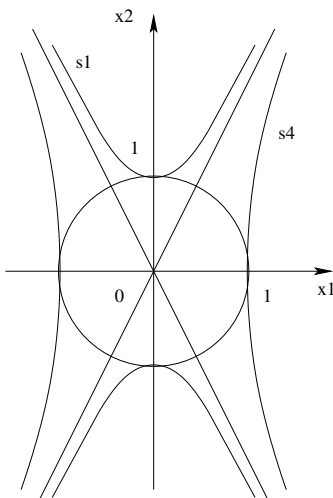


Abbildung 13.5:

Auf S^1 nimmt s ebenfalls ihr Maximum an der Stelle $(1, 0)$ und ihr Minimum an der Stelle $(0, 1)$ an.

- c) Allgemein hat jede symmetrische Form auf $(\mathbb{R}^2, \langle \cdot, \cdot \rangle)_{\mathbb{R}^2}$ die Gestalt:

$$s(x, y) = \left\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right\rangle = ax_1y_1 + bx_2y_1 + bx_1y_2 + cx_2y_2.$$

Insbesondere ist $s(x, x) = ax_1^2 + 2bx_1x_2 + cx_2^2$. Um die Hauptachse zu berechnen, muss man also $\begin{pmatrix} a & b \\ b & c \end{pmatrix}$ diagonalisieren. Die Vorzeichen der Eigenwerte bestimmen die geometrische Bedeutung von s . (Sind sie ungleich Null, so gilt: Haben sie gleiches Vorzeichen, so sind die Niveaulinien von s Ellipsen, sonst Hyperbeln).
Aufgabe: Welche Möglichkeiten ergeben sich in \mathbb{R}^3 ?

Definition 13.7.8. Sei $s : V \times V \rightarrow K$ eine hermitesche Form.

$$V^0 := \{x \in V \mid s(x, y) = 0 \text{ für alle } y \in V\}$$

heißt der *Nullraum* von s . Eine Form heißt *nicht ausgeartet*, falls $V^0 = \{0\}$.

Bemerkungen.

- Jedes Skalarprodukt s ist nicht ausgeartet, denn $s(x, x) > 0$, falls $x \neq 0$.
- s ist genau dann nicht ausgeartet, falls $\tilde{s} : V \rightarrow V^*$ mit $x \rightarrow s(\cdot, x)$ injektiv ist, denn

$$s(\cdot, x) = 0 \Leftrightarrow 0 = s(y, x) = s(x, y) \text{ für alle } y \in V \Leftrightarrow x \in V^0.$$

- Ist $\langle \cdot, \cdot \rangle$ ein Skalarprodukt, so existiert wegen Satz 13.7.2 eine selbstadjungierte Abbildung $f : V \rightarrow V$ mit $s(x, y) = \langle x, f(y) \rangle = \langle f(x), y \rangle$. Dann gilt: $V^0 = \text{Kern}(f)$.

Satz 13.7.9. Sei s eine hermitesche bzw. symmetrische Form auf V und sei V' ein lineares Komplement von V^0 . Dann ist $s|_{V'}$ nicht ausgeartet.

Beweis. Sei $x \in V'$ und $s(x, y') = 0$ für alle $y' \in V'$. Dann gilt für alle $y = y_0 + y' \in V^0 \oplus V' = V$:

$$s(x, y) = s(x, y_0 + y') = s(x, y_0) + s(x, y') = 0,$$

d.h. $x \in V^0 \cap V' = \{0\}$. Dies zeigt: $s|_{V'}$ ist nicht ausgeartet. \square

Definition 13.7.10. Sei s eine hermitesche Form auf V . Wir nennen einen Unterraum $W \subset V$ *positiv definit* (*negativ definit*) bezüglich s , falls gilt: $s(x, x) > 0$ ($s(x, x) < 0$) für alle $x \in W$ mit $x \neq 0$.

Die Zahlen (n_+, n_-, n_0) mit

$$\begin{aligned} n_+ &= \max\{\dim W \mid W \subset V \text{ positiv definit}\}. \\ n_- &= \max\{\dim W \mid W \subset V \text{ negativ definit}\}. \\ n_0 &= \dim V^0. \end{aligned}$$

heißen die *Signatur* von s . Ist s nicht ausgeartet, d.h. $n_0 = 0$, so schreiben wir auch (n_+, n_-) .

Satz 13.7.11. (Trägheitssatz von Sylvester)

Sei s eine hermitesche Form auf V . Dann besitzt V eine direkte Summenzerlegung

$$V = V^+ \oplus V^- \oplus V^0,$$

wobei V^+ positiv definit und V^- negativ definit ist. Für jede Zerlegung dieser Art gilt:

$$\dim V^+ = n_+, \quad \dim V^- = n_-.$$

Beweis. Wir zeigen zunächst die Existenz der Zerlegung.

Dazu wählt man auf V ein Skalarprodukt $\langle \cdot, \cdot \rangle$. Dann existiert wegen Satz 13.7.6 ein Hauptachsenssystem, d.h. eine ON-Basis $B = (b_1, \dots, b_n)$ mit $s(b_i, b_j) = \lambda_i \delta_{ij}$. Dann gilt:

$$V^0 = \text{Span}\{b_i \in B \mid s(b_i, b_i) = \lambda_i = 0\},$$

denn $x = \sum_{j=1}^n \alpha_j b_j \in V^0$ genau dann, falls für jedes $i \in \{1, \dots, n\}$

$$0 = s(x, b_i) = \sum_{j=1}^n \alpha_j s(b_j, b_i) = \alpha_i s(b_i, b_i) = \alpha_i \lambda_i$$

gilt. Dies ist genau dann der Fall, falls gilt: $\alpha_i = 0$ für alle i mit $\lambda_i \neq 0$, d.h. $x \in \text{Span}\{b_i \in B \mid s(b_i, b_i) = \lambda_i = 0\}$. Man definiere nun

$$\begin{aligned} V^+ &:= \text{Span}\{b_i \in B \mid s(b_i, b_i) = \lambda_i > 0\}, \\ V^- &:= \text{Span}\{b_i \in B \mid s(b_i, b_i) = \lambda_i < 0\}. \end{aligned}$$

Dann ist V^+ positiv definit und V^- negativ definit und da B eine Basis ist, gilt:

$$V = V^+ \oplus V^- \oplus V^0.$$

Sei nun $V = W^+ \oplus W^- \oplus V^0$ eine beliebige direkte Summenzerlegung, so dass W^+ positiv definit und W^- negativ definit ist.

Sei $\text{pr}^+ : V = W^+ \oplus W^- \oplus V^0 \rightarrow W^+$ mit $x = x_+ + x_- + x_0 \rightarrow x_+$ die Projektion auf W^+ .

Sei $W \subset V$ ein beliebiger positiv definiter Unterraum. Wir zeigen:

$\text{Kern}(\text{pr}^+|_W) = 0$. Dann folgt: $\dim W \leq \dim W^+$, d.h. hat man eine Zerlegung der obigen Form, so hat kein positiv definiter Unterraum eine grössere Dimension als W^+ . Insbesondere ist $n_+ = \dim W^+$. Sei $x \in W$ und $\text{pr}^+(x) = 0$. Dann ist $x = x_- + x_0 \in W^- \oplus V^0$ und damit folgt:

$$s(x, x) = s(x_- + x_0, x_- + x_0) = s(x_-, x_-) \leq 0.$$

Da W positiv definit ist, zeigt dies: $x = 0$.

Mit analogen Argumenten zeigt man: $n_- = \dim W^-$. □

Kapitel 14

Metrische Räume

Viele Begriffe aus der Analysis auf \mathbb{R} lassen sich fast wörtlich auf normierte Vektorräume oder sogar metrische Räume ausdehnen.

14.1 Normierte und metrische Räume

Wir erinnern hier nochmals an den Begriff der Norm, den wir in Zusammenhang mit dem Begriff des Skalarproduktes in Kapitel 13 eingeführt haben (siehe Definition 13.2.3).

Definition 14.1.1. Sei V ein Vektorraum über K mit $K = \mathbb{R}$ oder \mathbb{C} . Eine Abbildung $\|\cdot\| : V \rightarrow \mathbb{R}_+ \cup \{0\}$ heißt *Norm* auf V , falls für alle $x \in V$ und $\lambda \in K$ Folgendes erfüllt ist:

- (1) $\|x\| = 0 \Leftrightarrow x = 0$,
- (2) $\|\lambda x\| = |\lambda| \|x\|$,
- (3) $\|x + y\| \leq \|x\| + \|y\|$ (*Dreiecksungleichung*).

Bemerkung. Vektorräume mit Normen heißen *normierte Vektorräume*.

Beispiele.

- (a) $V = \mathbb{R}^n$. Ist $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, so definieren

- (1) $\|x\|_2 = \sqrt{x_1^2 + \dots + x_n^2}$,
- (2) $\|x\|_\infty = \max\{|x_i| \mid i \in \{1, \dots, n\}\}$,
- (3) $\|x\|_p = \sqrt[p]{|x_1|^p + \dots + |x_n|^p}$ für $p \geq 1$

Normen auf \mathbb{R}^n . $\|\cdot\|_2$ heißt *euklidische Norm*, $\|\cdot\|_\infty$ heißt *Maximumsnorm*. Es gilt:
 $\|x\|_\infty = \lim_{p \rightarrow \infty} \|x\|_p$ (Siehe Übungsblatt 4 Mathematik für Physiker III, Aufgabe 4(b)).

- (b) Die euklidische Norm ist Spezialfall einer *durch ein Skalarprodukt induzierten Norm*. Sei $(V, \langle \cdot, \cdot \rangle)$ ein Vektorraum über K mit Skalarprodukt, so definiert $\|x\| = \sqrt{\langle x, x \rangle}$ eine Norm. Die Norm $\|\cdot\|_p$ wird für $p \neq 2$ nicht von einem Skalarprodukt induziert (Übungsblatt 4 Mathematik für Physiker III, Aufgabe 4(c)).

(c) Sei M eine Menge und $B(M, \mathbb{R}) = \{f : M \rightarrow \mathbb{R} \mid f \text{ ist beschränkt}\}$. Dann definiert

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

eine Norm auf $B(M, \mathbb{R})$. Sie heißt *Supremumsnorm*.

Die Eigenschaften (1), (2) und (3) folgen aus den entsprechenden Eigenschaften für den Betrag.

Die Norm eines Vektorraumes V induziert durch

$$d(x, y) := \|x - y\|$$

für $x, y \in V$ eine Metrik auf V . Man interpretiert $d(x, y)$ als Abstand zwischen x und y . Im Gegensatz zur Norm benötigt die Definition einer Metrik keine Vektorraumstruktur.

Definition 14.1.2. Sei M eine Menge und $d : M \times M \rightarrow \mathbb{R}^+ \cup \{0\}$ eine Abbildung. Dann heißt d *Metrik*, falls für alle $x, y \in M$ gilt:

- (1) $d(x, y) = 0 \Leftrightarrow x = y$,
- (2) $d(x, y) = d(y, x)$ (*Symmetrie*),
- (3) $d(x, z) \leq d(x, y) + d(y, z)$ (*Dreiecksungleichung*).

Die Menge M zusammen mit einer Metrik d heißt *metrischer Raum*. Man schreibt (M, d) .

Bemerkungen.

(a) Ist $(V, \|\cdot\|)$ ein normierter Vektorraum, so definiert

$$d(x, y) := \|x - y\|$$

eine Metrik auf V , denn

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

und

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

für alle $x, y, z \in V$. Eigenschaft (1) ist trivialerweise erfüllt.

(b) Ist $A \subset V$ eine beliebige Teilmenge eines normierten Vektorraumes $(V, \|\cdot\|)$, so induziert die Norm eine Metrik $d_A : A \times A \rightarrow \mathbb{R}$ auf A (*induzierte Metrik*) mit

$$d_A(x, y) = \|x - y\|.$$

Die Menge A ist natürlich im Allgemeinen kein Vektorraum. Man betrachte zum Beispiel $A = \{(x_1, x_2) \mid x_1^2 + x_2^2 = 1\} \subset \mathbb{R}^2$.

14.2 Konvergenz und Stetigkeit in metrischen Räumen

Begriffe, wie Konvergenz und Stetigkeit, die wir in den Kapiteln 2 und 3 kennengelernt haben, lassen sich ohne Schwierigkeit auf metrische Räume übertragen.

Definition 14.2.1. Sei (M, d) ein metrischer Raum. Eine Folge $(a_n)_{n \in \mathbb{N}}$ in M ist eine Abbildung der natürlichen Zahlen in die Menge M mit $n \mapsto a_n$. Sie heißt *konvergent* mit Grenzwert $a \in M$, falls die reelle Zahlenfolge $d(a, a_n)$ eine Nullfolge ist. Ist die Folge (a_n) konvergent mit Grenzwert $a \in M$, so schreiben wir

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

Bemerkungen.

- (a) Ist $(V, \|\cdot\|)$ ein normierter Vektorraum, so heißt eine Folge (a_n) in V konvergent, falls sie in dem metrischen Raum (V, d) mit $d(x, y) = \|x - y\|$ konvergent ist.
- (b) Wie bei den reellen Zahlenfolgen gilt, dass der Grenzwert eindeutig bestimmt ist, falls er existiert. Denn ist $\lim_{n \rightarrow \infty} a_n = a$ und $\lim_{n \rightarrow \infty} a_n = b$, so gilt:

$$0 \leq d(a, b) \leq d(a, a_n) + d(a_n, b) \rightarrow 0, \quad n \rightarrow \infty,$$

d.h. $d(a, b) = 0$ und somit ist $a = b$.

- (c) Ist $a \in M$ und $\epsilon > 0$, so nennen wir die Menge

$$B(a, \epsilon) := \{x \in M \mid d(x, a) < \epsilon\}$$

den *durch d induzierten Ball* von Radius ϵ um den Punkt a (kurz: ϵ -Ball oder ϵ -Umgebung um a). Dann ist $\lim_{n \rightarrow \infty} a_n = a$ äquivalent dazu, dass in jedem ϵ -Ball um a fast alle Folgenglieder (alle bis auf endlich viele) enthalten sind.

- (d) Sei $(a_n)_{n \in \mathbb{N}}$ eine Folge in M . Dann heißt $a \in M$ *Häufungspunkt* (HP) der Folge, falls für alle $\epsilon > 0$ die Menge $\{n \in \mathbb{N} \mid a_n \in B(a, \epsilon)\}$ unendlich ist.
- (e) Sei $(a_n)_{n \in \mathbb{N}}$ eine Folge in M und $k_n \in \mathbb{N}$ eine streng monoton steigende Folge, d.h. $k_1 < k_2 < \dots < k_n < \dots$. Dann heißt die Folge $b_n = a_{k_n}$ *Teilfolge* von (a_n) . Außerdem gilt (siehe Satz 2.3.9): $a \in M$ ist HP der Folge $(a_n)_{n \in \mathbb{N}}$ genau dann, falls eine konvergente Teilfolge a_{k_n} existiert mit $\lim_{n \rightarrow \infty} a_{k_n} = a$.

Beispiel. Sei M eine Menge und $(B(M, \mathbb{R}), \|\cdot\|)$ der Vektorraum der beschränkten Funktionen mit der Supremumsnorm. Dann gilt: $f_n \rightarrow f$ genau dann, falls f_n gleichmäßig gegen f konvergiert, d.h. für alle $\epsilon > 0$ existiert ein $n_0 \in \mathbb{N}$ mit

$$|f_n(x) - f(x)| < \epsilon$$

für alle $n \geq n_0$ und $x \in M$.

Auf \mathbb{R}^k ist die Konvergenz einer Folge bezüglich der euklidischen Norm äquivalent zur Konvergenz ihrer Komponentenfolgen.

Lemma 14.2.2. Man betrachte $(\mathbb{R}^k, \|\cdot\|_2)$ mit der euklidischen Norm $\|x\|_2 = \sqrt{x_1^2 + \dots + x_k^2}$. Ist $a_n = (a_{n1}, \dots, a_{nk})$ eine Folge von Vektoren im \mathbb{R}^k , so gilt:

$$\lim_{n \rightarrow \infty} a_n = b = (b_1, \dots, b_k) \Leftrightarrow \lim_{n \rightarrow \infty} a_{nj} = b_j$$

für alle $j \in \{1, \dots, k\}$.

Beweis. Für alle $j \in \{1, \dots, k\}$ gilt:

$$\begin{aligned} |x_j - y_j| &\leq \|x - y\| = \sqrt{(x_1 - y_1)^2 + \dots + (x_k - y_k)^2} \\ &\leq |x_1 - y_1| + \dots + |x_k - y_k|. \end{aligned}$$

Damit folgt: $\|b - a_n\|$ ist Nullfolge $\Leftrightarrow |b_j - a_{nj}|$ ist Nullfolge für alle $j \in \{1, \dots, k\}$. \square

Bemerkung. In einem unendlich dimensionalen Vektorraum hängt die Konvergenz einer Folge im Allgemeinen von der Wahl der Norm ab. Wir werden bald sehen, dass im \mathbb{R}^k die Konvergenz nicht von der Wahl der Norm abhängt. Insbesondere ist die Konvergenz einer Folge bezüglich einer beliebigen Norm äquivalent zur Konvergenz ihrer Komponentenfolgen.

Genauso leicht wie der Begriff der Konvergenz, lässt sich der Begriff der Cauchyfolge auf metrische Räume übertragen.

Definition 14.2.3. Sei (M, d) ein metrischer Raum. Eine Folge $(a_n)_{n \in \mathbb{N}}$ heißt *Cauchyfolge* genau dann, falls gilt:

$$\forall \epsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n, m \geq n_0 : d(a_n, a_m) \leq \epsilon.$$

Bemerkungen.

- (a) Jede konvergente Folge ist eine Cauchyfolge.
- (b) Jede Cauchyfolge (a_n) ist beschränkt, d.h. es existiert ein $p_0 \in M$ und $r > 0$ mit $a_n \in B(p_0, r)$. Man wähle hierzu ein $n_0 \in \mathbb{N}$ mit $d(a_n, a_{n_0}) \leq 1$ für alle $n \geq n_0$. Ist $r \geq 1$ so gewählt, dass $d(a_n, a_{n_0}) \leq r$ für alle $n < n_0$, so folgt $a_n \in B(p_0, r)$ mit $p_0 := a_{n_0}$.

Definition 14.2.4. (a) Ein metrischer Raum (M, d) heißt *vollständig*, falls jede Cauchyfolge einen Grenzwert besitzt.

- (b) Ein normierter Vektorraum $(V, \|\cdot\|)$ heißt *Banachraum*, falls V mit der durch die Norm $\|\cdot\|$ induzierten Metrik vollständig ist.
- (c) Ein Vektorraum V mit Skalarprodukt $\langle \cdot, \cdot \rangle$ heißt *Hilbertraum*, wenn $(V, \|\cdot\|)$ mit der Norm $\|x\| = \sqrt{\langle x, x \rangle}$ ein Banachraum ist.

Bemerkungen.

- (a) Die reellen und komplexen Zahlen sind vollständig (siehe Satz 2.3.7 und Satz 2.4.8). Dies gilt nicht für beliebige metrische Räume. Als Beispiel betrachte man

$$M = \mathbb{R} \setminus \{0\}, \quad \text{mit} \quad d(x, y) = |x - y|.$$

Dann ist $x_n = \frac{1}{n}$ eine Cauchyfolge, die keinen Grenzwert in M besitzt.

(b) Hilberträume spielen in der Quantenmechanik eine zentrale Rolle.

Satz 14.2.5. *Der Vektorraum \mathbb{R}^k ist bezüglich der Euklidischen Norm $\|\cdot\|_2$ vollständig.*

Beweis. Sei (a_n) eine Cauchyfolge in $(\mathbb{R}^k, \|\cdot\|_2)$. Dann gilt:

$$|a_{nj} - a_{mj}| \leq \|a_n - a_m\|_2$$

für alle $j \in \{1, \dots, k\}$. Somit ist $(a_{nj})_{n \in \mathbb{N}}$ eine Cauchyfolge in \mathbb{R} . Dann gilt $\lim_{n \rightarrow \infty} a_{nj} = b_j$ und aus Lemma 14.2.2 folgt die Konvergenz von a_n mit

$$\lim_{n \rightarrow \infty} a_n = b = (b_1, \dots, b_k).$$

□

Bemerkung. Wie wir bald sehen werden, ist der Vektorraum \mathbb{R}^k bezüglich einer beliebigen Norm vollständig. Allgemeiner ist sogar jeder endlich dimensionale Vektorraum vollständig.

Satz 14.2.6. *Sei M eine Menge und $B(M, \mathbb{R})$ der Vektorraum der auf M beschränkten Funktionen. Dann ist $B(M, \mathbb{R})$ bezüglich der Supremumsnorm $\|\cdot\|$ vollständig.*

Beweis. Der Beweis folgt im Wesentlichen aus der Vollständigkeit von \mathbb{R} . Sei $(f_n)_{n \in \mathbb{N}}$ eine Cauchyfolge bezüglich der Supremumsnorm. Wegen

$$|f_n(x) - f_m(x)| \leq \|f_n - f_m\|$$

für alle $x \in M$, ist auch $(f_n(x))_{n \in \mathbb{N}}$ für jedes $x \in M$ eine Cauchyfolge und

$$\lim_{n \rightarrow \infty} f_n(x) =: f(x),$$

d.h. f_n konvergiert punktweise gegen f (Siehe Definition 6.1.1).

Es bleibt zu zeigen: $\|f_n - f\| \rightarrow 0$ und $f \in B(M, \mathbb{R})$. Sei $\epsilon > 0$, so wähle $n_0 \in \mathbb{N}$, so dass

$$\|f_n - f_m\| < \epsilon$$

für alle $n, m \geq n_0$. Ist insbesondere $n \geq n_0$, so gilt für alle $m \geq n_0$ und $x \in M$:

$$|f(x) - f_n(x)| \leq |f(x) - f_m(x)| + |f_m(x) - f_n(x)| < |f(x) - f_m(x)| + \epsilon.$$

Wegen $\lim_{m \rightarrow \infty} f_m(x) =: f(x)$ gilt dann auch: $|f(x) - f_n(x)| \leq \epsilon$ und somit ist $\|f_n - f\|$ eine Nullfolge (dies ist äquivalent zu gleichmäßigen Konvergenz der Folge f_n).

Wählt man $n \in \mathbb{N}$, so dass $\|f - f_n\| \leq 1$, so folgt:

$$\|f\| \leq \|f - f_n\| + \|f_n\| \leq 1 + \|f_n\|.$$

Also ist f beschränkt.

□

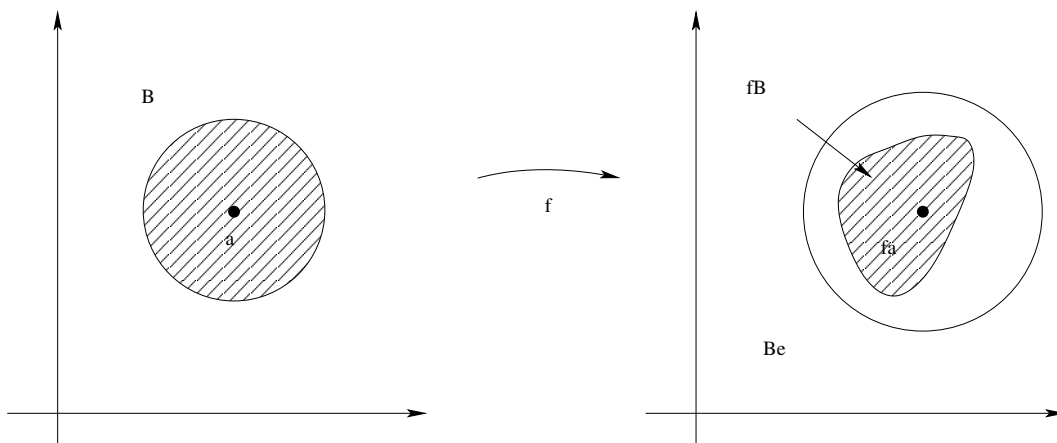


Abbildung 14.1: Stetigkeit einer Abbildung $f : M_1 \rightarrow M_2$ im Punkt a .

Bemerkung. Wesentlich für den Beweis war nur die Vollständigkeit von \mathbb{R} . Falls $(X, \|\cdot\|)$ ein Banachraum und $B(M, X)$ der Raum der beschränkten Funktionen auf M ist, so ist $B(M, X)$ auch bezüglich der Supremumsnorm $\|f\| = \sup\{\|f(x)\| \mid x \in M\}$ ein Banachraum.

Nun wollen wir den Begriff der Stetigkeit auf metrische Räume übertragen.

Definition 14.2.7. Es seien (M_1, d_1) und (M_2, d_2) zwei metrische Räume und $f : M_1 \rightarrow M_2$ eine Abbildung.

- (a) f heißt *stetig* in $a \in M_1$, falls für jedes $\epsilon > 0$ ein $\delta > 0$ existiert mit

$$f(B(a, \delta)) \subset B(f(a), \epsilon)$$

(siehe Abbildung 14.1).

- (b) f heißt *stetig*, wenn f in jedem Punkt $a \in M_1$ stetig ist.

- (c) f heißt *Lipschitz-stetig*, falls eine Konstante $c > 0$ existiert mit

$$d_2(f(x), f(y)) \leq c \cdot d_1(x, y)$$

für alle $x, y \in M_1$.

Bemerkungen.

- (a) Ist f Lipschitz-stetig, so ist f auch stetig.
- (b) Ist $D \subset M_1$, so ist D bezüglich der auf D eingeschränkten Metrik wieder ein metrischer Raum. Dann heißt eine Abbildung $f : D \rightarrow M_2$ stetig in $a \in D$, falls f bezüglich der auf D und M_2 gegebenen Metriken stetig ist.
- (c) Auch der Begriff des Grenzwertes einer Abbildung lässt sich auf metrische Räume übertragen (siehe Definition 3.4.3 und nachfolgende Bemerkungen). Seien (M_1, d_1) und

(M_2, d_2) zwei metrische Räume und $a \in M_1$ sei Berührungspunkt einer Menge $D \subset M_1$, d.h. $B(a, \epsilon) \cap D \neq \emptyset$ für alle $\epsilon > 0$ (siehe auch Definition 14.3.2). Sei $f : D \rightarrow M_2$ eine Abbildung, so heißt $b \in M_2$ Grenzwert von f in a , falls die Abbildung $g : D \cup \{a\} \rightarrow M_2$ mit

$$g(x) = \begin{cases} f(x), & \text{für } x \in D, \\ b, & \text{für } x = a. \end{cases}$$

stetig in a ist. Insbesondere lässt sich f stetig in a fortsetzen. Ist $b \in M_2$ Grenzwert von f in a , so schreiben wir

$$\lim_{\substack{x \rightarrow a \\ x \in D}} f(x) = b,$$

oder auch $\lim_{x \rightarrow a} f(x)$, wenn klar ist, auf welchen Definitionsbereich sich der Grenzwert der Funktion f bezieht.

Satz 14.2.8. Seien (M_1, d_1) und (M_2, d_2) zwei metrische Räume und $f : M_1 \rightarrow M_2$ eine Abbildung. Dann ist $f : M_1 \rightarrow M_2$ stetig in $a \in M_1$ genau dann, falls für jede Folge (a_n) in M_1 mit $\lim_{n \rightarrow \infty} a_n = a$ folgt:

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

Beweis. Fast wörtliche Übertragung des entsprechenden Satzes 3.1.3. □

Außerdem gilt:

Satz 14.2.9. Sei (M, d) ein metrischer Raum und seien $f, g : M \rightarrow \mathbb{C}$ Funktionen, die stetig in $a \in M$ sind. Dann sind

$$f + g : M \rightarrow \mathbb{C} \quad \text{sowie} \quad f \cdot g : M \rightarrow \mathbb{C}$$

stetig in a . Seien (M_1, d_1) , (M_2, d_2) und (M_3, d_3) drei metrische Räume und $f : M_1 \rightarrow M_2$, $g : M_2 \rightarrow M_3$ zwei Abbildungen. Ist f stetig in $a \in M_1$ und g stetig in $f(a) \in M_2$, so ist

$$g \circ f : M_1 \rightarrow M_3$$

stetig in a .

Beweis. Die Beweise sind analog zu den Beweisen der Sätze 3.1.4 und 3.1.5. □

Satz 14.2.10. Sei (M, d) ein metrischer Raum und $f_n : M \rightarrow \mathbb{R}$ eine gleichmäßig konvergente Folge von stetigen Funktionen mit Grenzwert f . Dann ist f stetig.

Beweis. Sei $a \in M$. Für jedes $x \in M$ gilt:

$$\begin{aligned} d(f(x), f(a)) &\leq d(f(x), f_n(x)) + d(f_n(x), f(a)) \\ &\leq d(f(x), f_n(x)) + d(f_n(x), f_n(a)) + d(f_n(a), f(a)). \end{aligned}$$

Sei $\epsilon > 0$. Dann existiert wegen der gleichmäßigen Konvergenz der Folge f_n ein $n \in \mathbb{N}$, so dass

$$d(f(y), f_n(y)) < \epsilon/3$$

für alle $y \in M$. Wähle $\delta > 0$, so dass

$$d(f_n(x), f_n(a)) < \epsilon/3$$

für alle $x \in B(a, \delta)$ (siehe Abbildung 14.2). Dann folgt:

$$d(f(x), f(a)) < \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon$$

für alle $x \in B(a, \delta)$, d.h. f ist stetig in a . □

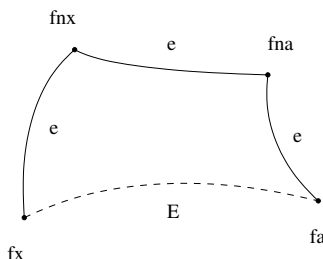


Abbildung 14.2:

Mit $C^0(M, \mathbb{R})$ bezeichnen wir die stetigen Funktionen auf (M, d) und mit $C_b^0(M, \mathbb{R})$ die beschränkten stetigen Funktionen.

Korollar 14.2.11. *Der Vektorraum $C_b^0(M, \mathbb{R})$ der auf M beschränkten stetigen Funktionen ist ein Untervektorraum von $B(M, \mathbb{R})$. Bezüglich der Supremumsnorm ist $C_b^0(M, \mathbb{R})$ vollständig.*

Beweis. Ist (f_n) eine Cauchyfolge in $C_b^0(M, \mathbb{R})$, so auch in $B(M, \mathbb{R})$. Da $B(M, \mathbb{R})$ vollständig ist, existiert $f \in B(M, \mathbb{R})$ mit $\|f_n - f\| \rightarrow 0$, $n \rightarrow \infty$. Wegen Satz 14.2.10 ist f stetig, d.h. $f \in C_b^0(M, \mathbb{R})$ und somit ist $C_b^0(M, \mathbb{R})$ vollständig. \square

Bemerkung. Sei (M, d) ein metrischer Raum und $f : M \rightarrow \mathbb{R}^m$ eine Abbildung. Sei $f(x) = (f_1(x), \dots, f_m(x))$, wobei $f_j : M \rightarrow \mathbb{R}$ die Komponentenfunktionen von f beschreiben. Dann gilt: f ist genau dann stetig, falls alle ihre Komponentenfunktionen stetig sind. Denn ist $a_n \in M$ eine konvergente Folge mit $\lim_{k \rightarrow \infty} a_n = a$, so folgt aus Lemma 14.2.2 und nachfolgender Bemerkung:

$$\lim_{k \rightarrow \infty} f(a_k) = f(a) = (f_1(a), \dots, f_m(a)) \Leftrightarrow \lim_{k \rightarrow \infty} f_j(a_k) = f_j(a)$$

für alle $j \in \{1, \dots, m\}$.

Beispiel. Polynome mehrerer Veränderlichen sind Abbildungen $p : \mathbb{R}^n \rightarrow \mathbb{R}$ mit

$$p(x_1, \dots, x_n) = \sum_{i_1=0}^{k_1} \cdots \sum_{i_n=0}^{k_n} a_{i_1 \dots i_n} x_1^{i_1} \cdots x_n^{i_n}$$

mit Koeffizienten $a_{i_1 \dots i_n} \in \mathbb{R}$. p ist stetig, denn jedes Monom $x_1^{i_1} \cdots x_n^{i_n}$ als Produkt stetiger Funktionen auf \mathbb{R}^n stetig ist. Mit

$$\text{Grad } p := \max\{i_1 + \cdots + i_n \mid a_{i_1 \dots i_n} \neq 0\}$$

bezeichnen wir den Grad von p . Insbesondere ist jede lineare Abbildung $L : \mathbb{R}^n \rightarrow \mathbb{R}$ stetig, denn $L(x_1, \dots, x_n) = a_1 x_1 + \cdots + a_n x_n$.

Wegen obiger Bemerkung ist auch jede lineare Abbildung $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$ stetig, denn die Komponentenfunktionen sind linear und damit stetig.

Zwischen unendlich dimensionalen Vektorräumen müssen lineare Abbildungen nicht stetig sein (siehe Beispiel unten). Es gilt aber der Satz:

Satz 14.2.12. *Es seien $(V, \|\cdot\|_V)$ und $(W, \|\cdot\|_W)$ zwei normierte Vektorräume und $L : V \rightarrow W$ eine lineare Abbildung. Dann sind folgende Aussagen äquivalent:*

- (1) L ist stetig auf ganz V .
- (2) L ist im Nullpunkt stetig.
- (3) Es existiert ein $c > 0$ mit $\|Lx\|_W \leq c\|x\|_V$ für alle $x \in V$.

Beweis. (1) \Rightarrow (2) ist trivial.

(2) \Rightarrow (3). Ist L im Nullpunkt stetig, so existiert zu $\epsilon = 1$ ein $\delta > 0$ mit $L(B(0, \delta)) \subset B(0, 1) \subset W$. Daraus folgt für alle $x \neq 0$:

$$\left\| L \left(\frac{x}{\|x\|_V} \cdot \frac{\delta}{2} \right) \right\|_W < 1,$$

denn

$$\left\| \frac{x}{\|x\|_V} \cdot \frac{\delta}{2} \right\|_V = \frac{\delta}{2} \left\| \frac{x}{\|x\|_V} \right\|_V = \frac{\delta}{2} < \delta,$$

und da L linear ist, folgt:

$$\|L(x)\|_W \leq \frac{2}{\delta} \|x\|_V.$$

(3) \Rightarrow (1), denn

$$\|L(x) - L(y)\|_W = \|L(x - y)\|_W \leq c\|x - y\|_V,$$

d.h. L ist Lipschitz-stetig und somit ist L stetig. \square

Beispiel. Seien $C^0[0, \pi]$ bzw. $C^1[0, \pi]$ die Vektorräume der auf dem Intervall $[0, \pi]$ stetigen bzw. 1-mal stetig differenzierbaren reellwertigen Funktionen. Dann ist

$$D : C^1[0, \pi] \rightarrow C^0[0, \pi] \quad \text{mit} \quad Df = f'$$

eine lineare Abbildung. Wählen wir auf $C^1[0, \pi]$ und $C^0[0, \pi]$ die Supremumsnorm, so ist D nicht stetig. Man betrachte hierzu die Folge $f_n \in C^1[0, \pi]$ mit $f_n(x) = \frac{1}{n} \sin(nx)$. Dann gilt

$$\|f_n\| = \frac{1}{n} \rightarrow 0 \quad \text{und} \quad \|Df_n\| = \|\cos(nx)\| = 1,$$

und somit ist D nicht im Nullpunkt stetig (siehe Abbildung 14.3).

Definition 14.2.13. Sind $(X, \|\cdot\|)$ und $(Y, \|\cdot\|)$ normierte Vektorräume, so bezeichnen wir mit $\mathcal{L}(X, Y)$ die Menge der *stetigen linearen Abbildungen* $A : X \rightarrow Y$. Außerdem ist $(\mathcal{L}(X, Y), \|\cdot\|)$ wieder ein normierter Vektorraum mit

$$\|A\| = \sup\{\|Ax\| \mid x \in X, \|x\| = 1\}$$

für alle $A \in \mathcal{L}(X, Y)$. Diese Norm heißt auch *Operatornorm*.

Bemerkungen.

- (a) Die Überprüfung der Normaxiome ist als Übung überlassen (Übungsblatt 5, Mathematik für Physiker III, Aufgabe 3).

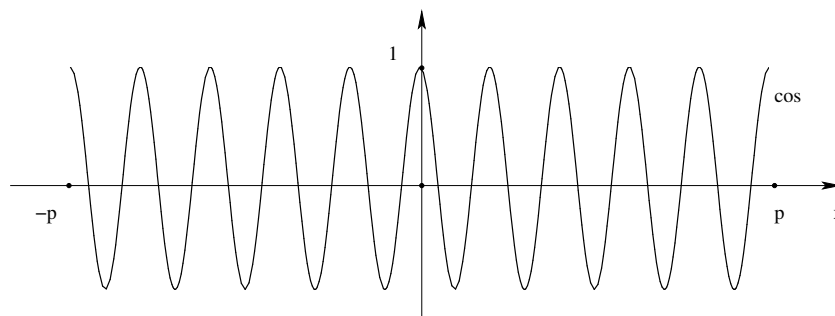


Abbildung 14.3:

- (b) Obwohl in der obigen Definition die Normen der Vektorräume X, Y im Allgemeinen verschieden sind, werden wir sie in Zukunft häufig mit dem gleichen Symbol $\|\cdot\|$ bezeichnen.
- (c) Aus der Definition der Operatornorm folgt: $\|A(x)\| \leq \|A\|\|x\|$ für alle $x \in X$.
- (d) Ist Y vollständig, so auch $\mathcal{L}(X, Y)$ bezüglich der Operatornorm. Denn ist $A_n \in \mathcal{L}(X, Y)$ eine Cauchyfolge, so ist für jedes $x \in X$ auch die Folge $A_n(x)$ wegen

$$\|A_n(x) - A_m(x)\| \leq \|A_n - A_m\|\|x\|$$

eine Cauchyfolge. Dann zeigt man analog zu Satz 14.2.6 und Satz 14.2.10, dass der Grenzwert

$$A(x) := \lim_{n \rightarrow \infty} A_n(x)$$

eine stetige lineare Abbildung definiert mit

$$\lim_{n \rightarrow \infty} \|A_n - A\| = 0.$$

14.3 Topologie in metrischen Räumen

Definition 14.3.1. Sei (M, d) ein metrischer Raum und $U \subset M$. Dann heißt $a \in U$ *innerer Punkt*, falls ein $\epsilon > 0$ existiert mit $B(a, \epsilon) \subset U$. Die Menge U heißt *offen*, falls jeder Punkt in U innerer Punkt ist.

Bemerkungen.

- (a) Für $a \in M$ und $r > 0$ ist die Menge $B(a, r) = \{x \in M \mid d(x, a) < r\}$ offen. Denn ist $x \in B(a, r)$, so ist wegen der Dreiecksungleichung $B(x, \epsilon) \subset B(a, r)$ für $\epsilon = r - d(x, a)$.
- (b) M und die leere Menge sind offen.
- (c) Die Vereinigung beliebig vieler offener Mengen ist offen. Der Durchschnitt endlich vieler offener Mengen ist offen. Der Durchschnitt unendlich vieler offener Mengen ist hingegen nicht immer offen. Man betrachte z.B. die in \mathbb{R} offenen Intervalle $I_n = (-\frac{1}{n}, \frac{1}{n}), n \in \mathbb{N}$. Der Durchschnitt all dieser offenen Mengen besteht nur aus der Null und ist somit nicht offen.

Definition 14.3.2. Sei (M, d) ein metrischer Raum und $A \subset M$. Dann heißt $x \in M$ *Berührungspunkt* von A , falls $B(x, \epsilon) \cap A \neq \emptyset$ für alle $\epsilon > 0$.

Definition 14.3.3. Sei $A \subset M$, so heißt

$$\bar{A} = \{x \in M \mid x \text{ Berührungspunkt von } A\}$$

die *abgeschlossene Hülle* von A . Die Menge A heißt *abgeschlossen*, falls $A = \bar{A}$.

Genau wie in \mathbb{R} gelten die folgenden Charakterisierungen von abgeschlossenen Mengen.

Satz 14.3.4. Sei (M, d) ein metrischer Raum und $A \subset M$. Dann sind folgende Aussagen äquivalent:

- (a) A abgeschlossen.
- (b) $M \setminus A$ offen.
- (c) Jede konvergente Folge $(x_n)_{n \in \mathbb{N}}$ in M mit $x_n \in A$ hat ihren Grenzwert in A .

Beweis. Wir zeigen nur die Äquivalenz von (a) und (c).

Sei A abgeschlossen und $x_n \in A$ eine Folge mit $\lim_{n \rightarrow \infty} x_n = x$. Da $d(x, x_n)$ eine Nullfolge ist, ist x somit Berührungspunkt von A .

Angenommen, jede konvergente Folge in A habe ihren Grenzwert in A . Sei $x \in M$ Berührungspunkt von A . Dann existiert zu jedem $n \in \mathbb{N}$ ein $x_n \in B(x, 1/n) \cap A$. Da $\lim_{n \rightarrow \infty} x_n = x$, ist nach Annahme $x \in A$. □

Bemerkungen. Für metrische Räume M gilt:

- (a) \emptyset und M sind abgeschlossen (und offen).
- (b) Für jedes $p \in M$ und $r \geq 0$ ist $K(p, r) := \{x \in M \mid d(p, x) \leq r\}$ abgeschlossen. $K(p, r)$ heißt *abgeschlossener Ball*.

Warnung:

In allgemeinen metrischen Räumen gilt nicht: $\overline{B(p, r)} = K(p, r)$. Als Beispiel betrachte man auf einer Menge M (z.B. $M = \mathbb{R}$) die „diskrete“ Metrik d , d.h.

$$d(x, y) = \begin{cases} 1, & x \neq y, \\ 0, & x = y. \end{cases}$$

Dann ist $B(p, 1) = \{p\}$ und $\overline{B(p, 1)} = \{p\}$, aber $K(p, 1) = M$. Für normierte Vektorräume ist aber $\overline{B(p, 1)} = K(p, 1)$ (Siehe Übungsblatt 6, Mathematik für Physiker III, Aufgabe 1).

Stetige Abbildungen können mittels offener Mengen wie folgt charakterisiert werden.

Satz 14.3.5. Seien $(M_1, d_1), (M_2, d_2)$ metrische Räume und $f : M_1 \rightarrow M_2$ eine Abbildung. Dann sind äquivalent:

- (1) f ist stetig.

(2) Für alle offenen Mengen $U \subset M_2$ ist $f^{-1}(U) \subset M_1$ offen.

Beweis. Sei (1) erfüllt und $U \subset M_2$ offen. Ist $x \in f^{-1}(U)$, so existiert zu $f(x) \in U$ ein $\epsilon > 0$ mit $B(f(x), \epsilon) \subset U$. Man wähle $\delta > 0$, so dass $f(B(x, \delta)) \subset B(f(x), \epsilon)$. Dann ist $B(x, \delta) \subset f^{-1}(U)$, d.h. x ist innerer Punkt von $f^{-1}(U)$.

Sei (2) erfüllt und $x \in M_1$. Für jedes $\epsilon > 0$ ist $B(f(x), \epsilon)$ offen und somit ist auch $f^{-1}(B(f(x), \epsilon))$ offen. Insbesondere ist x innerer Punkt von $f^{-1}(B(f(x), \epsilon))$, d.h. es existiert ein $\delta > 0$ mit $B(x, \delta) \subset f^{-1}(B(f(x), \epsilon))$. Also gilt $f(B(x, \delta)) \subset B(f(x), \epsilon)$ und somit ist f stetig in x . \square

Bemerkungen.

- (a) f ist also genau dann stetig, falls das Urbild offener Mengen offen ist. f ist aber auch genau dann stetig, falls das Urbild abgeschlossener Mengen abgeschlossen ist. Denn ist $A \subset M_2$ abgeschlossen, so ist $M_2 \setminus A$ offen. Ist f stetig, so ist auch $f^{-1}(M_2 \setminus A) = M_1 \setminus f^{-1}(A)$ offen. Dann ist aber $f^{-1}(A)$ abgeschlossen. Die Umkehrung beweist man genauso.
- (b) Bilder offener Mengen müssen unter stetigen Abbildungen nicht offen sein. Man betrachte $f: \mathbb{R} \rightarrow \mathbb{R}$ mit $f(x) = x^2$. Dann ist $f((-1, +1)) = [0, 1)$.

Nun wollen wir den Begriff der kompakten Menge verallgemeinern. In \mathbb{R} sind es die abgeschlossenen und beschränkten Mengen. In allgemeinen metrischen Räumen ist die Definition komplizierter. Die folgende Definition gilt sogar für alle „topologischen Räume“. Diese werden wir aber im Rahmen dieser Vorlesung nicht behandeln.

Definition 14.3.6. Sei (M, d) ein metrischer Raum. Eine Menge $K \subset M$ heißt *kompakt* genau dann, wenn für jede Überdeckung von K durch offene Mengen eine endliche Teilüberdeckung existiert. Dies bedeutet:

Ist $(U_\alpha)_{\alpha \in I}$ eine Familie von offenen Mengen mit $K \subset \bigcup_{\alpha \in I} U_\alpha$, so existieren $\alpha_1, \dots, \alpha_\ell \in I$

mit $K \subset \bigcup_{i=1}^{\ell} U_{\alpha_i}$.

In metrischen Räumen lässt sich Kompaktheit auch mittels Folgen charakterisieren.

Definition 14.3.7. Sei $K \subset M$ Teilmenge eines metrischen Raumes M . Dann heißt K *folgenkompakt*, wenn zu jeder Folge (x_n) mit $x_n \in K$ eine konvergente Teilfolge x_{k_n} mit Grenzwert in K existiert.

Satz 14.3.8. Sei $K \subset M$ folgenkompakt. Dann ist K abgeschlossen und beschränkt.

Bemerkung. Eine Menge $A \subset (M, d)$ heißt *beschränkt*, falls ein $r > 0$ und $p \in M$ existieren mit $A \subset B(p, r)$.

Beweis. Sei K folgenkompakt und $x_n \in K$ eine konvergente Folge mit Grenzwert x . Dann ist $x \in K$, denn jede Teilfolge hat ebenfalls x als Grenzwert. Damit ist K abgeschlossen.

Sei K nicht beschränkt und $p \in K$. Dann existiert zu jedem $n \in \mathbb{N}$ ein $x_n \in K$ mit $d(x_n, p) \geq n$. Somit hat x_n keine konvergente Teilfolge, denn wäre $\lim_{n \rightarrow \infty} x_{k_n} = x$ für eine Teilfolge x_{k_n} , so würde ein $n_0 \in \mathbb{N}$ existieren mit $d(x_{k_n}, x) \leq 1$ für alle $n \geq n_0$. Dann würde

$$d(x_{k_n}, p) \leq d(x_{k_n}, x) + d(x, p) \leq 1 + d(x, p)$$

für $n \geq n_0$ folgen. Da $d(x_{k_n}, p) \geq k_n \geq n$ ist dies ein Widerspruch. \square

Bemerkung. Ist eine Menge abgeschlossen und beschränkt, so muss sie nicht folgenkompakt sein. Man betrachte dazu den normierten Vektorraum

$$B(\mathbb{N}, \mathbb{R}) = \{f : \mathbb{N} \rightarrow \mathbb{R} \mid f \text{ ist beschränkt}\}$$

mit der Supremumsnorm $\|\cdot\|$ ($B(\mathbb{N}, \mathbb{R})$ ist nichts Anderes, als der Raum der beschränkten Folgen). Dann ist

$$K(0, 1) = \{f \in B(\mathbb{N}, \mathbb{R}) \mid \|f\| \leq 1\}$$

nicht folgenkompakt. Man betrachte hierzu die Folge $f_n \in B(\mathbb{N}, \mathbb{R})$ mit

$$f_n(m) = \begin{cases} 0, & n \neq m, \\ 1, & n = m. \end{cases}$$

Dann ist $\|f_n\| = 1$ und $\|f_n - f_k\| = 1$, falls $n \neq k$. Dies gilt wegen

$$(f_n - f_k)(m) = \begin{cases} 1, & \text{falls } m = n, \\ -1, & \text{falls } m = k, \\ 0, & \text{sonst.} \end{cases}$$

Daher besitzt f_n keine HP und somit keine konvergente Teilfolge.

In metrischen Räumen sind Folgenkompaktheit und Kompaktheit äquivalent.

Satz 14.3.9. Sei (M, d) ein metrischer Raum. $K \subset M$ ist genau dann kompakt, wenn K folgenkompakt ist.

Beweis. Wir zeigen zunächst: K kompakt $\Rightarrow K$ folgenkompakt. Wir benutzen das Kontrapositionsgesetz und zeigen die äquivalente Aussage:

K ist nicht folgenkompakt $\Rightarrow K$ nicht kompakt.

Ist K nicht folgenkompakt, so existiert eine Folge $(x_n), x_n \in K$, so dass x_n keine konvergente Teilfolge mit Grenzwert in K besitzt. Daraus folgt: Kein $x \in K$ ist Häufungspunkt der Folge x_n , d.h. für alle $x \in K$ existiert ein $\epsilon = \epsilon(x) > 0$, sodass die Menge

$$\{n \in \mathbb{N} \mid x_n \in B(x, \epsilon(x))\}$$

endlich ist (siehe Bemerkung (c) nach Definition 14.2.1). Man betrachte die offene Überdeckung

$$\bigcup_{x \in K} B(x, \epsilon(x)) \supset K$$

von K . Diese hat keine endliche Teilüberdeckung, denn sonst hätte die Folge (x_n) nur endlich viele Folgenglieder.

Um die Umkehrung zeigen zu können, benötigen wir folgendes Lemma.

Lemma 14.3.10. Sei $K \subset M$ folgenkompakt. Dann folgt:

(a) Zu jedem $r > 0$ existieren endlich viele $x_1, \dots, x_k \in K$ mit $\bigcup_{i=1}^k B(x_i, r) \supset K$.

- (b) Zu jeder Familie von offenen Mengen $(U_\alpha)_{\alpha \in I}$, die K überdeckt, existiert ein $r > 0$, so dass für alle $x \in K$ ein $\alpha \in I$ existiert mit $B(x, r) \subset U_\alpha$. Die Zahl $r > 0$ heißt auch Lebesgue-Zahl der Überdeckung.

Beweis.

- (a) Wir benutzen wieder das Kontrapositionsgesetz und zeigen: existiert ein $r > 0$, so dass für jede endliche Menge $T \subset K$ gilt:

$$\bigcup_{y \in T} B(y, r)$$

ist keine Überdeckung von K , d.h.

$$K \setminus \bigcup_{y \in T} B(y, r) \neq \emptyset,$$

so ist K nicht folgenkompakt. Wir konstruieren eine Folge in K , die keine konvergente Teilfolge besitzt. Man wähle $y = y_1 \in K$ beliebig. Dann ist

$$K \setminus B(y_1, r) \neq \emptyset.$$

Seien $y_1, \dots, y_n \in K$ schon konstruiert, so ist $K \setminus \bigcup_{i=1}^n B(y_i, r) \neq \emptyset$. Man wähle

$$y_{n+1} \in K \setminus \bigcup_{i=1}^n B(y_i, r).$$

Insbesondere gilt für die so konstruierte Folge: $y_n \notin B(y_m, r)$ für $n > m$. Also ist $d(y_n, y_m) \geq r$ für $n \neq m$. Damit hat die Folge (y_n) keine Häufungspunkte und besitzt somit auch keine konvergente Teilfolge.

- (b) Sei also $(U_\alpha)_{\alpha \in I}$ eine Familie von offenen Mengen, die K überdeckt. Würde zu dieser Überdeckung keine positive Lebesgue-Zahl existieren, so gäbe es zu jedem $n \in \mathbb{N}$ ein $x_n \in K$, mit $B(x_n, \frac{1}{n}) \not\subset U_\alpha$ für alle $\alpha \in I$. Da K folgenkompakt ist, existiert dann eine konvergente Teilfolge x_{k_n} mit $\lim_{n \rightarrow \infty} x_{k_n} = x \in K$ und ein $\alpha \in I$ mit $x \in U_\alpha$. Da U_α offen ist, ist x innerer Punkt und es existiert ein $\epsilon > 0$ mit $B(x, \epsilon) \subset U_\alpha$. Man wähle $n \in \mathbb{N}$ so groß, dass $x_{k_n} \in B(x, \epsilon/2)$ für $n \geq n_0$. Dann ist

$$B(x_{k_n}, \epsilon/2) \subset B(x, \epsilon) \subset U_\alpha,$$

und wir erhalten für $\frac{1}{k_n} \leq \epsilon/2$ und $n \geq n_0$ einen Widerspruch.

□

Nun können wir den Beweis von Satz 14.3.9 vervollständigen, indem wir zeigen: Ist $K \subset M$ folgenkompakt $\Rightarrow K$ ist kompakt.

Sei $(U_\alpha)_{\alpha \in I}$ eine Familie von Mengen, die K überdeckt und $r > 0$ eine Lebesgue-Zahl dieser Überdeckung. Dann folgt aus Teil (a) von Lemma 14.3.10 die Existenz von endlich vielen Punkten $x_1, \dots, x_k \in K$ mit

$$\bigcup_{i=1}^k B(x_i, r) \supset K.$$

Da r eine Lebesgue-Zahl der Überdeckung von K mit den offenen Mengen $(U_\alpha)_{\alpha \in I}$ ist, existiert zu jedem x_i ein $\alpha_i \in U_{\alpha_i}$ mit $B(x_i, r) \subset U_{\alpha_i}$. Somit erhalten wir:

$$K \subset \bigcup_{i=1}^k B(x_i, r) \subset \bigcup_{i=1}^k U_{\alpha_i}.$$

□

Satz 14.3.11. *Seien (M_1, d_1) und (M_2, d_2) metrische Räume und $f : M_1 \rightarrow M_2$ stetig. Ist $K \subset M_1$ kompakt, so ist auch $f(K) \subset M_2$ kompakt.*

Beweis. Sei $(U_\alpha)_{\alpha \in I}$ eine offene Überdeckung von $f(K)$, d.h. $f(K) \subset \bigcup_{\alpha \in I} U_\alpha$. Da wegen Satz 1.1.12

$$f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B) \quad \text{und} \quad f^{-1}(f(A)) \supset A,$$

folgt: $K \subset f^{-1}(f(K)) \subset f^{-1}\left(\bigcup_{\alpha \in I} U_\alpha\right) = \bigcup_{\alpha \in I} f^{-1}(U_\alpha)$. Da $f^{-1}(U_\alpha)$ offen ist, existiert wegen der Kompaktheit von K eine endliche Teilüberdeckung $f^{-1}(U_{\alpha_1}), \dots, f^{-1}(U_{\alpha_n})$ von K , d.h. $\bigcup_{i=1}^n f^{-1}(U_{\alpha_i}) \supset K$. Dann folgt: $U_{\alpha_1}, \dots, U_{\alpha_n}$ überdecken $f(K)$ (denn $f(A \cup B) = f(A) \cup f(B)$, Satz 1.1.12). □

Bemerkung. Man kann den Beweis unter Benutzung der Folgenkompaktheit von K durchführen.

Korollar 14.3.12. *Sei M ein metrischer Raum und $K \subset M$ eine kompakte Teilmenge. Ist $f : M \rightarrow \mathbb{R}$ stetig, so nimmt f auf K Maximum und Minimum an.*

Beweis. Wegen Satz 14.3.11 ist $f(K) \subset \mathbb{R}$ kompakt. In \mathbb{R} sind kompakte Mengen abgeschlossen und beschränkt. Aus den Definitionen von Supremum und Infimum folgt: $\sup f(K) \in f(K)$ und $\inf f(K) \in f(K)$. Somit existieren $x_M, x_m \in K$ mit $f(x_M) = \sup f(K)$ und $f(x_m) = \inf f(K)$. □

Korollar 14.3.13. *Sei $K \subset M$ eine kompakte Teilmenge eines metrischen Raumes. Dann ist jede abgeschlossene Menge $A \subset K$ kompakt.*

Außerdem sind kompakte metrische Räume immer vollständig.

Beweis. Zum Beweis nutzen wir die Folgenkompaktheit von K . Sei $x_n \in A \subset K$ eine Folge, so besitzt x_n eine konvergente Teilfolge x_{k_n} . Da A abgeschlossen ist, gilt: $\lim_{n \rightarrow \infty} x_{k_n} \in A$. Der zweite Teil der Aussage sei als Übung überlassen (Übungsblatt 6, Mathematik für Physiker III, Aufgabe 2(b)). Zum Beweis muss man sich nur davon überzeugen, dass Cauchyfolgen konvergieren, wenn sie eine konvergente Teilfolge besitzen. □

Im \mathbb{R}^k sind die kompakten genau die abgeschlossenen und beschränkten Mengen.

Wir zeigen dies zunächst bezüglich der Maximumsnorm $\|x\|_\infty = \max\{|x_i| \mid i \in \{1, \dots, k\}\}$.

Satz 14.3.14. (Heine-Borel)

Sei $A \subset (\mathbb{R}^k, \|\cdot\|_\infty)$. Dann ist A kompakt genau dann, falls A abgeschlossen und beschränkt ist.

Beweis. Zu zeigen bleibt: ist A abgeschlossen und beschränkt, so ist A auch kompakt. Es genügt zu zeigen: die Menge

$$I^k := I \times \cdots \times I$$

ist für jedes Intervall $I = [-a, a]$ mit $a > 0$ kompakt. Denn ist A eine beliebige abgeschlossene und beschränkte Teilmenge von \mathbb{R}^k , so existiert ein $a > 0$ mit

$$A \subset K(0, a) = \{x \in \mathbb{R}^k \mid \|x\|_\infty \leq a\} = I^k$$

mit $I = [-a, a]$. Ist I^k kompakt, so ist wegen Korollar 14.3.13 auch die Menge A als abgeschlossene Teilmenge kompakt.

Wir zeigen durch Induktion über k , dass I^k folgenkompakt ist. Für $k = 1$ folgt dies aus Analysis I (siehe Satz 3.3.3 und Beispiel (1) nach Definition 3.3.1). Sei I^k folgenkompakt und man betrachte $I^{k+1} = I^k \times I$. Ist $z_n = (x_n, y_n) \in I^k \times I$ eine Folge, so besitzt x_n eine konvergente Teilfolge $\tilde{x}_n := x_{k_n}$ mit $\lim_{n \rightarrow \infty} \tilde{x}_n = x \in I^k$. Man betrachte die Teilfolge $\tilde{z}_n = (\tilde{x}_n, \tilde{y}_n)$ der Folge (z_n) mit $\tilde{z}_n = z_{k_n}$. Da I kompakt ist, besitzt \tilde{y}_n eine konvergente Teilfolge \tilde{y}_{l_n} mit $\lim_{n \rightarrow \infty} \tilde{y}_{l_n} = y \in [-a, +a]$. Dann gilt:

$$\lim_{n \rightarrow \infty} \tilde{z}_{l_n} = \lim_{n \rightarrow \infty} (\tilde{x}_{l_n}, \tilde{y}_{l_n}) = (x, y) \in I^k \times I,$$

d.h. \tilde{z}_n und somit auch z_n besitzen eine konvergente Teilfolge mit Grenzwert in $I^k \times I = I^{k+1}$. \square

Definition 14.3.15. Es sei V ein Vektorraum. Eine Norm $\|\cdot\|_1$ heißt äquivalent zu einer Norm $\|\cdot\|_2$, falls Konstanten $0 < a \leq b$ existieren mit

$$a\|x\|_1 \leq \|x\|_2 \leq b\|x\|_1 \quad (*)$$

für alle $x \in V$.

Bemerkungen.

(a) Ist $\|\cdot\|_1$ äquivalent zu $\|\cdot\|_2$, so ist $\|\cdot\|_2$ auch äquivalent zu $\|\cdot\|_1$, denn aus (*) folgt

$$\frac{1}{b}\|x\|_2 \leq \|x\|_1 \leq \frac{1}{a}\|x\|_2.$$

Ist $\|\cdot\|_1$ äquivalent zu $\|\cdot\|_2$ und $\|\cdot\|_2$ äquivalent zu $\|\cdot\|_3$, so ist $\|\cdot\|_1$ äquivalent zu $\|\cdot\|_3$. Damit definiert der Begriff der Äquivalenz von Normen eine Äquivalenzrelation auf der Menge aller Normen eines Vektorraumes V .

(b) Zwei äquivalente Normen haben die gleichen konvergenten Folgen und Cauchyfolgen. Außerdem stimmen die beschränkten, abgeschlossen, offen oder kompakten Mengen bezüglich äquivalenter Normen überein.

Wir zeigen nun, dass auf \mathbb{R}^k zwei beliebige Normen äquivalent sind. Wegen Teil (a) der obigen Bemerkung genügt es zu zeigen, dass jede Norm äquivalent zur Maximumsnorm ist. Dies wird eine Konsequenz aus folgendem Satz sein.

Satz 14.3.16. Sei $\|\cdot\|$ eine Norm auf \mathbb{R}^k . Dann ist $f : (\mathbb{R}^k, \|\cdot\|_\infty) \rightarrow \mathbb{R}$ mit $f(x) = \|x\|$ stetig.

Beweis.

$$|f(x) - f(y)| = | \|x\| - \|y\| | \leq \|x - y\|,$$

denn es gilt

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

wegen

$$\|x\| = \|x - y + y\| \leq \|x - y\| + \|y\| \quad \text{und} \quad \|y\| \leq \|y - x\| + \|x\|.$$

Ist $x = \sum_{i=1}^k x_i e_i$ mit $e_i = (0, \dots, 1, \dots, 0)$, so gilt:

$$\begin{aligned} \|x\| &\leq \sum_{i=1}^k |x_i| \|e_i\| \leq \left(\max_{1 \leq i \leq k} \|e_i\| \right) \sum_{i=1}^k |x_i| \\ &\leq \max_{1 \leq i \leq k} \|e_i\| \cdot k \cdot \|x\|_\infty = c \|x\|_\infty, \end{aligned}$$

mit $c := \max_{1 \leq i \leq k} \|e_i\| \cdot k$.

Insbesondere gilt: $|f(x) - f(y)| \leq \|x - y\| \leq c \|x - y\|_\infty$. Damit ist f Lipschitz-stetig, also auch stetig. \square

Korollar 14.3.17. Sei $\|\cdot\|$ eine beliebige Norm auf \mathbb{R}^k . Dann existieren Konstanten $0 < a \leq b$ mit

$$a \|x\|_\infty \leq \|x\| \leq b \|x\|_\infty.$$

Beweis. Die Menge

$$S(0, 1) := \{x \in \mathbb{R}^k \mid \|x\|_\infty = 1\}$$

ist abgeschlossen und beschränkt in \mathbb{R}^k und damit kompakt bezüglich $\|\cdot\|_\infty$. Da die Abbildung f stetig ist, nimmt f Minimum und Maximum auf $S(0, 1)$ an, d.h. es existieren $x_M, x_m \in S(0, 1)$ mit

$$f(x_M) = \|x_M\| = \sup\{\|x\| \mid \|x\|_\infty = 1\} =: b$$

und

$$f(x_m) = \|x_m\| = \inf\{\|x\| \mid \|x\|_\infty = 1\} =: a.$$

Dann gilt für jedes $x \neq 0$: $a \leq \left\| \frac{x}{\|x\|_\infty} \right\| \leq b$, denn $\left\| \frac{x}{\|x\|_\infty} \right\|_\infty = 1$, d.h. $\frac{x}{\|x\|_\infty} \in S(0, 1)$. Somit erhalten wir

$$a \|x\|_\infty \leq \|x\| \leq b \|x\|_\infty.$$

\square

Bemerkung. Es folgt sogar aus obigem Korollar und Bemerkung (a) nach Definition 14.3.15, dass auf jedem endlichdimensionalen Vektorraum je zwei Normen äquivalent sind.

Kapitel 15

Differentialrechnung auf Vektorräumen

15.1 Differenzierbarkeit

Wir werden die Differenzierbarkeit in allgemeinen normierten Vektorräumen einführen. Hauptsächlich sind wir aber an der Differenzierbarkeit im \mathbb{R}^n interessiert. Differenzierbare Funktionen werden solche sein, die sich gut durch stetige lineare Abbildungen approximieren lassen. Wie wir gesehen haben, ist die Stetigkeit von linearen Abbildungen zwischen endlich dimensionalen Vektorräumen automatisch erfüllt.

Definition 15.1.1. Es seien X, Y normierte Vektorräume und $U \subset X$ eine offene Menge. Eine Abbildung $f : U \rightarrow Y$ heißt in $a \in U$ *differenzierbar*, falls eine stetige lineare Abbildung $A \in \mathcal{L}(X, Y)$ und eine in a stetige Abbildung $r : U \rightarrow Y$ mit $r(a) = 0$ existiert, mit

$$f(x) = f(a) + A(x - a) + r(x)\|x - a\| \quad (15.1)$$

für alle $x \in U$.

Bemerkungen.

- (a) Um die Differenzierbarkeit einer Abbildung $f : U \rightarrow Y$ in $a \in U$ nachzuweisen, muss man also eine stetige lineare Abbildung A finden, so dass die für $x \neq a$ erklärte Abbildung

$$\tilde{r}(x) := \frac{f(x) - f(a) - A(x - a)}{\|x - a\|}$$

die Bedingung $\lim_{x \rightarrow a} \tilde{r}(x) = 0$ erfüllt. Dies ist äquivalent dazu, dass die Funktion \tilde{r} in a *stetig fortsetzbar* ist (siehe Seite 283), d.h.

$$r(x) = \begin{cases} \tilde{r}(x), & x \neq a, \\ 0, & x = a \end{cases}$$

ist stetig.

- (b) Die Definition hängt von der Wahl der Normen ab. Sind X und Y endlich dimensional, so sind alle Normen äquivalent und die Differenzierbarkeit ist in diesem Falle unabhängig von der Wahl der Normen.

(c) Setzen wir $x = a + h$, so entspricht (15.1) der Gleichung

$$f(a + h) = f(a) + A(h) + \varphi(h)\|h\|, \quad (15.2)$$

wobei $\varphi(h) = r(a + h)$ eine in $0 \in X$ stetige Abbildung ist, mit $\varphi(0) = 0$.

$T(h) := f(a) + A(h)$ kann man als lineare (*affin lineare*) Approximation von f in der Umgebung von a interpretieren.

(d) Ist f in a differenzierbar, so existiert nur eine lineare Abbildung, die (15.1) bzw. (15.2) erfüllt.

Zum Beweis betrachten wir zwei lineare Abbildungen $A_1, A_2 \in \mathcal{L}(X, Y)$ und zwei in 0 stetige Abbildungen φ_1, φ_2 mit $\varphi_1(0) = \varphi_2(0) = 0$, so dass gilt:

$$f(a + h) - f(a) = A_1(h) + \varphi_1(h)\|h\| = A_2(h) + \varphi_2(h)\|h\|$$

für alle $h \in X$ mit $a + h \in U$. Ist $v \in X$ und $t > 0$, so dass $a + tv \in U$, so gilt insbesondere

$$A_1(tv) - A_2(tv) = (\varphi_2(tv) - \varphi_1(tv))\|tv\|$$

und damit

$$t(A_1(v) - A_2(v)) = t(\varphi_2(tv) - \varphi_1(tv))\|v\|.$$

Daraus folgt:

$$A_1(v) - A_2(v) = \|v\| \lim_{t \rightarrow 0} (\varphi_2(tv) - \varphi_1(tv)) = 0.$$

Für den Beweis der Eindeutigkeit haben wir verwendet, dass U eine offene Menge ist.

(e) Ist f in a differenzierbar, so folgt auch die Stetigkeit von f in a , denn wegen (15.2) gilt

$$\lim_{h \rightarrow 0} f(a + h) = f(a) + \lim_{h \rightarrow 0} A(h) + \lim_{h \rightarrow 0} \varphi(h)\|h\| = f(a),$$

denn A ist stetig und $\lim_{h \rightarrow 0} \varphi(h) = 0$.

(f) Die eindeutig bestimmte lineare Abbildung A heißt *Ableitung* oder *Differential* von f an der Stelle a . Wir schreiben $A =: Df(a) = f'(a) \in \mathcal{L}(X, Y)$. Ist $f : U \rightarrow Y$ in jedem $a \in U$ differenzierbar, so heißt $Df : U \rightarrow \mathcal{L}(X, Y)$ mit $a \rightarrow Df(a) \in \mathcal{L}(X, Y)$ die *Ableitung* oder das *Differential* von f .

Beispiele.

- (a) Konstante Abbildungen $f : X \rightarrow Y$ sind differenzierbar. Denn ist $c \in Y$ mit $f(x) = c$ für alle $x \in X$ und $0 \in \mathcal{L}(X, Y)$ die Nullabbildung, so gilt: $f(x) = f(a) + 0(x - a)$. Damit gilt: $Df(a) = 0$.
- (b) Ist $A \in \mathcal{L}(X, Y)$ für alle $a \in X$, so ist A differenzierbar mit Ableitung A , denn $A(x) = A(a) + A(x - a)$. Es gilt also $DA(a) = A$ für alle $a \in X$.

Ist $X = \mathbb{R}$ und $I \subset \mathbb{R}$ ein offenes Intervall, so lässt sich die Differenzierbarkeit auch mittels eines Differenzenquotienten charakterisieren.

Satz 15.1.2. Seien $I \subset \mathbb{R}$, Y ein normierter Vektorraum und $f : I \rightarrow Y$ eine Abbildung. Die Abbildung f ist in $a \in I$ genau dann differenzierbar, falls der Grenzwert

$$\lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h} = b$$

existiert. Außerdem gilt dann $Df(a)(1) = b$ für das Differential $Df(a) : \mathbb{R} \rightarrow Y$.

Beweis. Ist f in a differenzierbar, so folgt:

$$f(a+h) - f(a) = Df(a)h + \varphi(h)|h|,$$

wobei $Df(a) \in \mathcal{L}(\mathbb{R}, Y)$ und φ stetig in $0 \in \mathbb{R}$ ist, mit $\varphi(0) = 0$. Daraus folgt:

$$\frac{f(a+h) - f(a)}{h} = \frac{h}{h} Df(a)(1) + \varphi(h) \frac{|h|}{h} = Df(a)(1) + \varphi(h) \frac{|h|}{h}$$

Da $\lim_{h \rightarrow 0} \varphi(h) = 0$, folgt

$$\lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h} = Df(a)(1).$$

Existiert der Grenzwert

$$\lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h} = b,$$

so definiere man $Df(a) \in \mathcal{L}(\mathbb{R}, Y)$ durch $Df(a)h = h \cdot b$. Dann gilt:

$$\begin{aligned} 0 &= \lim_{h \rightarrow 0} \left\| \frac{f(a+h) - f(a)}{h} - b \right\| = \lim_{h \rightarrow 0} \left\| \frac{f(a+h) - f(a) - Df(a)h}{h} \right\| \\ &= \lim_{h \rightarrow 0} \left\| \frac{f(a+h) - f(a) - Df(a)h}{|h|} \right\| = \lim_{h \rightarrow 0} \|\varphi(h)\|, \end{aligned}$$

und somit ist f in $a \in I$ differenzierbar. □

Bemerkungen.

- (a) Ist $I \subset \mathbb{R}$ ein Intervall, so heißt eine Abbildung $f : I \rightarrow Y$ auch *Kurve* (siehe auch Definition 4.2.1). Ist f in a differenzierbar, so heißt $\dot{f}(a) := Df(a)(1)$ auch die *Geschwindigkeit* von f zum Zeitpunkt a . Ist $\dot{f}(a) \neq 0$, so nennen wir $\dot{f}(a)$ *Tangentenvektor* von f in a .

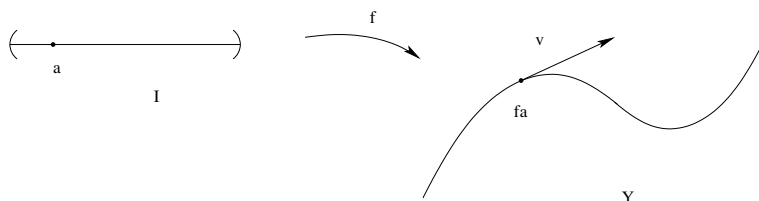


Abbildung 15.1:

(b) Ist $Y = \mathbb{R}^n$, so können wir $f(t) = (f_1(t), \dots, f_n(t))$ komponentenweise schreiben. Dann gilt:

$$\frac{f(a+h) - f(a)}{h} = \left(\frac{f_1(a+h) - f_1(a)}{h}, \dots, \frac{f_n(a+h) - f_n(a)}{h} \right).$$

Daher ist f genau dann in a differenzierbar, falls die Komponentenfunktionen diese Eigenschaft besitzen. Insbesondere gilt, falls f in a differenzierbar ist:

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h} &= \left(\lim_{h \rightarrow 0} \frac{f_1(a+h) - f_1(a)}{h}, \dots, \lim_{h \rightarrow 0} \frac{f_n(a+h) - f_n(a)}{h} \right) \\ &= (f'_1(a), \dots, f'_n(a)) = \dot{f}(a) = Df(a)(1) \end{aligned}$$

Beispiel. Sei $f : \mathbb{R} \rightarrow \mathbb{R}^2$ mit $f(t) = (\cos t, \sin t)$. Dann ist f differenzierbar und es gilt:

$$\dot{f}(t) = (-\sin t, \cos t).$$

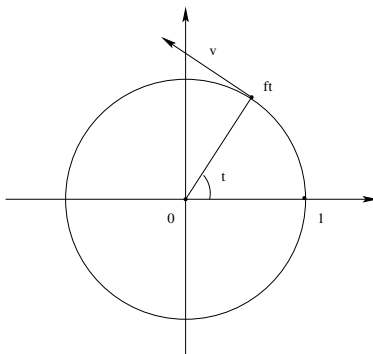


Abbildung 15.2:

Sei nun $U \subset X$ offen und $f : U \rightarrow Y$ eine beliebige Abbildung. Ist $a \in U$, $v \in X$ und $\epsilon > 0$ so gewählt, dass $a + tv \in U$ für alle $t \in (-\epsilon, \epsilon)$, so beschreibt die Kurve

$$g(t) = f(a + tv)$$

das Verhalten von f längs der Geraden $t \mapsto a + tv$. Es gilt nun folgender Satz:

Satz 15.1.3. Sei f in a differenzierbar, so ist auch $g(t) = f(a + tv)$ in $0 \in \mathbb{R}$ differenzierbar und es gilt

$$\dot{g}(0) = Df(a)v.$$

Beweis.

$$\begin{aligned} \frac{g(t) - g(0)}{t} &= \frac{f(a + tv) - f(a)}{t} = \frac{Df(a)tv + \varphi(tv)}{t} \\ &= Df(a)v + \varphi(tv) \frac{|t|}{t} \|v\|, \end{aligned}$$

wobei φ in 0 stetig ist mit $\varphi(0) = 0$. Daraus folgt

$$\lim_{t \rightarrow 0} \frac{g(t) - g(0)}{t} = Df(a)v,$$

denn

$$\lim_{t \rightarrow 0} \left\| \varphi(tv) \frac{|t|}{t} \right\| = \lim_{t \rightarrow 0} \|\varphi(tv)\| = 0.$$

□

Bemerkungen.

- (a) Ist $v \in X$ und $v \neq 0$, so heißt $\dot{g}(0) = Df(a)v$ *Richtungsableitung* von f im Punkte a in Richtung v .
- (b) Ist $B \subset X$ eine Basis von X , so ist $Df(a)$ durch die Richtungsableitungen $Df(a)v$ mit $v \in B$ bestimmt, denn lineare Abbildungen sind durch ihre Werte auf einer Basis eindeutig bestimmt.
- (c) Aus der Existenz der Richtungsableitungen folgt nicht, dass f differenzierbar ist. Als Beispiel betrachte man die Abbildung $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ mit

$$f(x, y) = \begin{cases} \frac{x^2 y}{x^2 + y^2}, & \text{für } (x, y) \neq 0, \\ 0, & \text{sonst.} \end{cases}$$

Dann besitzt f in Nullpunkt in jeder Richtung $v = (v_1, v_2)$ eine Ableitung, denn

$$\frac{f(tv)}{t} = \frac{t^2 v_1^2 \cdot tv_2}{t(t^2 v_1^2 + t^2 v_2^2)} = \frac{v_1^2 \cdot v_2}{v_1^2 + v_2^2} = f(v_1, v_2).$$

Wäre f differenzierbar, so würde aus

$$Df(0)v = \lim_{t \rightarrow 0} \frac{f(tv)}{t} = f(v_1, v_2)$$

auch die Linearität von f folgen. Jedoch ist f nicht linear. Selbst wenn die Richtungsableitungen eine lineare Abbildung definieren, folgt die Differenzierbarkeit im Allgemeinen nicht (siehe z.B. S. 102, Barner-Flohr II). Auch die Stetigkeit von f in einem Punkt folgt nicht aus der Existenz der Richtungsableitungen in diesem Punkt (siehe Aufgabe 3(a), Übungsblatt 7, Mathematik für Physiker III).

Aus Bemerkung (b) folgt, dass sich das Differenzial aus den Richtungsableitungen berechnen lässt. Insbesondere erhalten wir im Falle eines endlich dimensionalen Definitionsbereiches den folgenden Satz.

Satz 15.1.4. *Seien $U \subset X$ offen, X ein Vektorraum endlicher Dimension und Y ein Vektorraum beliebiger Dimension. Sei $f : U \rightarrow Y$ eine Abbildung, die differenzierbar in $a \in U$ ist. Dann gilt: Ist $B = (v_1, \dots, v_n) \subset X$ eine Basis von X , so folgt für $v = x_1 \cdot v_1 + \dots + x_n v_n$:*

$$Df(a)v = \sum_{j=1}^n x_j Df(a)v_j = \sum_{j=1}^n x_j \dot{g}_j(0),$$

wobei

$$\dot{g}_j(0) = \lim_{t \rightarrow 0} \frac{f(a + tv_j) - f(a)}{t}$$

die Richtungsableitungen von f im Punkte a in die Basisrichtungen v_j bezeichnen.

Beweis. Dies folgt unmittelbar aus Satz 15.1.3 und der Linearität von $Df(a)$. \square

Bemerkungen.

- (a) Sind $X = \mathbb{R}^n$, $Y = \mathbb{R}^m$ und (e_1, \dots, e_n) die kanonische Basis auf \mathbb{R}^n , so heißt die Richtungsableitung

$$\frac{\partial f}{\partial x_j}(a) = \lim_{t \rightarrow 0} \frac{f(a + te_j) - f(a)}{t} \in \mathbb{R}^m$$

die j -te partielle Ableitung. Sind $f_1(x), \dots, f_m(x)$ die Komponentenfunktionen von f , so gilt:

$$\begin{aligned} \frac{\partial f}{\partial x_j}(a) &= \left(\lim_{t \rightarrow 0} \frac{f_1(a + te_j) - f_1(a)}{t}, \dots, \lim_{t \rightarrow 0} \frac{f_m(a + te_j) - f_m(a)}{t} \right) \\ &= \left(\frac{\partial f_1}{\partial x_j}(a), \dots, \frac{\partial f_m}{\partial x_j}(a) \right). \end{aligned}$$

- (b) Ist $F_{jk} : \mathbb{R} \rightarrow \mathbb{R}$ die Funktion mit $t \mapsto f_k(a_1, \dots, a_{j-1}, t, a_{j+1}, \dots, a_n) = F_{jk}(t)$, so gilt:
 $F'_{jk}(a_j) = \frac{\partial f_k}{\partial x_j}(a)$.

Mit diesen Bezeichnungen erhalten wir folgendes Korollar zu Satz 15.1.4:

Korollar 15.1.5. Sei $U \subset \mathbb{R}^n$ offen und $f : U \rightarrow \mathbb{R}^m$ eine Abbildung, die in $a \in U$ differenzierbar ist. Dann gilt für $x = (x_1, \dots, x_n)$:

$$Df(a)x = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \dots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \dots & \frac{\partial f_m}{\partial x_n}(a) \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Beweis. Aus Satz 15.1.4 folgt für die kanonische Basis $(e_1, \dots, e_n) \subset \mathbb{R}^n$ und $x = \sum_{j=1}^n x_j e_j = (x_1, \dots, x_n)$:

$$Df(a)x = \sum_{j=1}^n x_j Df(a)e_j = \sum_{j=1}^n x_j \frac{\partial f}{\partial x_j}(a) = \sum_{j=1}^n \begin{pmatrix} \frac{\partial f_1}{\partial x_j}(a) \\ \vdots \\ \frac{\partial f_m}{\partial x_j}(a) \end{pmatrix} x_j.$$

\square

Definition 15.1.6. Seien $U \subset \mathbb{R}^n$ und $f : U \rightarrow \mathbb{R}^m$ in $a \in U$ partiell differenzierbar, d.h. ihre partiellen Ableitungen $\frac{\partial f_i}{\partial x_j}(a)$ existieren. Dann heißt die Matrix

$$\left(\frac{\partial f_i}{\partial x_j}(a) \right) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \cdots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \cdots & \frac{\partial f_m}{\partial x_n}(a) \end{pmatrix}$$

die *Jacobimatrix* von f an der Stelle $a \in U$.

Bemerkung. Wie wir gesehen haben, folgt aus der Existenz der partiellen Ableitungen nicht, dass f differenzierbar ist. Ist f differenzierbar in $a \in U$, so ist die Jacobimatrix die Ableitung von f in a .

Existieren die partiellen Ableitungen in einer Umgebung von a und sind die partiellen Ableitungen stetig, so folgt auch die Umkehrung.

Satz 15.1.7. Seien $U \subset \mathbb{R}^n$ und $f : U \rightarrow \mathbb{R}^m$, so dass die partiellen Ableitungen $\frac{\partial f_i}{\partial x_j}(x)$ für alle $x \in U$ und $i \in \{1, \dots, m\}$, $j \in \{1, \dots, n\}$ existieren. Sind die partiellen Ableitungen in $a \in U$ stetig, so ist f in a differenzierbar.

Beweis. Es genügt, dies im Falle $m = 1$ zu zeigen. (Denn sind die Komponentenfunktionen von f differenzierbar, so auch f .) Wir zeigen:

$$\lim_{h \rightarrow 0} \frac{f(a+h) - f(a) - Df(a)h}{\|h\|_\infty} = 0,$$

wobei $Df(a) : \mathbb{R}^n \rightarrow \mathbb{R}$ eine lineare Abbildung ist, die durch

$$Df(a)h = \frac{\partial f}{\partial x_1}(a)h_1 + \cdots + \frac{\partial f}{\partial x_n}(a)h_n,$$

mit $h = \sum_{i=1}^n h_i e_i = (h_1, \dots, h_n)$ definiert ist. Aufgrund der Äquivalenz der Normen auf \mathbb{R}^n dürfen wir bei der Berechnung des Grenzwertes eine beliebige Norm verwenden. Besonders zwecksmäßig ist die Verwendung der Maximumsnorm. Sei $a_0 := a$, so definiere man

$$\begin{aligned} a_1 &= h_1 e_1 + a_0 \\ a_2 &= h_2 e_2 + a_1 = h_2 e_2 + h_1 e_1 + a_0 \\ a_n &= h_n e_n + a_{n-1} = \sum_{i=1}^n h_i e_i + a_0 = h + a_0 = h + a. \end{aligned}$$

Man betrachte die Teleskopsumme

$$f(a+h) - f(a) = f(a_n) - f(a_{n-1}) + f(a_{n-1}) - f(a_{n-2}) + \cdots + f(a_1) - f(a_0).$$

Sei I_j das abgeschlossene Intervall, das 0 und h_j als Randpunkte besitzt, d.h.

$$I_j = \begin{cases} [0, h_j], & \text{falls } h_j \geq 0, \\ [h_j, 0], & \text{falls } h_j < 0. \end{cases}$$

Man betrachte die Funktion $g_j : I_j \rightarrow \mathbb{R}$ mit $g_j(t) = f(a_{j-1} + te_j)$. Dann gilt: $g_j(0) = f(a_{j-1})$, $g_j(h_j) = f(a_j)$ und $g'_j(t) = \frac{\partial f}{\partial x_j}(a_{j-1} + te_j)$.

Indem wir den Mittelwertsatz 4.4.5 auf die Funktionen g_j anwenden, erhalten wir

$$f(a+h) - f(a) = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(c_j) h_j,$$

wobei $c_j = a_{j-1} + t_j e_j$ mit $t_j \in I_j$. Dann folgt:

$$|f(a+h) - f(a) - Df(a)h| = \left| \sum_{j=1}^n \left(\frac{\partial f}{\partial x_j}(c_j) - \frac{\partial f}{\partial x_j}(a) \right) h_j \right| \leq \|h\|_\infty \sum_{j=1}^n \left| \frac{\partial f}{\partial x_j}(c_j) - \frac{\partial f}{\partial x_j}(a) \right|.$$

Aus der Stetigkeit der partiellen Ableitungen folgt somit:

$$\lim_{h \rightarrow 0} \frac{|f(a+h) - f(a) - Df(a)h|}{\|h\|_\infty} = 0,$$

denn $c_j \rightarrow a$, falls $h_j \rightarrow 0$. □

Sei X ein reeller endlichdimensionaler Vektorraum und $U \subset X$ offen, so ist die Ableitung einer in $a \in U$ differenzierbaren Funktion $f : U \rightarrow \mathbb{R}$ eine lineare Abbildung $Df(a) : X \rightarrow \mathbb{R}$. Also ist $Df(a)$ ein Element des Dualraumes X^* . Ist auf X ein Skalarprodukt gegeben, so lässt sich $Df(a)$ mit Hilfe des Riesz'schen Darstellungssatzes ein Vektor in X zuordnen, der der Gradient von f in a genannt wird.

Definition 15.1.8. Sei $U \subset X$ eine offene Teilmenge eines endlich dimensionalen Vektorraumes mit Skalarprodukt $\langle \cdot, \cdot \rangle$. Sei $f : U \rightarrow \mathbb{R}$ eine in $a \in U$ differenzierbare Abbildung. Dann heißt der Vektor $\text{grad } f(a)$ mit

$$\langle \text{grad } f(a), v \rangle = Df(a)(v) \tag{15.3}$$

für alle $v \in X$ der *Gradient* von f in $a \in U$. Ist f auf U differenzierbar, so nennt man die Abbildung $\text{grad } f : U \rightarrow X$ auch das *Gradientenvektorfeld* von f .

Bemerkungen.

- (a) Es folgt aus dem Riesz'schen Darstellungssatz 13.3.10, dass der Vektor $\text{grad } f(a)$ durch die Beziehung (15.3) eindeutig ist.
- (b) Sind $U \subset \mathbb{R}^n$, $\langle \cdot, \cdot \rangle$ das Standardskalarprodukt und (e_1, \dots, e_n) die kanonische Basis, so gilt:

$$\text{grad } f(a) = \sum_{j=1}^n \langle \text{grad } f(a), e_j \rangle e_j = \left(\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_n}(a) \right).$$

- (c) Ist $\|v\| = \sqrt{\langle v, v \rangle} = 1$, so misst

$$\langle \text{grad } f(a), v \rangle = Df(a)(v)$$

die Änderung von f in Richtung v . Ist $\text{grad } f(a) \neq 0$, so folgt

$$\langle \text{grad } f(a), v \rangle = \|\text{grad } f(a)\| \|v\| \cos \alpha = \|\text{grad } f(a)\| \cos \alpha,$$

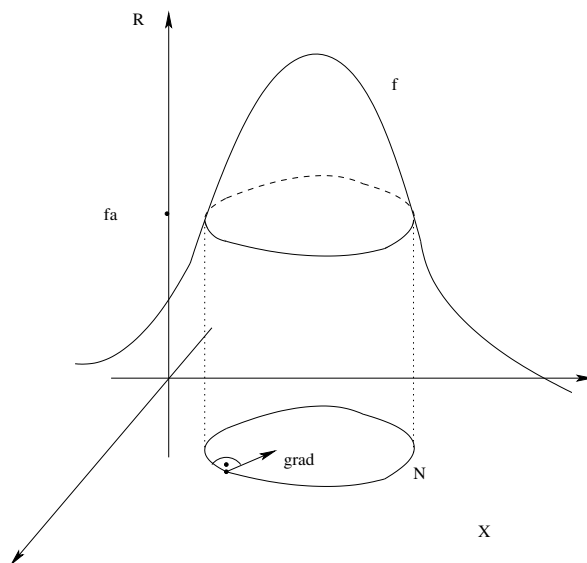


Abbildung 15.3:

wobei α den Winkel zwischen $\text{grad } f(a)$ und v bezeichnet (siehe 13.2.5 für die Definition des Winkels). Ist $\alpha = 0$ und somit $v = \frac{\text{grad } f(a)}{\|\text{grad } f(a)\|}$, so ist $\langle \text{grad } f(a), v \rangle$ maximal, d.h. die Funktion steigt in Richtung des Gradienten am stärksten. Ist $\alpha = \pi$ und somit $v = -\frac{\text{grad } f(a)}{\|\text{grad } f(a)\|}$, so ist $\langle \text{grad } f(a), v \rangle$ minimal, d.h. in Richtung $-\text{grad } f(a)$ fällt die Funktion am stärksten. Senkrecht zu $\text{grad } f(a)$ ändert sich die Funktion nicht. Diese Richtung ist tangential zu der Niveaumenge von f

$$N_{f(a)} = \{x \in X \mid f(x) = f(a)\}$$

im Punkte a . Wir werden auf diesen Sachverhalt im Rahmen der Theorie der impliziten Funktionen zurückkommen.

Definition 15.1.9. Ein *Vektorfeld* auf einem endlich-dimensionalen Vektorraum X ist eine Abbildung $v : U \rightarrow X$, wobei $U \subset X$. Ist $U \subset X$ offen, und existiert eine Funktion $f : U \rightarrow \mathbb{R}$ mit

$$\text{grad } f(x) = v(x),$$

so heißt f auch *Stammfunktion* oder *Potential* von v .

Bemerkung. Wie wir bald sehen werden, besitzt nicht jedes Vektorfeld ein Potential. Eine wichtige Größe ist die Divergenz eines Vektorfeldes.

Definition 15.1.10. Seien X ein endlich dimensionaler Vektorraum, $v : U \rightarrow X$ ein differenzierbares Vektorfeld auf der offenen Teilmenge U von X . Dann heißt

$$\text{div } v(a) = \text{Spur } Dv(a)$$

die *Divergenz* von v im Punkte $a \in U$. Ist $X = \mathbb{R}^n$ und $v(x) = (v_1(x), \dots, v_n(x))$, so gilt:

$$\text{div } v(a) = \sum_{j=1}^n \frac{\partial v_j}{\partial x_j}(a).$$

15.2 Rechenregeln für differenzierbare Abbildungen

Satz 15.2.1. *Es seien X, Y normierte Vektorräume und $U \subset X$ offen. Es seien $f, g : U \rightarrow Y$ in $a \in U$ differenzierbare Abbildungen und $c \in \mathbb{R}$. Dann sind $f + g : U \rightarrow Y$ und $c \cdot f : U \rightarrow Y$ in a differenzierbar und es gilt:*

$$D(f + g)(a) = Df(a) + Dg(a) \quad \text{und} \quad D(cf)(a) = cDf(a).$$

Beweis. Da f, g in $a \in U$ differenzierbar sind, folgt für alle $h \in X$ mit $a + h \in U$:

$$f(a + h) = f(a) + Df(a)(h) + \varphi_1(h)\|h\| \quad \text{und} \quad g(a + h) = g(a) + Dg(a)(h) + \varphi_2(h)\|h\|,$$

mit $Df(a), Dg(a) \in \mathcal{L}(X, Y)$. Außerdem sind φ_1, φ_2 in $0 \in X$ stetige Abbildungen mit $\varphi_1(0) = \varphi_2(0) = 0$. Dann erhalten wir durch Addition der beiden Gleichungen:

$$\begin{aligned} (f + g)(a + h) &= f(a + h) + g(a + h) \\ &= f(a) + g(a) + Df(a)(h) + Dg(a)(h) + (\varphi_1(h) + \varphi_2(h))\|h\| \\ &= (f + g)(a) + (Df(a) + Dg(a))(h) + \varphi(h)\|h\|. \end{aligned}$$

Dabei ist $\varphi := \varphi_1 + \varphi_2$ in 0 stetig mit $\varphi(0) = 0$. Damit ist $f + g$ in a differenzierbar mit

$$D(f + g)(a) = Df(a) + Dg(a).$$

Genauso folgt: $D(cf)(a) = cDf(a)$. □

Nun wollen wir die Produktregel für reellwertige Funktionen beweisen.

Satz 15.2.2 (Produktregel).

Seien X ein normierter Vektorraum, $U \subset X$ offen und $f, g : U \rightarrow \mathbb{R}$ in $a \in U$ differenzierbar. Dann ist $f \cdot g$ in a differenzierbar und es gilt:

$$D(f \cdot g)(a) = f(a) \cdot Dg(a) + g(a) \cdot Df(a).$$

Beweis. Zu zeigen ist

$$\lim_{h \rightarrow 0} \frac{|(f \cdot g)(a + h) - (f \cdot g)(a) - f(a) \cdot Dg(a)h - g(a) \cdot Df(a)(h)|}{\|h\|} = 0.$$

Da

$$f(a + h) = f(a) + Df(a)h + \varphi_1(h)\|h\| \quad \text{und} \quad g(a + h) = g(a) + Dg(a)h + \varphi_2(h)\|h\|$$

mit $\lim_{h \rightarrow 0} \varphi_1(h) = \lim_{h \rightarrow 0} \varphi_2(h) = 0$, erhalten wir durch Multiplikation der beiden Gleichungen:

$$\begin{aligned} f(a + h) \cdot g(a + h) &= f(a) \cdot g(a) + f(a) \cdot Dg(a)h + g(a)Df(a)h + \\ &\quad (Df(a)(h))(Dg(a)(h)) + \varphi_1(h)\|h\|(g(a) + Dg(a)h) \\ &\quad + \varphi_2(h)\|h\|(f(a) + Df(a)h) + \varphi_1(h)\varphi_2(h)\|h\|^2. \end{aligned}$$

Damit gilt:

$$\begin{aligned} &|(f \cdot g)(a + h) - (f \cdot g)(a) - f(a) \cdot Dg(a)h - g(a)Df(a)h| = \\ &|Df(a)h \cdot Dg(a)h + \varphi_1(h)\|h\|(g(a) + Dg(a)h) + \varphi_2(h)\|h\|(f(a) + Df(a)h) + \varphi_1(h)\varphi_2(h)\|h\|^2| \\ &\leq |Df(a)h||Dg(a)h| + |\varphi_1(h)| \|h\|(|g(a)| + |Dg(a)h|) + \\ &\quad |\varphi_2(h)| \|h\| (|f(a)| + |Df(a)h|) + |\varphi_1(h)\varphi_2(h)|\|h\|^2 \\ &\leq c_1 c_2 \|h\|^2 + |\varphi_1(h)| \|h\|(|g(a)| + c_2 \|h\|) + |\varphi_2(h)| \|h\|(f(a) + c_1 \|h\|) + |\varphi_1(h)\varphi_2(h)|\|h\|^2, \end{aligned}$$

denn wegen Satz 14.2.12 gilt: $|Df(a)h| \leq c_1\|h\|$ und $|Dg(a)(h)| \leq c_2\|h\|$ für Konstanten $c_1, c_2 > 0$. Daraus folgt die Behauptung. \square

Nun wollen wir Ableitungen von verketteten Funktionen berechnen.

Satz 15.2.3 (Kettenregel).

Es seien X, Y, Z normierte Vektorräume und $U \subset X, V \subset Y$ offen. Es sei $a \in U$, und $f : U \rightarrow V$ bzw. $g : V \rightarrow Z$ seien in a bzw. in $f(a)$ differenzierbare Abbildungen. Dann ist auch $g \circ f : U \rightarrow Z$ in $a \in U$ differenzierbar und es gilt:

$$D(g \circ f)(a) = Dg(f(a)) \circ Df(a).$$

Bemerkung. Sind die linearen Abbildungen

$$Df(a) \in \mathcal{L}(X, Y) \quad \text{und} \quad Dg(f(a)) \in \mathcal{L}(Y, Z)$$

stetig, so auch die Hintereinanderschaltung

$$Dg(f(a)) \circ Df(a) \in \mathcal{L}(X, Z).$$

Beweis. Sei $k(h) = f(a + h) - f(a)$. Da g in $f(a)$ differenzierbar, gilt:

$$g \circ f(a + h) = g(f(a) + k(h)) = g \circ f(a) + Dg(f(a))k(h) + \varphi_2(k(h))\|k(h)\|,$$

wobei $\lim_{h \rightarrow 0} \varphi_2(h) = 0$ gilt.

Desweiteren folgt aus der Differenzierbarkeit von f in a :

$$k(h) = f(a + h) - f(a) = Df(a)h + \varphi_1(h)\|h\|.$$

Setzen wir dies in die obige Beziehung ein, so erhalten wir:

$$\begin{aligned} g \circ f(a + h) &= g \circ f(a) + Dg(f(a)) \circ Df(a)(h) + \|h\|Dg(f(a))\varphi_1(h) \\ &\quad + \varphi_2(Df(a)h + \varphi_1(h)\|h\|)\|Df(a)h + \varphi_1(h)\|h\|. \end{aligned}$$

Da $Dg(f(a))$ und $Df(a)$ stetig sind, existieren nach Satz 14.2.12 Konstanten $c_1, c_2 > 0$ mit

$$\|Dg(f(a))(h)\| \leq c_2\|h\| \quad \text{und} \quad \|Df(a)h\| \leq c_1\|h\|.$$

Damit ergibt sich folgende Abschätzung:

$$\begin{aligned} &\|g \circ f(a + h) - g \circ f(a) - Dg(f(a)) \circ Df(a)h\| \\ &\quad \leq \|h\| \cdot \|Dg(f(a))\varphi_1(h)\| \\ &\quad + \|\varphi_2(Df(a)h + \varphi_1(h)\|h\|)\| \cdot (\|Df(a)(h)\| + \|\varphi_1(h)\| \|h\|) \\ &\leq c_2\|h\|\|\varphi_1(h)\| + \|\varphi_2(Df(a)h + \varphi_1(h)\|h\|)\| \cdot (c_1 + \|\varphi_1(h)\|)\|h\|. \end{aligned}$$

Daraus folgt:

$$\begin{aligned} &\lim_{h \rightarrow 0} \frac{\|g \circ f(a + h) - g \circ f(a) - Dg(f(a))Df(a)h\|}{\|h\|} \leq \\ &\lim_{h \rightarrow 0} (c_2\|\varphi_1(h)\| + (c_1 + \|\varphi_1(h)\|)\|\varphi_2(Df(a)h + \varphi_1(h)\|h\|)\|) = 0. \end{aligned}$$

\square

Beispiele.

- (a) Sei $A \in \mathcal{L}(X, Y)$, eine stetige lineare Abbildung, $b \in Y$ und $g : X \rightarrow Y$ die *affin lineare Abbildung* mit $g(x) = A(x) + b$. Man betrachte $F(x) = f(g(x)) = f(Ax + b)$. Dann gilt:

$$DF(x) = Df(Ax + b) \circ A.$$

- (b) Sei $c : (a, b) \rightarrow Y$ eine differenzierbare Kurve, so ist $f \circ c : (a, b) \rightarrow Z$ eine differenzierbare Kurve und es gilt:

$$\widehat{f \circ c}(t) = D(f \circ c)(t)(1) = Df(c(t)) \circ Dc(t)(1) = Df(c(t))(\dot{c}(t)).$$

15.3 Mittelwertsatz

Der Mittelwertsatz der Differentialrechnung 4.4.5 lässt sich mit Hilfe der Kettenregel unmittelbar auf reellwertige Funktionen übertragen.

Satz 15.3.1 (*Mittelwertsatz für reellwertige Funktionen*).

Seien X ein normierter Vektorraum und $U \subset X$ offen. Sei $f : U \rightarrow \mathbb{R}$ eine differenzierbare Funktion und $x, y \in U$, so dass die Verbindungsgerade $c(t) = x + t(y - x)$ ganz in U verläuft. Dann gilt:

$$f(y) - f(x) = Df(x + \delta(y - x))(y - x)$$

für ein $\delta \in (0, 1)$.

Beweis. Man betrachte die Funktion $f \circ c : [0, 1] \rightarrow \mathbb{R}$. Wegen des Mittelwertsatzes 4.4.5 existiert ein $\delta \in (0, 1)$ mit

$$f(y) - f(x) = f(c(1)) - f(c(0)) = \widehat{(f \circ c)}(\delta) \cdot (1 - 0) = Df(c(\delta))(\dot{c}(\delta)) = Df(x + \delta(y - x))(y - x),$$

denn $\dot{c}(t) = y - x$ für alle $t \in (0, 1)$:

$$\dot{c}(t) = \lim_{h \rightarrow 0} \frac{c(t+h) - c(t)}{h} = \lim_{h \rightarrow 0} \frac{(t+h)(y-x) - t(y-x)}{h} = (y-x).$$

□

Wir wollen nun einen Mittelwertsatz für vektorwertige Abbildungen beweisen. Dazu werden wir das in Kapitel 5 eingeführte Integral auf Kurven in Banachräumen ausdehnen. Wir werden uns hierbei relativ kurz fassen, da die dort verwendeten Definitionen und Sätze sich ohne große Probleme auf den allgemeinen Fall übertragen lassen.

Definition 15.3.2. Sei $(X, \| \cdot \|)$ ein Banachraum und $I = [a, b] \subset \mathbb{R}$ ein abgeschlossenes Intervall. Eine Abbildung $\varphi : [a, b] \rightarrow X$ heißt *Treppenfunktion*, falls eine Unterteilung

$$a = x_0 < x_1 < \dots < x_n = b$$

existiert, so dass φ auf jedem offenen Teilintervall (x_{k-1}, x_k) konstant ist. Mit $T([a, b], X)$ bezeichnen wir die Menge der Treppenfunktionen auf X .

Bemerkung. $T([a, b], X)$ ist ein Untervektorraum der beschränkten Funktionen $B([a, b], X)$ von I nach X . (Beweis wie im Falle $X = \mathbb{R}$, siehe Satz 5.1.2). Wegen Satz 14.2.6 und der anschließenden Bemerkung ist $B(I, X)$ bezüglich der Supremumsnorm

$$\|f\|_{[a,b]} = \{\sup \|f(t)\| \mid t \in [a, b]\}$$

ein vollständiger Vektorraum.

Genau wie für $X = \mathbb{R}$ lässt sich das Integral definieren.

Definition 15.3.3. Sei $\varphi \in T([a, b], X)$ und $a = x_0 < x_1 < \dots < x_n = b$ eine zu φ gehörige Unterteilung von $I = [a, b]$. Dann heißt

$$\int \varphi := \sum_{k=1}^n c_k(x_k - x_{k-1}) \in X$$

mit $c_k = \varphi(x)$ für $x \in (x_{k-1}, x_k)$ das *Integral* von φ .

Bemerkungen.

(a) Diese Definition ist unabhängig von der Unterteilung.

(b) Ist $\varphi \in T([a, b], X)$, so schreiben wir auch $\int_a^b \varphi(t) dt$ statt $\int \varphi$.

(c) Ist $X = \mathbb{R}^n$ und $\varphi(t) = (\varphi_1(t), \dots, \varphi_n(t))$, so ist

$$\int \varphi = \left(\int \varphi_1, \dots, \int \varphi_n \right).$$

Satz 15.3.4. Sei $(T([a, b], X), \|\cdot\|_{[a,b]})$ der bezüglich der Supremumsnorm auf $T([a, b], X)$ normierte Vektorraum der Treppenfunktionen. Dann gilt für alle $\varphi \in T([a, b], X)$:

$$\left\| \int_a^b \varphi(t) dt \right\| \leq \int_a^b \|\varphi(t)\| dt \leq \|\varphi\|_{[a,b]}(b-a).$$

Außerdem ist

$$\int_a^b : T([a, b], X) \rightarrow X$$

mit $\varphi \rightarrow \int_a^b \varphi(t) dt$ eine stetige lineare Abbildung.

Beweis. Sei nun $\varphi \in T([a, b], X)$ eine Treppenfunktion mit $\varphi(x) = c_k$ für $x \in (x_{k-1}, x_k)$ und $a = x_0 < \dots < x_n = b$. Dann gilt:

$$\left\| \int_a^b \varphi(t) dt \right\| = \left\| \sum_{k=1}^n c_k(x_k - x_{k-1}) \right\| \leq \sum_{k=1}^n \|c_k\|(x_k - x_{k-1}) = \int_a^b \|\varphi(t)\| dt \leq \|\varphi\|_{[a,b]}(b-a).$$

Die Linearität zeigt man wie im Falle $X = \mathbb{R}$ (siehe Satz 5.1.4). Die Stetigkeit folgt aus der Abschätzung

$$\left\| \int_a^b \varphi(t) dt \right\| \leq \|\varphi\|_{[a,b]}(b-a).$$

□

Nun wollen wir analog zur Definition 5.1.7 die Regelfunktionen einführen.

Definition 15.3.5. Eine beschränkte Funktion $f : [a, b] \rightarrow X$ heißt *Regelfunktion*, wenn für jedes $\epsilon > 0$ eine Treppenfunktion $\varphi \in T([a, b], X)$ existiert mit

$$\|f - \varphi\|_{[a,b]} \leq \epsilon.$$

Die Menge der Regelfunktionen auf $[a, b]$ bezeichnen wir mit $R([a, b], X)$.

Bemerkung. Nach Definition des Abschlusses, gilt also

$$R([a, b], X) = \overline{T([a, b], X)},$$

denn in jedem ϵ -Ball um eine Regelfunktion liegt eine Treppenfunktion.

Analog zur Definition 5.1.9 lässt sich nun das Integral auf $R([a, b], X) = \overline{T([a, b], X)}$ ausdehnen.

Hier brauchen wir, dass X vollständig ist.

Definition 15.3.6. Sei $(X, \|\cdot\|)$ ein Banachraum und $f \in R([a, b], X)$. Sei $\varphi_n \in T([a, b], X)$ eine Folge von Treppenfunktionen mit $\|\varphi_n - f\|_{[a,b]} \rightarrow 0$. Dann definiere man:

$$\int f := \lim_{n \rightarrow \infty} \int \varphi_n \in X.$$

Bemerkungen.

(a) Der Limes existiert, denn wegen Satz 15.3.4 gilt

$$\left\| \int_a^b \varphi_n(t) dt - \int_a^b \varphi_m(t) dt \right\| = \left\| \int_a^b (\varphi_n - \varphi_m)(t) dt \right\| \leq \|\varphi_n - \varphi_m\|_{[a,b]} (b - a).$$

Da $\varphi_n \in T([a, b], X)$ eine Cauchyfolge ist, ist auch $\int \varphi_n \in X$ eine Cauchyfolge und der Grenzwert existiert wegen der Vollständigkeit von X .

(b) Der Grenzwert hängt nicht von der Wahl der Folge φ_n ab, denn ist ψ_n eine weitere Folge mit $\psi_n \rightarrow f$, so gilt:

$$\left\| \int_a^b \varphi_n(t) dt - \int_a^b \psi_n(t) dt \right\| \leq \|\varphi_n - \psi_n\|_{[a,b]} (b - a) \rightarrow 0.$$

(c) Die stetigen Funktionen $C^0([a, b], X)$ sind Regelfunktionen. Der Beweis ist analog zum Beweis des Satzes 5.1.8.

(d) Die Abschätzung in Lemma 15.3.4 überträgt sich auf Regelfunktionen, d.h. es gilt

$$\left\| \int_a^b f(t) dt \right\| \leq \int_a^b \|f(t)\| dt$$

für alle $f \in R([a, b], X)$.

(e) Ist $b < a$, so setze man $\int_a^b f(t)dt = -\int_b^a f(t)dt$ (siehe Abschnitt 5.2).

Satz 15.3.7 (verallgemeinerter Hauptsatz der Differential- und Integralrechnung). Sei X ein Banachraum und $f \in C^0([a, b], X)$ eine auf $[a, b]$ stetige Funktion. Gegeben sei

$$F(x) = \int_a^x f(t)dt$$

für $x \in [a, b]$. Dann ist $F : [a, b] \rightarrow X$ differenzierbar auf (a, b) und es gilt:

$$\dot{F}(x) = f(x), \quad x \in (a, b).$$

Beweis. Zu zeigen ist:

$$\lim_{h \rightarrow 0} \frac{F(x+h) - F(x)}{h} = f(x). \quad (15.4)$$

Ist $h > 0$, so gilt:

$$\begin{aligned} \left\| \frac{F(x+h) - F(x)}{h} - f(x) \right\| &= \left\| \frac{\int_a^{x+h} f(t)dt - \int_a^x f(t)dt - \int_x^{x+h} f(x)dt}{h} \right\| = \left\| \frac{\int_x^{x+h} f(t) - f(x)dt}{h} \right\| \\ &\leq \frac{|h|}{|h|} \sup\{\|f(t) - f(x)\| \mid t \in [x, x+h]\} \\ &= \|f(t) - f(x)\|_{[x, x+h]} \rightarrow 0, \end{aligned}$$

für $h \rightarrow 0$ und $h > 0$. Im Fall $h < 0$ ersetzt man das Intervall $[x, x+h]$ durch $[x+h, x]$ und erhält eine analoge Abschätzung. \square

Bemerkung. Der Grenzwert (15.4) existiert auch für die Randpunkte, allerdings haben wir die Differenzierbarkeit auf höherdimensionalen Vektorräumen nur für die offenen Mengen definiert.

Genau wie für Funktionen von \mathbb{R} nach \mathbb{R} führen wir den Begriff der Stammfunktion ein.

Definition 15.3.8. Sei $D \subset \mathbb{R}$ offen, X ein Banachraum und $f : D \rightarrow X$. Dann heißt $F : D \rightarrow X$ *Stammfunktion* von f , falls $\dot{F}(t) = f(t)$ für alle $t \in D$.

Nach dem Hauptsatz existieren Stammfunktionen von stetigen Funktionen auf Intervallen. Um zu zeigen, dass Stammfunktionen eindeutig bis auf eine Konstante bestimmt sind, benötigen wir das folgende Lemma.

Lemma 15.3.9 (Hahn-Banach).

Sei X ein normierter Vektorraum und $x \in X$, $x \neq 0$. Dann existiert eine stetige lineare Abbildung $\lambda \in \mathcal{L}(X, \mathbb{R})$ mit $\lambda(x) \neq 0$.

Bemerkung. Wir werden dieses Lemma nicht beweisen. Es ist trivial für endlich dimensionale Vektorräume: Ist (b_1, \dots, b_n) eine Basis von X und $x = x_1 b_1 + \dots + x_n b_n \neq 0$, so ist $x_i \neq 0$ für ein $i \in \{1, \dots, n\}$. Dann ist $\lambda(y) = y_i$ für $y = y_1 b_1 + \dots + y_n b_n \in X$ eine stetige lineare Abbildung mit $\lambda(x) \neq 0$.

Korollar 15.3.10. *Ist $f : I \rightarrow X$ differenzierbar und $\dot{f}(t) = 0$ für alle $t \in I$, so ist f konstant.*

Beweis. Sei $t_0 \in I$ fest gewählt. Ist f nicht konstant, so existiert ein $t \in I$ mit $f(t) \neq f(t_0)$. Wegen des obigen Lemmas existiert ein $\lambda \in \mathcal{L}(X, \mathbb{R})$ mit

$$0 \neq \lambda(f(t) - f(t_0)) = \lambda(f(t)) - \lambda(f(t_0)).$$

Da jede stetige lineare Abbildung differenzierbar ist, ist $\lambda \circ f : I \rightarrow \mathbb{R}$ differenzierbar und aus der Kettenregel folgt:

$$(\lambda \circ f)'(t) = D(\lambda \circ f)(t)(1) = D\lambda \circ Df(t)(1) = \lambda(\dot{f}(t)) = 0.$$

Dann folgt aus dem Mittelwertsatz der Differentialrechnung 4.4.5: $\lambda \circ f : I \rightarrow \mathbb{R}$ ist konstant. Insbesondere ist $\lambda \circ f(t) = \lambda \circ f(t_0)$ im Widerspruch zur Wahl von λ . \square

Korollar 15.3.11. *Seien I ein offenes Intervall und $F, G : I \rightarrow X$ Stammfunktionen von $f : I \rightarrow X$. Dann ist $F(t) - G(t) = c$, wobei c eine Konstante in X ist.*

Beweis. Da $\widehat{F - G} = \dot{F} - \dot{G} = 0$, folgt dies aus Korollar 15.3.10. \square

Korollar 15.3.12 (*Mittelwertsatz der Integralrechnung für Kurven*).

Sei X ein Banachraum und $f \in R([a, b], X)$. Dann gilt für jede Stammfunktion F von f :

$$F(b) - F(a) = \int_a^b f(t) dt.$$

Beweis. Da $F(x)$ nach Annahme und $\int_a^x f(t) dt$ nach Satz 15.3.7 Stammfunktionen von f

sind, existiert wegen des Korollars 15.3.11 eine Konstante $c \in X$ mit $F(x) - \int_a^x f(t) dt = c$.

Setzen wir für $x = a$, so erhalten wir

$$c = F(a) \quad \text{und somit} \quad F(b) = F(a) + \int_a^b f(t) dt.$$

\square

Nun können wir den Mittelwertsatz wie folgt für stetig differenzierbare Abbildungen formulieren.

Definition 15.3.13. Seien X, Y normierte Vektorräume und $U \subset X$ offen. Eine differenzierbare Abbildung $f : U \rightarrow Y$ heißt *stetig differenzierbar*, falls f differenzierbar ist und das Differential $Df : U \rightarrow \mathcal{L}(X, Y)$ stetig ist, wobei wir wie üblich auf dem Vektorraum $\mathcal{L}(X, Y)$ die Operatornorm wählen.

Satz 15.3.14 (*Mittelwertsatz der Integralrechnung*).

Es seien X, Y Banachräume und $U \subset X$ offen. Sei $f : U \rightarrow Y$ eine stetig differenzierbare Abbildung und $x, y \in U$, so dass die Verbindungsgerade $c(t) = x + t(y - x)$ ganz in U verläuft. Dann gilt:

$$f(y) - f(x) = \int_0^1 Df(x + t(y - x))(y - x) dt.$$

Bemerkung. Sei $A : [a, b] \rightarrow \mathcal{L}(X, Y)$ stetig mit $t \mapsto A(t)$. Dann ist für jedes $h \in X$ die Abbildung $[a, b] \rightarrow Y$ mit $t \mapsto A(t)h$ stetig. Die Abbildung $\left(\int_0^1 A(t)dt\right)$ mit

$$\left(\int_0^1 A(t)dt\right)(h) := \int_0^1 A(t)h dt$$

ist linear und stetig und somit ein Element in $\mathcal{L}(X, Y)$. Die Stetigkeit folgt aus Satz 15.3.4, denn

$$\left\| \left(\int_0^1 A(t)dt\right)h \right\| \leq \int_0^1 \|A(t)h\| dt \leq \int_0^1 \|A(t)\| \cdot \|h\| dt \leq \|h\| \int_0^1 \|A(t)\| dt.$$

Mit Satz 15.3.14 folgt:

$$f(y) - f(x) = \left(\int_0^1 Df(x + t(y-x))dt\right)(y-x).$$

Beweis. Die Abbildung $f \circ c : [0, 1] \rightarrow Y$ mit $t \mapsto (f \circ c)(t)$ ist stetig differenzierbar, denn c ist differenzierbar mit $\dot{c}(t) = y - x$ (siehe Beweis von Satz 15.3.1). Aus der Kettenregel folgt:

$$\widehat{f \circ c}(t) = Df(c(t))(\dot{c}(t)) = Df(x + t(y-x))(y-x).$$

Aus dem Mittelwertsatz der Integralrechnung für Kurven folgt dann:

$$f(y) - f(x) = \int_0^1 \widehat{f \circ c}(t) dt = \int_0^1 Df(x + t(y-x))(y-x) dt.$$

□

Korollar 15.3.15. Sei $f : U \rightarrow X$ stetig differenzierbar und $\|Df(a)\| \leq c$ für alle $a \in U$. Dann gilt:

$$\|f(y) - f(x)\| \leq c \cdot \|y - x\|,$$

falls die Gerade zwischen x und y in ganz U verläuft.

Beweis. Aus dem Mittelwertsatz 15.3.14 und aus der Bemerkung (d) nach der Definition 15.3.6 folgt

$$\|f(y) - f(x)\| = \left\| \int_0^1 Df(x + t(y-x))(y-x) dt \right\| \leq \int_0^1 \|Df(x + t(y-x))(y-x)\| dt \leq c \|y - x\|.$$

□

Definition 15.3.16. Sei X ein Vektorraum und $A \subset X$. Dann heißt A *konvex*, falls für alle $x, y \in A$ die Gerade die x mit y verbindet, ganz in A verläuft.

Bemerkungen.

- (a) Die offenen Bälle in einem normierten Vektorraum sind Beispiele für konvexe Mengen
- (b) Aus Korollar 15.3.15 folgt: Ist $U \subset X$ konvex und offen und $\|Df(x)\| \leq c$, so ist f Lipschitz-stetig mit Lipschitzkonstante c . Ist $Df(x) = 0$, so ist f konstant. Dies gilt auch für eine viel größere Klasse von Mengen, nämlich die zusammenhängenden Mengen.

Definition 15.3.17. Sei (M, d) ein metrischer Raum. M heißt *zusammenhängend*, falls sich M nicht als nichttriviale disjunkte Vereinigung zweier offener Mengen schreiben lässt, d.h. sind $U, V \subset M$ offen in M mit $U \cap V = \emptyset$ und $U \cup V = M$, so ist $U = \emptyset$ oder $U = M$.

Zusammenhängende metrische Räume lassen sich auch wie folgt charakterisieren:

Satz 15.3.18. Sei (M, d) ein metrischer Raum. Dann sind äquivalent:

- (a) M ist zusammenhängend.
- (b) Ist $U \subset M$ offen und abgeschlossen, so folgt $U = M$ oder $U = \emptyset$.

Beweis. Sei M zusammenhängend und U offen und abgeschlossen. Da U abgeschlossen ist, ist $M \setminus U = V$ offen. Da U offen ist, ist dann $M = U \cup V$ die disjunkte Vereinigung der offenen Mengen U und V . Da M zusammenhängend ist, folgt $U = M$ oder $U = \emptyset$. Gilt nun (b) und sind $U, V \subset M$ offen mit $U \cup V = M$ und $U \cap V = \emptyset$, so ist $U = M \setminus V$ abgeschlossen. Also folgt $U = M$ oder $U = \emptyset$. \square

Beispiel. Sei $M = I \subset \mathbb{R}$ ein Intervall, so ist I bezüglich $d(x, y) := |x - y|$ ein metrischer Raum. Dann ist I zusammenhängend, denn ist $U \subset I$ offen und abgeschlossen in I , so ist wie wir sehen werden, $U = I$ oder $U = \emptyset$. Denn ist $x_0 \in U$ und $x \in I$ mit $x > x_0$, so betrachte man

$$\sup\{t \mid t \in [x_0, x] \cap U\} =: b \in [x_0, x] \subset I.$$

Nach Definition des Supremum ist b Berührungspunkt von U , denn in jedem ϵ -Ball um b liegt ein Element aus U . Da U abgeschlossen ist, ist somit $b \in U$. Dann ist aber $b = x$, denn ist $b < x$, so existiert, da U offen ist, ein $\epsilon > 0$ mit $B(b, \epsilon) \subset U$ im Widerspruch zur Definition von b . Insbesondere ist $b = x \in U$. Ist $x < x_0$, so betrachte man

$$\inf\{t \mid t \in [x, x_0] \cap U\} =: a \in [x, x_0] \subset I$$

und wir erhalten wie oben: $x = a \in U$. Außer den Intervallen existieren keine weiteren zusammenhängenden Mengen in \mathbb{R} .

Man kann auch einen Zusammenhangsbegriff mit Hilfe stetiger Wege definieren.

Definition 15.3.19. Sei (M, d) ein metrischer Raum. M heißt *wegzusammenhängend*, falls für $x, y \in M$ eine stetige Abbildung $c : [0, 1] \rightarrow M$ (*stetige Kurve* oder *stetiger Weg*) existiert mit $c(0) = x, c(1) = y$.

Satz 15.3.20. Ist M wegzusammenhängend, so ist M auch zusammenhängend.

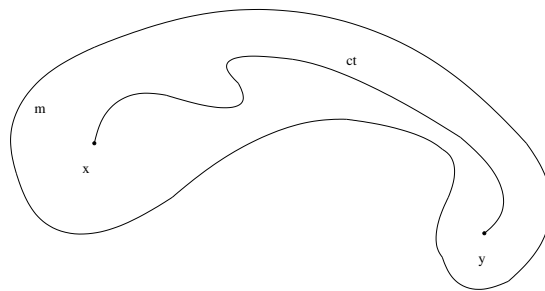


Abbildung 15.4:

Beweis. Sei M wegzusammenhängend. Würden offene Mengen $U, V \subset M$ existieren, mit $U \cap V = \emptyset$ und $U \cup V = M$, sowie $U \neq \emptyset$ und $V \neq \emptyset$, so gäbe es zu $x \in U$ und $y \in V$ einen stetigen Weg mit $c : [0, 1] \rightarrow M$ mit $c(0) = x, c(1) = y$. Wegen der Stetigkeit von c , sind dann die Mengen $c^{-1}(U)$ und $c^{-1}(V)$ offen in $[0, 1]$. Da $U \cap V = \emptyset$ gilt auch $c^{-1}(U) \cap c^{-1}(V) = \emptyset$. Da $U \cup V = M$ ist $[0, 1] = c^{-1}(M) = c^{-1}(U) \cup c^{-1}(V)$. Dies steht aber im Widerspruch dazu, dass $[0, 1]$ zusammenhängend ist. \square

Satz 15.3.21. *Es sei $U \subset X$ eine offene und zusammenhängende Teilmenge des Banachraumes X . Ist Y ein Banachraum und $f : U \rightarrow Y$ stetig differenzierbar mit $Df(x) = 0$ für alle $x \in U$, so ist f auf U konstant.*

Beweis. Die Menge $U \subset X$ mit $d(x, y) = \|x - y\|$ für $x, y \in U$ ist ein metrischer Raum. Zu $x_0 \in U$ betrachte man die Menge $V = \{x \in U \mid f(x) = f(x_0)\}$. V ist abgeschlossen, denn ist (x_n) eine konvergente Folge mit $x_n \in V$, so ist auch der Grenzwert $x \in V$, denn aus der Stetigkeit von f und der Definition von V folgt:

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

V ist offen, denn ist $x \in V$, so existiert (da U offen in X ist) ein $\epsilon > 0$ mit $B(x, \epsilon) \subset U$. Da $B(x, \epsilon)$ konvex ist, gilt wegen Bemerkung (b) nach Definition 15.3.16 $f(y) = f(x) = f(x_0)$ für alle $y \in B(x, \epsilon)$, d.h. $B(x, \epsilon) \subset V$. Damit ist x innerer Punkt und somit ist V offen.

Die einzigen gleichzeitig offenen und abgeschlossenen Mengen sind aber U und \emptyset . Da V nicht leer ist, gilt: $V = U$. \square

Bemerkung. Der Satz folgt auch, wenn f nur als differenzierbar vorausgesetzt wird.

15.4 Multilineare Abbildungen

Die Definition von Ableitungen höherer Ordnung macht es erforderlich, multilineare Abbildungen zu betrachten.

Definition 15.4.1. Es seien X_1, \dots, X_n, Y Vektorräume über einem Körper K . Eine Abbildung $\alpha : X_1 \times \dots \times X_n \rightarrow Y$ heißt *multilinear*, falls α in jeder Variablen linear ist, d.h. für jedes $i \in \{1, \dots, n\}$ und $(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$ ist die Abbildung

$$x \rightarrow \alpha(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_n)$$

eine lineare Abbildung von X_i nach Y . Wir bezeichnen mit $L(X_1, \dots, X_n; Y)$ die Menge der multilinearen Abbildungen von $X_1 \times \dots \times X_n$ nach Y . Ist $X_1 = \dots = X_n = X$, so schreiben wir auch $L^n(X, Y)$ statt $L(X, \dots, X; Y)$.

Beispiele.

- (a) $\det : \mathbb{R}^n \times \dots \times \mathbb{R}^n \rightarrow \mathbb{R}$.
- (b) Ist X ein Vektorraum über \mathbb{R} , so ist jedes Skalarprodukt $\alpha : X \times X \rightarrow \mathbb{R}$ eine bilineare Abbildung.
- (c) Sind X, Y Vektorräume über $K = \mathbb{R}$ oder $K = \mathbb{C}$, so ist

$$\alpha : \mathcal{L}(X, Y) \times X \rightarrow Y \quad \text{mit} \quad (A, x) \mapsto A(x)$$

bilinear. Denn sind $A_1, A_2, A \in \mathcal{L}(X, Y)$, $x \in X$ und $\lambda \in K$, so gilt

$$\alpha(A_1 + A_2, x) = (A_1 + A_2)(x) = A_1(x) + A_2(x) = \alpha(A_1, x) + \alpha(A_2, x)$$

und

$$\alpha(\lambda A, x) = \lambda \alpha(A, x).$$

Außerdem folgt für $x, y \in X$

$$\alpha(A, x + y) = A(x + y) = A(x) + A(y) = \alpha(A, x) + \alpha(A, y)$$

und

$$\alpha(A, \lambda x) = \lambda \alpha(A, x).$$

Sind X_1, \dots, X_n normierte Vektorräume, so betrachte man die Maximumsnorm auf dem Produktraum $X_1 \times \dots \times X_n$, definiert durch

$$\|x\| = \{\max \|x_i\| \mid 1 \leq i \leq n\}$$

für alle $x = (x_1, \dots, x_n) \in X_1 \times \dots \times X_n$. Durch Verallgemeinerung der Charakterisierung der stetigen Abbildungen in Satz 14.2.12, erhalten wir die folgende Charakterisierung der stetigen multilinearen Abbildungen.

Satz 15.4.2. *Es seien X_1, \dots, X_n, Y normierte Vektorräume und $\alpha : X_1 \times \dots \times X_n \rightarrow Y$ eine multilineare Abbildung. Dann ist α genau dann stetig, falls eine Konstante $c \geq 0$ existiert, so dass*

$$\|\alpha(x_1, \dots, x_n)\| \leq c \|x_1\| \cdot \|x_2\| \cdot \dots \cdot \|x_n\|. \quad (15.5)$$

Bemerkung. Sind die Vektorräume X_1, \dots, X_n, Y endlich dimensional, so ist diese Bedingung (15.5) immer erfüllt. Zum Beweis wähle man Basen in X_1, \dots, X_n und nutze die Multilinearität von α .

Beweis. Sei $\alpha : X_1 \times \dots \times X_n \rightarrow Y$ stetig, so ist α insbesondere stetig in $0 = (0, \dots, 0) \in X_1 \times \dots \times X_n$. Daher existiert zu $\epsilon = 1$ ein $\delta > 0$, so dass $\|\alpha(x_1, \dots, x_n)\| \leq 1$, falls

$$\|(x_1, \dots, x_n)\| = \max\{\|x_i\| \mid i \in \{1, \dots, n\}\} \leq \delta.$$

Sei $x = (x_1, \dots, x_n)$, wobei $x_i \neq 0$ für alle $i \in \{1, \dots, n\}$. Dann gilt:

$$1 \geq \left\| \alpha \left(\delta \frac{x_1}{\|x_1\|}, \dots, \delta \frac{x_n}{\|x_n\|} \right) \right\|$$

und aus der Multilinearität von α folgt

$$\|\alpha(x_1, \dots, x_n)\| \leq \frac{1}{\delta^n} \|x_1\| \cdot \dots \cdot \|x_n\| = c \|x_1\| \cdot \dots \cdot \|x_n\|$$

mit $c = \frac{1}{\delta^n}$. Ist $x = (x_1, \dots, x_n)$ mit $x_i = 0$ für ein $i \in \{1, \dots, n\}$, so ist $\alpha(x_1, \dots, x_n) = 0$ und die Abschätzung ist trivialerweise erfüllt. Wir nehmen nun umgekehrt an, dass eine Konstante c mit obiger Eigenschaft existiert. Sei $a = (a_1, \dots, a_n) \in X_1 \times \dots \times X_n$. Um die Stetigkeit in a zu zeigen, betrachten wir die Teleskopsumme

$$\begin{aligned} & \alpha(x_1, \dots, x_n) - \alpha(a_1, \dots, a_n) = \alpha(x_1, \dots, x_n) - \alpha(a_1, x_2, \dots, x_n) \\ & + \alpha(a_1, x_2, \dots, x_n) - \alpha(a_1, a_2, x_2, \dots, x_n) + \dots + \alpha(a_1, a_2, \dots, x_n) - \alpha(a_1, \dots, a_n) \\ & = \alpha(x_1 - a_1, x_2, \dots, x_n) + \alpha(a_1, x_2 - a_2, x_3, \dots, x_n) + \dots + \alpha(a_1, \dots, a_{n-1}, x_n - a_n). \end{aligned}$$

Damit folgt:

$$\begin{aligned} & \|\alpha(x_1, \dots, x_n) - \alpha(a_1, \dots, a_n)\| \leq c \|x_1 - a_1\| \|x_2\| \cdot \dots \cdot \|x_n\| \\ & + c \|a_1\| \|x_2 - a_2\| \|x_3\| \cdot \dots \cdot \|x_n\| + \dots + c \|a_1\| \cdot \dots \cdot \|a_{n-1}\| \|x_n - a_n\|. \end{aligned}$$

Es sei $x = (x_1, \dots, x_n) \in B(a, 1)$, d.h. $\|(x_1, \dots, x_n) - (a_1, \dots, a_n)\| = \|x - a\| \leq 1$. Dann gilt

$$\max\{\|x_i\| \mid i \in \{1, \dots, n\}\} = \|x\| \leq \|x - a\| + \|a\| \leq 1 + \|a\| = b$$

und es folgt:

$$\|\alpha(x_1, \dots, x_n) - \alpha(a_1, \dots, a_n)\| \leq b^{n-1} c \left(\sum_{i=1}^n \|x_i - a_i\| \right) \leq n b^{n-1} c \|x - a\|.$$

Daher ist α Lipschitz-stetig und somit auch stetig in $a = (a_1, \dots, a_n)$. □

Definition 15.4.3. Sind X_1, \dots, X_n, Y normierte Vektorräume, so bezeichnen wir mit

$$\mathcal{L}(X_1, \dots, X_n; Y)$$

die Menge aller stetigen multilinearen Abbildungen $\alpha : X_1 \times \dots \times X_n \rightarrow Y$. Ist $X_1 = \dots = X_n = X$, so schreiben wir auch $\mathcal{L}^n(X, Y)$ statt $\mathcal{L}(X, \dots, X; Y)$.

Bemerkung. Genau wie die linearen Abbildungen, bilden auch die multilinearen Abbildungen einen Vektorraum. Die stetigen multilinearen Abbildungen bilden einen Untervektorraum.

Satz 15.4.4. Sei $\alpha : X_1 \times \dots \times X_n \rightarrow Y$ eine stetige multilineare Abbildung. Dann ist α in jedem Punkt $(a_1, \dots, a_n) \in X_1 \times \dots \times X_n$ differenzierbar. Ihre Ableitung

$$D\alpha(a_1, \dots, a_n) \in \mathcal{L}(X_1 \times \dots \times X_n, Y)$$

ist durch

$$D\alpha(a_1, \dots, a_n)(h_1, \dots, h_n) = \alpha(h_1, a_2, \dots, a_n) + \alpha(a_1, h_2, a_3, \dots, a_n) + \dots + \alpha(a_1, \dots, a_{n-1}, h_n).$$

gegeben.

Bemerkung. Für $n > 1$ ist natürlich $\mathcal{L}(X_1 \times \dots \times X_n, Y)$ verschieden von $\mathcal{L}^n(X_1, \dots, X_n; Y)$.

Beweis. Wir beweisen dies für $n = 2$.

$$\begin{aligned}\alpha(a_1 + h_1, a_2 + h_2) &= \alpha(a_1, a_2 + h_2) + \alpha(h_1, a_2 + h_2) \\ &= \alpha(a_1, a_2) + \alpha(a_1, h_2) + \alpha(h_1, a_2) + \alpha(h_1, h_2) \\ &= \alpha(a_1, a_2) + D\alpha(a_1, a_2)(h_1, h_2) + \alpha(h_1, h_2).\end{aligned}$$

Da α stetig ist, folgt $\|\alpha(h_1, h_2)\| \leq c\|h_1\| \|h_2\| \leq c\|h\|^2$ für ein $c \geq 0$ und auch die Stetigkeit von $D\alpha(a_1, a_2)$. Denn es gilt:

$$\begin{aligned}\|D\alpha(a_1, a_2)(h_1, h_2)\| &= \|\alpha(a_1, h_2) + \alpha(h_1, a_2)\| \leq \|\alpha(a_1, h_2)\| + \|\alpha(h_1, a_2)\| \\ &\leq c\|a_1\| \cdot \|h_2\| + c\|h_1\| \cdot \|a_2\| \leq c(\|a_1\| + \|a_2\|)\|h\|.\end{aligned}$$

□

Mit Satz 15.4.4 lässt sich die folgende Verallgemeinerung der Produktregel mit Hilfe der Kettenregel beweisen.

Sind X, Y_i normierte Vektorräume, $i \in \{1, \dots, n\}$ und $f_i : U \rightarrow Y_i$ in $a \in U$ differenzierbare Abbildungen, so ist auch

$$f = (f_1, \dots, f_n) : U \rightarrow Y_1 \times \dots \times Y_n$$

in a differenzierbar (wobei wir in $Y_1 \times \dots \times Y_n$ die Maximumsnorm wählen) und es gilt: $Df(a)h = (Df_1(a)h, \dots, Df_n(a)h)$. Damit erhalten wir:

Satz 15.4.5. Sei $\alpha : Y_1 \times \dots \times Y_n \rightarrow Z$ eine multilineare stetige Abbildung und

$$f = (f_1, \dots, f_n) : U \rightarrow Y_1 \times \dots \times Y_n$$

sei in a differenzierbar. Dann ist auch $\alpha \circ f : U \rightarrow Z$ in $a \in U$ differenzierbar und es gilt:

$$\begin{aligned}D(\alpha \circ f)(a)h &= \alpha(Df_1(a)h, f_2(a), \dots, f_n(a)) + \alpha(f_1(a), Df_2(a)h, f_3(a), \dots, f_n(a)) + \dots \\ &\quad + \alpha(f_1(a), \dots, f_{n-1}(a), Df_n(a)h).\end{aligned}$$

Beweis. Aus der Kettenregel und Satz 15.4.4 folgt:

$$\begin{aligned}D(\alpha \circ f)(a)(h) &= D\alpha(f(a))Df(a)(h) = D\alpha(f(a))(Df_1(a)h, \dots, Df_n(a)h) \\ &= \alpha(Df_1(a)h, f_2(a), \dots, f_n(a)) + \alpha(f_1(a), Df_2(a)h, f_3(a), \dots, f_n(a)) + \dots \\ &\quad + \alpha(f_1(a), \dots, f_{n-1}(a), Df_n(a)h).\end{aligned}$$

□

Beispiele.

- (a) Sei Y ein Hilbertraum über \mathbb{R} , d.h. ein Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$, so dass Y bezüglich der Norm $\|y\| = \sqrt{\langle y, y \rangle}$ vollständig ist. Dann ist $\alpha : Y \times Y \rightarrow \mathbb{R}$ mit $\alpha(x, y) = \langle x, y \rangle$ wegen der Cauchy-Schwarzschen Ungleichung 13.2.1

$$|\langle x, y \rangle| \leq \|x\| \|y\| \text{ für alle } x, y \in Y$$

stetig. Seien $(X, \|\cdot\|)$ ein Banachraum, $U \subset X$ offen und $f_1, f_2 : U \rightarrow Y$ differenzierbare Abbildungen. Dann sind auch $f : U \rightarrow Y \times Y$ mit $f(x) = (f_1(x), f_2(x))$ und

$$\langle f_1, f_2 \rangle := \alpha \circ f : U \rightarrow \mathbb{R}$$

differenzierbar und aus Satz 15.4.5 folgt:

$$\begin{aligned} D\langle f_1, f_2 \rangle(a)(h) &= D(\alpha \circ f)(h) = \alpha(Df_1(a)h, f_2(a)) + \alpha(f_1(a), Df_2(a)h) \\ &= \langle Df_1(a)(h), f_2(a) \rangle + \langle f_1(a), Df_2(a)(h) \rangle \end{aligned}$$

Sind insbesondere $Y = \mathbb{R}$ und $\alpha : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ mit $(u, v) \mapsto u \cdot v$, so erhalten wir wieder die Produktregel 15.2.2.

- (b) Sei $\alpha = \det : \mathbb{R}^n \times \cdots \times \mathbb{R}^n \rightarrow \mathbb{R}$ die Determinante und $f_i : (a, b) \rightarrow \mathbb{R}^n$ differenzierbare Kurven, mit $i \in \{1, \dots, n\}$. Dann betrachte man $f : \mathbb{R} \rightarrow \mathbb{R}^n \times \cdots \times \mathbb{R}^n$ mit $f(t) = (f_1(t), \dots, f_n(t))$. Dann ist

$$Df(t)(1) = \dot{f}(t) = (\dot{f}_1(t), \dots, \dot{f}_n(t)) \in \mathbb{R}^n \times \cdots \times \mathbb{R}^n$$

und wir erhalten mit Satz 15.4.4:

$$\begin{aligned} \det(f_1(t), \dots, f_n(t))' &= D(\det \circ f)(t)(1) \\ &= D \det(f(t))(Df(t)(1)) = D \det(f(t))(\dot{f}_1(t), \dots, \dot{f}_n(t)) \\ &= \det(\dot{f}_1(t), f_2(t), \dots, f_n(t)) + \det(f_1(t), \dot{f}_2(t), \dots, f_n(t)) + \dots + \det(f_1(t), \dots, f_{n-1}(t), \dot{f}_n(t)). \end{aligned}$$

- (c) Sei $\alpha : \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) \times \mathbb{R}^n \rightarrow \mathbb{R}^m$ die Bilinearform, gegeben durch $(A, x) \mapsto Ax$. Seien

$$B : (a, b) \rightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) \quad \text{und} \quad c : (a, b) \rightarrow \mathbb{R}^n.$$

differenzierbare Kurven. Dann sind $f : (a, b) \rightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) \times \mathbb{R}^n$ mit $f(t) = (B(t), c(t))$ und $\alpha(B(t), c(t)) = B(t)(c(t))$ differenzierbar mit

$$\begin{aligned} \widehat{B(t)(c(t))} &= D(\alpha \circ f)(t)(1) = D\alpha(f(t))(\dot{f}(t)) \\ &= \alpha(\dot{B}(t), c(t)) + \alpha(B(t), \dot{c}(t)) \\ &= \dot{B}(t)c(t) + B(t)\dot{c}(t). \end{aligned}$$

15.5 Ableitungen höherer Ordnung

Es seien X, Y normierte Vektorräume, $U \subset X$ offen, und $f : U \rightarrow Y$ eine Abbildung. Eine solche Abbildung heißt *stetig differenzierbar* (siehe Def 15.3.13), falls f differenzierbar ist und das Differential $Df : U \rightarrow \mathcal{L}(X, Y)$ stetig ist. Auf $\mathcal{L}(X, Y)$ ist dabei die Operatornorm gewählt. Wir bezeichnen mit

$$C^1(U, Y) = \{f : U \rightarrow Y \mid f \text{ stetig differenzierbar}\}$$

die auf U stetig differenzierbaren Abbildungen mit Werten in Y .

Satz 15.5.1. Sei $f : U \rightarrow Y$ stetig differenzierbar. Dann ist für jedes $h \in X$ die Abbildung $g_h : U \rightarrow Y$ mit $x \mapsto g_h(x) := Df(x)h$ stetig.

Beweis. Man betrachte die lineare Abbildung $\varphi_h : \mathcal{L}(X, Y) \rightarrow Y$ mit $A \mapsto A(h)$. Dann ist φ_h stetig, denn

$$\|\varphi_h(A)\| = \|A(h)\| \leq \|A\| \cdot \|h\| = c\|A\|$$

mit $c = \|h\|$. Da $g_h = \varphi_h \circ Df$ folgt die Behauptung. \square

Die Umkehrung gilt, falls X, Y endliche Dimensionen haben. Dazu benötigen zunächst folgendes Lemma.

Lemma 15.5.2. *Es seien X, Y endlich dimensionale normierte Vektorräume und b_1, \dots, b_n eine Basis von X . Dann ist $\varphi : \mathcal{L}(X, Y) \rightarrow Y \times \dots \times Y =: Y^n$ mit $A \mapsto (Ab_1, \dots, Ab_n)$ eine bijektive lineare Abbildung. Außerdem sind φ und φ^{-1} stetig.*

Bemerkung. Für den Beweis der Bijektivität benötigt man natürlich nur, dass X endlich dimensional ist.

Beweis. Da

$$\varphi(A) = (Ab_1, \dots, Ab_n) = (\varphi_{b_1}(A), \dots, \varphi_{b_n}(A)),$$

folgt die Linearität von φ aus der Linearität der φ_{b_i} . Außerdem ist φ bijektiv, denn zu jedem n -Tupel $(v_1, \dots, v_n) \in Y^n$ existiert genau eine lineare Abbildung $A \in \mathcal{L}(X, Y)$ mit $Ab_i = v_i$ für $i \in \{1, \dots, n\}$. Da X, Y endlichdimensional sind, sind $\mathcal{L}(X, Y)$ und Y^n endlich dimensional und daher ist φ und φ^{-1} stetig für jede Norm (als lineare Abbildung zwischen endlich dimensionalen Vektorräumen). \square

Korollar 15.5.3. *Seien X, Y endlich dimensionale Vektorräume, $U \subset X$ offen und $f : U \rightarrow Y$ differenzierbar. Sei $\{b_1, \dots, b_n\}$ eine Basis von X . Dann ist $Df : U \rightarrow \mathcal{L}(X, Y)$ genau dann stetig, falls $x \mapsto Df(x)b_i$ stetig ist für jedes $i \in \{1, \dots, n\}$.*

Beweis. Aus der Stetigkeit von $Df : U \rightarrow \mathcal{L}(X, Y)$ folgt aus Satz 15.5.1 die Stetigkeit von $x \mapsto Df(x)h$ für jedes $h \in X$.

Sind die Abbildungen $g_{b_i} : U \rightarrow Y$ mit $Df(x)b_i = g_{b_i}(x)$ stetig, so auch $F : U \rightarrow Y^n$ mit

$$F(x) = (g_{b_1}(x), \dots, g_{b_n}(x)).$$

Ist $\varphi : \mathcal{L}(X, Y) \rightarrow Y^n$ der oben definierter Isomorphismus, so folgt: $Df = \varphi^{-1} \circ F : U \rightarrow \mathcal{L}(X, Y)$. Damit erhalten wir die Stetigkeit von Df aus der Stetigkeit von F und φ^{-1} . \square

Korollar 15.5.4. *Seien $U \subset \mathbb{R}^n$ offen und $f : U \rightarrow \mathbb{R}^m$ mit $f(x) = (f_1(x), \dots, f_m(x))$. Dann ist f genau dann stetig differenzierbar, falls alle partiellen Ableitungen $\frac{\partial f_j}{\partial x_i} : U \rightarrow \mathbb{R}$ existieren und stetig sind.*

Beweis. Aus Satz 15.1.7 folgt: existieren alle partiellen Ableitungen $\frac{\partial f_j}{\partial x_i} : U \rightarrow \mathbb{R}$ und sind sie stetig, so ist $f : U \rightarrow \mathbb{R}^m$ differenzierbar. Ist $e_1, \dots, e_n \in \mathbb{R}^n$ die Standardbasis, so sind daher die Abbildungen

$$x \mapsto Df(x)e_i = \left(\frac{\partial f_1}{\partial x_i}(x), \dots, \frac{\partial f_m}{\partial x_i}(x) \right)$$

stetig. Wegen Korollar 15.5.3 ist dann auch $Df : U \rightarrow L(\mathbb{R}^n, \mathbb{R}^m)$ stetig.

Sei umgekehrt $f : U \rightarrow \mathbb{R}^m$ stetig differenzierbar. Dann sind wegen Korollar 15.5.3 die Funktionen

$$x \mapsto Df(x)e_i = \left(\frac{\partial f_1}{\partial x_i}(x), \dots, \frac{\partial f_m}{\partial x_i}(x) \right), \quad i = 1, \dots, n,$$

stetig. Daher sind auch alle partiellen Ableitungen $\frac{\partial f_j}{\partial x_i} : U \rightarrow \mathbb{R}$ stetig. \square

Nun wollen wir Ableitungen höherer Ordnung einführen. Sind X, Y normierte Vektorräume, so ist wie wir gesehen haben,

$$\mathcal{L}(X, Y) = \{A : X \rightarrow Y \mid A \text{ stetig und linear}\}$$

mit der Operatornorm $\|A\| = \sup\{\|A(x)\| \mid \|x\| = 1\}$ ein normierter Vektorraum. Sind X, Y vollständig, so ist auch $\mathcal{L}(X, Y)$ vollständig (siehe Mathematik für Physiker III, Übungsblatt 7, Aufgabe 2).

Seien nun $U \subset X$ offen und $f : U \rightarrow Y$ differenzierbar, so heißt f *zweimal differenzierbar* in $a \in U$, falls $Df : U \rightarrow \mathcal{L}(X, Y) =: Z$ in a differenzierbar ist. Dann ist

$$D(Df)(a) \in \mathcal{L}(X, Z) = \mathcal{L}(X, \mathcal{L}(X, Y)).$$

Statt $D(Df)(a)$ schreiben wir auch $D^2f(a)$. Wir können $\mathcal{L}(X, \mathcal{L}(X, Y))$ mit den stetigen bilinearen Abbildungen

$$\mathcal{L}^2(X; Y) = \{\alpha : X \times X \rightarrow Y \mid \alpha \text{ ist stetig und bilinear}\}$$

identifizieren. Denn ist $B \in \mathcal{L}(X, \mathcal{L}(X, Y))$, so definiere man $\alpha \in \mathcal{L}^2(X; Y)$ durch

$$\alpha(h_1, h_2) := B(h_1)(h_2)$$

für alle $h_1, h_2 \in X$. Es gilt:

Lemma 15.5.5. *Ist $B \in \mathcal{L}(X, \mathcal{L}(X, Y))$, so ist $\alpha : X \times X \rightarrow Y$ mit $\alpha(h_1, h_2) = B(h_1)(h_2)$ eine stetige bilineare Abbildung.*

Beweis. Da $B : X \rightarrow \mathcal{L}(X, Y)$ linear ist, folgt für $\lambda, \lambda' \in K$ und $h_1, h'_1 \in X$:

$$B(\lambda h_1 + \lambda' h'_1) = \lambda B(h_1) + \lambda' B(h'_1).$$

Da beide Seiten wiederum lineare Abbildungen darstellen, können wir sie auf einen weiteren Vektor $h_2 \in X$ anwenden und wir erhalten:

$$\begin{aligned} \alpha(\lambda h_1 + \lambda' h'_1, h_2) &= B(\lambda h_1 + \lambda' h'_1)(h_2) = \lambda B(h_1)(h_2) + \lambda' B(h'_1)(h_2) \\ &= \lambda \alpha(h_1, h_2) + \lambda' \alpha(h'_1, h_2) \end{aligned}$$

Die Linearität von α in der zweiten Variablen folgt, da $B(h_1) : X \rightarrow Y$ für jeden Vektor $h_1 \in X$ linear ist. Wegen

$$\|\alpha(h_1, h_2)\| = \|B(h_1)(h_2)\| \leq \|B(h_1)\| \cdot \|h_2\| \leq \|B\| \cdot \|h_1\| \cdot \|h_2\|$$

ist α stetig (Satz 15.4.2, mit $\|B\|$ wird die Operatornorm bezeichnet). □

Definition 15.5.6. Seien X, Y normierte Vektorräume, $\tilde{U} \subset X$ und $f : \tilde{U} \rightarrow Y$ eine Abbildung. Dann heißt f *k-mal differenzierbar* in $a \in \tilde{U}$, falls eine offene Menge $U \subset \tilde{U}$ existiert, so dass die Ableitungen $Df, D^2f, \dots, D^k f$ mit

$$D^j f = D(D^{j-1} f) : U \rightarrow \mathcal{L}(X, \mathcal{L}(X, \dots, \mathcal{L}(X, Y)) \dots)$$

auf U existieren und $D^{k-1} f$ in a differenzierbar ist. Ist $D^{k-1} f$ für alle $a \in U$ differenzierbar, so heißt die Abbildung $f : U \rightarrow Y$ *k-mal differenzierbar*.

Bemerkungen.

- (a) Wir können $D^k f : U \rightarrow \mathcal{L}(X, \dots, X; Y) = \mathcal{L}^k(X; Y)$ als Abbildung von U in die stetigen multilinearen Abbildungen interpretieren. Denn ist

$$B \in \mathcal{L}(X, \mathcal{L}(X, \dots, \mathcal{L}(X, Y) \dots)),$$

so definiert $\alpha(h_1, h_2, \dots, h_k) = B(h_1)(h_2) \dots (h_k)$ mit $h_i \in X$ eine stetige multilineare Abbildung. Mit

$$C^k(U, Y) := \{f : U \rightarrow Y \mid f \text{ } k\text{-mal differenzierbar und } D^k f \text{ stetig}\}$$

bezeichnen wir die k -mal auf U stetig differenzierbaren Abbildungen.

- (b) Analog zu Korollar 15.5.3 erhält man für endlich dimensionale Vektorräume X und Y und eine auf X gegebenen Basis (b_1, \dots, b_n) :

Eine k -mal differenzierbare Abbildung $f : U \rightarrow Y$ ist genau dann k -mal stetig differenzierbar, falls alle Abbildungen $g_{b_{i_1}, \dots, b_{i_k}} : U \rightarrow Y$ mit

$$g_{b_{i_1}, \dots, b_{i_k}}(x) := D^k f(x)(b_{i_1}, \dots, b_{i_k})$$

mit $i_j \in \{1, \dots, n\}$ stetig sind. Zum Beweis betrachte man zur Indexmenge

$$I_{k,n} = \{(i_1, \dots, i_k) \mid i_j \in \{1, \dots, n\}\}$$

den Vektorraum

$$Y^{I_{k,n}} := \{f \mid f : I_{k,n} \rightarrow Y \text{ Abbildung}\}$$

(siehe Beispiel 2 nach Definition 7.1.6). Nun ist die Abbildung $\varphi : \mathcal{L}^k(X; Y) \rightarrow Y^{I_{k,n}}$ mit

$$\varphi(\alpha)_{i_1, \dots, i_k} := \alpha(b_{i_1}, \dots, b_{i_k})$$

ein Vektorraumisomorphismus, denn zu jedem $f \in Y^{I_{k,n}}$ existiert genau ein $\alpha \in \mathcal{L}^k(X; Y)$ mit $\alpha(b_{i_1}, \dots, b_{i_k}) = f_{i_1, \dots, i_k}$. Da

$$\varphi(D^k f(x))_{i_1, \dots, i_k} = (D^k f(x))(b_{i_1}, \dots, b_{i_k})$$

und $\varphi : \mathcal{L}^k(X; Y) \rightarrow Y^{I_{k,n}}$ bijektiv und stetig ist, folgt die Behauptung. Im Falle $k = 1$ stimmt der Isomorphismus mit dem in Lemma 15.5.2 betrachteten Isomorphismus überein, wenn wir $Y^{I_{1,n}}$ über $f \mapsto (f_1, \dots, f_n) \in Y^n$ mit Y^n identifizieren.

Satz 15.5.7. *Seien X, Y normierte Vektorräume und $U \subset X$ offen. Sei $f : U \rightarrow Y$ in $a \in U$ 2-mal differenzierbar. Dann ist für jedes $h \in X$ die Abbildung $g_h : U \rightarrow Y$ mit $g_h(x) = Df(x)h$ in a differenzierbar und es gilt:*

$$Dg_h(a)(k) = D^2 f(a)(k, h).$$

Beweis. Sei $\varphi_h : \mathcal{L}(X, Y) \rightarrow Y$ mit $\varphi_h : A \rightarrow A(h)$. Diese Abbildung ist linear und stetig und somit auch differenzierbar. Da $Df : U \rightarrow \mathcal{L}(X, Y)$ in a differenzierbar ist, ist auch $g_h = \varphi_h \circ Df$ nach der Kettenregel in a differenzierbar mit

$$Dg_h(a) = D\varphi_h(Df(a)) \circ D^2 f(a) = \varphi_h \circ D^2 f(a).$$

Daraus ergibt sich:

$$Dg_h(a)(k) = \varphi_h(D^2 f(a)(k)) = D^2 f(a)(k)(h) = D^2 f(a)(k, h).$$

□

Für endlich dimensionale Vektorräume gilt folgende Umkehrung von Satz 15.5.7.

Satz 15.5.8. *Es seien X, Y endlich dimensionale Vektorräume, $U \subset X$ offen, $f : U \rightarrow Y$ differenzierbar und $b_1, \dots, b_n \in X$ eine Basis von X . Dann sind folgende Aussagen äquivalent:*

- (a) *f ist zweimal differenzierbar in $a \in U$, d.h. $Df : U \rightarrow \mathcal{L}(X, Y)$ ist in $a \in U$ differenzierbar.*
- (b) *Die Abbildungen $g_{b_i} : U \rightarrow Y$ mit $g_{b_i}(x) = Df(x)b_i$ sind in $a \in U$ differenzierbar.*

Beweis. Ist $Df : U \rightarrow \mathcal{L}(X, Y)$ in $a \in U$ differenzierbar, so sind die Abbildungen g_{b_i} wegen Satz 15.5.7 differenzierbar.

Sei nun umgekehrt die Abbildungen $g_{b_i} : U \rightarrow Y$ mit $g_{b_i}(x) = Df(x)b_i$ in $a \in U$ differenzierbar. Man betrachte den linearen Isomorphismus $\varphi : \mathcal{L}(X, Y) \rightarrow Y^n$ mit $\varphi(A) = (\varphi_{b_1}(A), \dots, \varphi_{b_n}(A))$. Dann gilt:

$$Df(x) = \varphi^{-1}(Df(x)b_1, \dots, Df(x)b_n) = \varphi^{-1}(g_{b_1}(x), \dots, g_{b_n}(x)).$$

Da $F : U \rightarrow Y^n$ mit $F(x) = (g_{b_1}(x), \dots, g_{b_n}(x))$ in $a \in U$ differenzierbar ist, folgt aus der Kettenregel auch die Differenzierbarkeit von $Df = \varphi^{-1} \circ F$ in a . □

Nun wollen wir den Spezialfall $U \subset \mathbb{R}^n$ offen und $Y = \mathbb{R}^m$ behandeln.

Korollar 15.5.9. *Es sei $f : U \rightarrow \mathbb{R}^m$ eine differenzierbare Abbildung. Dann sind folgende Aussagen äquivalent:*

- (a) *f ist zweimal differenzierbar in $a \in U$*
- (b) *Die partiellen Ableitungen $\frac{\partial f}{\partial x_i} : U \rightarrow \mathbb{R}^m$ mit $x \mapsto \frac{\partial f}{\partial x_i}(x)$ sind in $a \in U$ differenzierbar. Außerdem gilt: ist e_1, \dots, e_n die kanonische Basis des \mathbb{R}^n , so folgt:*

$$D^2 f(a)(e_j, e_i) = \frac{\partial}{\partial x_j} \left(\frac{\partial f}{\partial x_i} \right) (a) =: \frac{\partial^2 f}{\partial x_j \partial x_i} (a).$$

Beweis. Ist e_1, \dots, e_n die kanonische Basis des \mathbb{R}^n , so gilt:

$$g_{e_i}(x) = Df(x)e_i = \frac{\partial f}{\partial x_i}(x).$$

Daher ist Df wegen Satz 15.5.8 genau dann in a differenzierbar, falls die partiellen Ableitungen $\frac{\partial f}{\partial x_i}(x)$ in a differenzierbar sind. Aus Satz 15.5.7 folgt:

$$Dg_{e_i}(a)(e_j) = D^2 f(a)(e_j, e_i)$$

und somit $D^2 f(a)(e_j, e_i) = Dg_{e_i}(a)e_j = \frac{\partial}{\partial x_j} g_{e_i}(a) = \frac{\partial}{\partial x_j} \left(\frac{\partial f}{\partial x_i} \right) (a) = \frac{\partial^2 f}{\partial x_j \partial x_i} (a)$. □

Bemerkung. Aus der Bilinearität von $D^2 f(a)$ folgt für $h = \sum_{j=1}^n h_j e_j$ und $k = \sum_{i=1}^n k_i e_i$:

$$D^2 f(a)(h, k) = \sum_{i,j=1}^n h_j k_i \frac{\partial^2 f}{\partial x_j \partial x_i} (a).$$

Das folgende Korollar liefert eine sehr nützliche Charakterisierung für die zweimalige Differenzierbarkeit einer Abbildung $f : U \rightarrow \mathbb{R}^m$, $U \subset \mathbb{R}^n$ offen. Diese Charakterisierung benutzt nur die Existenz und Stetigkeit der ersten und zweiten partiellen Ableitungen und nicht ihre Differenzierbarkeit.

Korollar 15.5.10. *Es sei $f : U \rightarrow \mathbb{R}^m$ eine Abbildung. Dann sind äquivalent*

(a) $f : U \rightarrow \mathbb{R}^m$ ist zweimal stetig differenzierbar.

(b) Die partiellen Ableitungen

$$\frac{\partial f}{\partial x_i}(x) \quad \text{und} \quad \frac{\partial^2 f}{\partial x_j \partial x_i}(x)$$

existieren auf U und sind stetig.

Beweis. Aus (a) folgt (b) wegen Korollar 15.5.9. Da

$$D^2 f(x)(e_j, e_i) = \frac{\partial^2 f}{\partial x_j \partial x_i}(x),$$

folgt die Stetigkeit der zweiten partiellen Ableitungen aus der Stetigkeit von $D^2 f : U \rightarrow \mathcal{L}^2(\mathbb{R}^n; \mathbb{R}^m)$ und Bemerkung (b) nach Definition 15.5.6.

Ist (b) erfüllt, so folgt mit Korollar 15.5.3 die Differenzierbarkeit von f und die Differenzierbarkeit der partiellen Ableitungen $\frac{\partial f}{\partial x_i}$. Also folgt aus Korollar 15.5.9, die zweimalige Differenzierbarkeit von f . Da die Abbildungen

$$D^2 f(x)(e_j, e_i) = \frac{\partial^2 f}{\partial x_j \partial x_i}(x)$$

stetig sind, folgt wieder aus der oben genannten Bemerkung, dass f zweimal stetig differenzierbar ist. \square

Definition 15.5.11. Sei $f : U \rightarrow \mathbb{R}$ zweimal in $a \in U \subset \mathbb{R}^n$ differenzierbar. Dann heißt die $n \times n$ -Matrix

$$(H_f(a))_{j,i} = \frac{\partial^2 f}{\partial x_j \partial x_i}(a)$$

auch *Hessematrix* von f in a .

Bemerkung. Man kann die zweite Ableitung dann auch wie folgt schreiben. Ist $\langle \cdot, \cdot \rangle$ das Standardskalarprodukt auf \mathbb{R}^n , so gilt

$$D^2 f(a)(h, k) = \langle H_f(a)h, k \rangle.$$

Beispiel. Man betrachte $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ mit $f(x_1, x_2) = x_1 x_2^2 + \sin x_1$. Dann folgt: Die partiellen Ableitungen

$$\frac{\partial f}{\partial x_1}(x) = x_2^2 + \cos x_1 \quad \text{und} \quad \frac{\partial f}{\partial x_2}(x) = 2x_1 x_2$$

sind stetig. Außerdem existieren auch die partiellen Ableitungen von $\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2} : \mathbb{R}^2 \rightarrow \mathbb{R}$. Da

$$\begin{aligned} \frac{\partial^2 f}{\partial x_1 \partial x_1}(x) &= -\sin x_1, & \frac{\partial^2 f}{\partial x_2 \partial x_1}(x) &= 2x_2, \\ \frac{\partial^2 f}{\partial x_1 \partial x_2}(x) &= 2x_2, & \frac{\partial^2 f}{\partial x_2 \partial x_2}(x) &= 2x_1, \end{aligned}$$

folgt auch ihre Stetigkeit. Damit ist f zweimal stetig differenzierbar und es folgt:

$$H_f(a) = \begin{pmatrix} -\sin a_1 & 2a_2 \\ 2a_2 & 2a_1 \end{pmatrix}.$$

Dass diese Matrix symmetrisch ist, ist kein Zufall, denn es gilt:

Satz 15.5.12 (*H.A. Schwarz*).

Seien X, Y normierte Vektorräume, $U \subset X$ offen und $f : U \rightarrow Y$ zweimal in $a \in U$ differenzierbar. Dann gilt:

$$D^2 f(a)(h, k) = D^2 f(a)(k, h)$$

für alle $h, k \in X$.

Beweis. Wir zeigen für alle $h, k \in X$:

$$\lim_{\substack{s \rightarrow 0 \\ s \in \mathbb{R}}} \frac{f(a + sh + sk) - f(a + sh) - f(a + sk) + f(a)}{s^2} = D^2 f(a)(k, h).$$

Da die linke Seite symmetrisch in h und k ist, gilt dies dann auch für die rechte Seite.

Sei $\delta > 0$ so gewählt, dass $B(a, \delta) \subset U$. Dann ist $g(t) = f(a + th + k) - f(a + th)$ für $t \in [0, 1]$ und $\|h\|, \|k\| < \delta/2$ definiert.

Aus dem Mittelwertsatz angewandt auf $F(t) = g(t) - t\dot{g}(0)$ folgt:

$$\|g(1) - g(0) - \dot{g}(0)\| = \|F(1) - F(0)\| \leq \sup_{0 \leq t \leq 1} \|\dot{F}(t)\| = \sup_{0 \leq t \leq 1} \|\dot{g}(t) - \dot{g}(0)\|. \quad (15.6)$$

Wir zeigen zunächst: Es existiert eine in $0 \in X$ stetige Abbildung $\varphi : B(0, \delta) \rightarrow \mathcal{L}(X, Y)$ mit $\varphi(0) = 0$ und

$$\dot{g}(t) - D^2 f(a)(k, h) = \varphi(th + k)(h)\|th + k\| - \varphi(th)(h)\|th\|.$$

Da

$$\begin{aligned} \dot{g}(t) &= Df(a + th + k)h - Df(a + th)h = \\ &= (Df(a + th + k) - Df(a))h - (Df(a + th) - Df(a))(h) \end{aligned}$$

folgt aus der Differenzierbarkeit von Df an der Stelle a :

$$\begin{aligned} Df(a + th + k) - Df(a) &= D^2 f(a)(th + k) + \varphi(th + k)\|th + k\| \\ \text{und } Df(a + th) - Df(a) &= D^2 f(a)(th) + \varphi(th)\|th\|, \end{aligned}$$

wobei $\varphi : B(0, \delta) \rightarrow \mathcal{L}(X, Y)$ stetig in 0 ist mit $\varphi(0) = 0$. Durch Einsetzen erhalten wir:

$$\begin{aligned} \dot{g}(t) &= (D^2 f(a)(th + k) + \varphi(th + k)\|th + k\|)(h) - (D^2 f(a)(th) + \varphi(th)\|th\|)(h) \\ &= D^2 f(a)(k, h) + \varphi(th + k)(h)\|th + k\| - \varphi(th)(h)\|th\|. \end{aligned} \quad (15.7)$$

Insbesondere gilt:

$$\dot{g}(0) = D^2 f(a)(k, h) + \varphi(k)(h)\|k\|. \quad (15.8)$$

Damit erhalten wir:

$$\begin{aligned}
& \|f(a+h+k) - f(a+h) - f(a+k) + f(a) - D^2f(a)(k,h)\| = \|g(1) - g(0) - D^2f(a)(k,h)\| \\
& \leq \|g(1) - g(0) - \dot{g}(0)\| + \|\dot{g}(0) - D^2f(a)(k,h)\| \stackrel{(15.6),(15.8)}{\leq} \sup_{0 \leq t \leq 1} \|\dot{g}(t) - \dot{g}(0)\| + \|\varphi(k)(h)\| \|k\| \\
& \leq \sup_{0 \leq t \leq 1} \|\dot{g}(t) - D^2f(a)(k,h)\| + 2\|\varphi(k)(h)\| \|k\| \\
& \stackrel{(15.7)}{\leq} \sup_{0 \leq t \leq 1} (\|\varphi(th+k)(h)\| \|th+k\| + \|\varphi(th)h\| \|th\|) + 2\|\varphi(k)(h)\| \|k\| \\
& \leq \sup_{0 \leq t \leq 1} (\|\varphi(th+k)\| \|h\| \|th+k\| + \|\varphi(th)\| \|h\| \|th\|) + 2\|\varphi(k)\| \|h\| \|k\|.
\end{aligned}$$

Man substituiere $h \mapsto sh, k \mapsto sk$ so folgt:

$$\begin{aligned}
& \|f(a+sh+sk) - f(a+sh) - f(a+sk) + f(a) - s^2Df(a)(k,h)\| \\
& \leq s^2 \sup_{0 \leq t \leq 1} \|\varphi(t \cdot sh+sk)\| \|h\| \|th+k\| + \varphi(t \cdot sh) \|h\| \|th\| + 2s^2\|\varphi(sk)\| \|h\| \|k\|.
\end{aligned}$$

Da φ in 0 stetig, strebt die rechte Seite nach Division durch s^2 gegen null für $s \rightarrow 0$. \square

Bemerkung. Insbesondere folgt aus diesem Satz: Ist $f : U \rightarrow \mathbb{R}, U \subset \mathbb{R}^n$ offen, 2-mal in a differenzierbar, so gilt:

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a).$$

Den Satz von Schwarz werden wir am Ende dieses Abschnitts auf Ableitungen höherer Ordnung ausdehnen. Zunächst notieren wir noch die folgenden Verallgemeinerungen der Sätze 15.5.7 und 15.5.8:

Satz 15.5.13. Sei $f : U \rightarrow Y$ in a $(k+1)$ -mal differenzierbar. Dann gilt für alle $h_1, \dots, h_k \in X$: Die Abbildung $g_{h_1, \dots, h_k} : U \rightarrow Y$ mit

$$g_{h_1, \dots, h_k}(x) = D^k f(x)(h_1, \dots, h_k)$$

ist eine in a differenzierbare Abbildung mit

$$Dg_{h_1, \dots, h_k}(a)k = D^{k+1}f(a)(k, h_1, \dots, h_k).$$

Beweis. Der Beweis ist analog zum Beweis des Satzes 15.5.7. Man betrachte zu $h_1, \dots, h_k \in X$ die lineare Abbildung $\varphi_{h_1, \dots, h_k} : \mathcal{L}^k(X; Y) \rightarrow Y$ mit

$$\varphi_{h_1, \dots, h_k}(\alpha) = \alpha(h_1, \dots, h_k).$$

Man wende dann die Kettenregel auf

$$g_{h_1, \dots, h_k} = \varphi_{h_1, \dots, h_k} \circ D^k f$$

an. \square

Satz 15.5.14. Seien X und Y endlich dimensionale Vektorräume, (b_1, \dots, b_n) eine Basis von X und $f : U \rightarrow Y$ eine auf U k -mal differenzierbare Abbildung. Dann sind folgende Aussagen äquivalent:

(a) f ist $(k + 1)$ -mal in $a \in U$ differenzierbar, d.h. $D^k f : U \rightarrow \mathcal{L}^k(X; Y)$ ist in $a \in U$ differenzierbar.

(b) Die Abbildungen $g_{b_{i_1}, \dots, b_{i_k}} : U \rightarrow Y$ mit

$$g_{b_{i_1}, \dots, b_{i_k}}(x) = D^k f(x)(b_{i_1}, \dots, b_{i_k})$$

sind für alle $b_{i_1}, \dots, b_{i_k} \in X$ in $a \in U$ differenzierbar.

Beweis. Der Beweis ist analog zum Beweis des Satzes 15.5.8. Er benutzt den in Bemerkung (b) nach Definition 15.5.6 erklärten Isomorphismus $\varphi : \mathcal{L}^k(X; Y) \rightarrow Y^{I(k, n)}$. \square

Der folgende Satz ist analog zu 15.5.10. Er liefert ein Kriterium für die k -malige Differenzierbarkeit einer Abbildung $f : U \rightarrow \mathbb{R}^m$ mit $U \subset \mathbb{R}^n$ offen unter Benutzung der partiellen Ableitungen.

Mit $\frac{\partial^j f}{\partial x_{i_1} \dots \partial x_{i_j}}$ bezeichnen wir die partiellen Ableitungen der Ordnung j . Sie sind induktiv definiert: $\frac{\partial^j f}{\partial x_{i_1} \dots \partial x_{i_j}} = \frac{\partial}{\partial x_{i_1}} \left(\frac{\partial^{j-1} f}{\partial x_{i_2} \dots \partial x_{i_j}} \right)$.

Satz 15.5.15. Es sei $f : U \rightarrow \mathbb{R}^m$ eine Abbildung. Dann sind äquivalent:

(a) $f : U \rightarrow \mathbb{R}^m$ ist k -mal stetig differenzierbar.

(b) Die partiellen Ableitungen

$$\frac{\partial^j f}{\partial x_{i_1} \dots \partial x_{i_j}} : U \rightarrow \mathbb{R}^m$$

existieren für alle $j \leq k$ und sind auf U stetig. Dann gilt insbesondere

$$D^j f(x)(e_{i_1}, \dots, e_{i_j}) = \frac{\partial^j f}{\partial x_{i_1} \dots \partial x_{i_j}}(x)$$

für alle $j \leq k$.

Beweis. Sei e_1, \dots, e_n die kanonische Basis von \mathbb{R}^n . Aus (a) folgt (b) wegen

$$\frac{\partial^j f}{\partial x_{i_1} \dots \partial x_{i_j}}(x) = \varphi_{e_{i_1}, \dots, e_{i_j}} \circ D^j f(x).$$

Die Implikation (b) \Rightarrow (a) zeigt man durch Induktion über k . Für $k = 1$ ist dies Korollar 15.5.4. Man nehme nun an, dass die Aussage für $k \in \mathbb{N}$ bewiesen sei und dass alle partiellen Ableitungen

$$\frac{\partial^j f}{\partial x_{i_1} \dots \partial x_{i_j}} : U \rightarrow \mathbb{R}^m$$

für $j \leq k + 1$ existieren und stetig sind. Dann ist nach Induktionsannahme f k -mal stetig differenzierbar. Da die partiellen Ableitungen

$$\frac{\partial}{\partial x_i} \frac{\partial^j f}{\partial x_{i_1} \dots \partial x_{i_k}}(x) = \frac{\partial^{k+1} f}{\partial x_i \partial x_{i_1} \dots \partial x_{i_k}}(x)$$

existieren und stetig sind, sind wegen Korollar 15.5.4 die Funktionen

$$D^k f(x)(e_{i_1}, \dots, e_{i_k}) = \frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}}(x)$$

für alle $e_{i_1}, \dots, e_{i_k} \in X$ auf U differenzierbar. Wegen Satz 15.5.14 ist damit f $(k+1)$ -mal differenzierbar. Die Stetigkeit von $D^{k+1}f$ folgt aus der Stetigkeit der partiellen Ableitungen

$$\frac{\partial^{k+1} f}{\partial x_{i_1} \dots \partial x_{i_{k+1}}}(x).$$

□

Durch vollständige Induktion erhält man die Verallgemeinerung des Satzes von Schwarz.

Satz 15.5.16. *Seien X, Y normierte Vektorräume, $U \subset X$ offen und $f : U \rightarrow Y$ eine in $a \in U$ k -mal differenzierbare Abbildung. Dann ist die stetige multilineare Abbildung*

$$(h_1, \dots, h_k) \mapsto D^k f(a)(h_1, \dots, h_k)$$

symmetrisch, d.h. es gilt:

$$D^k f(a)(h_{\sigma(1)}, \dots, h_{\sigma(k)}) = D^k f(a)(h_1, \dots, h_k)$$

für jede Permutation $\sigma \in S_k$.

Beweis. Der Beweis wird durch Induktion über $k \geq 2$ geführt. Für $k = 2$ ist dies Satz 15.5.12. Wir nehmen an, dass die Aussage des Satzes für $k - 1 \geq 2$ richtig sei. Wir zeigen zunächst, dass $D^k f(a)(h_1, \dots, h_k)$ in den letzten $k - 1$ und den ersten beiden Variablen symmetrisch ist. Man betrachte dazu

$$g_{h_2, \dots, h_k}(x) = D^{k-1} f(x)(h_2, \dots, h_k) = D^{k-1} f(x)(h_2, \dots, h_k).$$

Nach Induktionsvoraussetzung sind beide Seiten symmetrisch in (h_2, \dots, h_k) . Wegen

$$Dg_{h_2, \dots, h_k}(a)(h_1) = D^k f(a)(h_1, \dots, h_k)$$

ist somit auch $D^k f(a)(h_1, \dots, h_k)$ symmetrisch in (h_2, \dots, h_k) . Man betrachte nun

$$\varphi(x) = D^{k-2} f(x)(h_3, \dots, h_k),$$

so gilt wegen Satz 15.5.7:

$$D^2 \varphi(a)(h_1, h_2) = D(D\varphi(a)(h_2))(h_1) = D^k f(a)(h_1, \dots, h_k).$$

Wegen des Satzes von Schwarz ist $D^2 \varphi(a)(h_1, h_2)$ symmetrisch in (h_1, h_2) und somit ist auch $D^k f(a)(h_1, \dots, h_k)$ symmetrisch in (h_1, h_2) . Daraus folgt aber

$$D^k f(a)(h_{\sigma(1)}, \dots, h_{\sigma(k)}) = D^k f(a)(h_1, \dots, h_k)$$

für jede Permutation $\sigma \in S_k$. Denn ist $\sigma(1) = 1$, so erhalten wir die Aussage durch Vertauschung der letzten $(k - 1)$ Variablen. Ist $\sigma(1) \neq 1$, so permutiere die letzten $k - 1$ Variablen, so dass an der zweiten Stelle h_1 steht, d.h.

$$D^k f(a)(h_{\sigma(1)}, \dots, h_{\sigma(k)}) = D^k f(a)((h_{\sigma(1)}, h_1, \dots)).$$

Vertauschen wir nun $h_{\sigma(1)}$ und h_1 , so folgt wieder die Behauptung durch Vertauschung der letzten $(k - 1)$ Variablen. □

15.6 Taylorformel

Zunächst wollen wir nochmal an die Taylorformel aus dem ersten Semester erinnern (siehe Satz 4.5.2 und anschließende Bemerkung (c)).

Sei $F : [a, b] \rightarrow \mathbb{R}$ eine n -mal stetig differenzierbare Funktion, die auf (a, b) $(n + 1)$ -mal differenzierbar ist. Sind $t_0 \in [a, b]$ und $t_0 + h \in [a, b]$, so gilt:

$$F(t_0 + h) = \sum_{k=0}^n \frac{F^{(k)}(t_0)}{k!} h^k + \frac{F^{(n+1)}(t_0 + \delta h)}{(n+1)!} h^{n+1}$$

für ein $\delta \in (0, 1)$.

Hierbei ist das Restglied von Lagrange gewählt. Dieser Satz lässt sich unmittelbar auf reellwertige Funktionen auf normierten Vektorräumen übertragen.

Satz 15.6.1. *Sei X ein normierter Vektorraum, $U \subset X$ offen und $f : U \rightarrow \mathbb{R}$ eine $(n + 1)$ -mal differenzierbare Abbildung. Es sei $x_0 \in U$ und $h \in X$, so dass die Verbindungsstrecke $\gamma(t) = x_0 + th$ für $t \in [0, 1]$ ganz in U verläuft. Dann gilt für ein $\delta \in (0, 1)$:*

$$\begin{aligned} f(x_0 + h) &= f(x_0) + Df(x_0)h + \frac{1}{2!} D^2 f(x_0)(h, h) \\ &\quad + \dots + \frac{1}{n!} D^n f(x_0)(h, \dots, h) + \frac{1}{(n+1)!} D^{n+1} f(x_0 + \delta h)(h, \dots, h). \end{aligned}$$

Beweis. Zu $F : [0, 1] \rightarrow \mathbb{R}$ mit $F(t) = f \circ \gamma(t)$ betrachte die Taylorformel mit Entwicklungspunkt $t_0 = 0$. Dann gilt

$$F(1) = \sum_{k=0}^n \frac{F^{(k)}(0)}{k!} + \frac{F^{(n+1)}(\delta)}{(n+1)!}$$

für ein $\delta \in (0, 1)$. Es gilt:

$$F^{(1)}(t) = (f \circ \gamma(t))' = Df(\gamma(t))(\dot{\gamma}(t)) = Df(\gamma(t))(h).$$

Definiere $g_1(x) := Df(x)h$, so folgt aus Satz 15.5.7

$$F^{(2)}(t) = (g_1 \circ \gamma(t))' = Dg_1(\gamma(t))(\dot{\gamma}(t)) = Dg_1(\gamma(t))(h) = D^2 f(\gamma(t))(h, h).$$

Sei

$$F^{(n)}(t) = D^n f(\gamma(t))(h, \dots, h),$$

so definiere $g_n(x) = D^n f(x)(h, \dots, h)$. Dann folgt mit Hilfe der Kettenregel und Satz 15.5.13:

$$F^{(n+1)}(t) = (g_n \circ \gamma(t))' = Dg_n(\gamma(t))(h) = D^{(n+1)} f(\gamma(t))(h, \dots, h).$$

Damit erhalten wir

$$f(x_0 + h) = F(1) = \sum_{k=0}^n \frac{1}{k!} D^k f(x_0)(h, \dots, h) + \frac{1}{(n+1)!} D^{n+1} f(x_0 + \delta h)(h, \dots, h).$$

□

Eine Taylorformel für vektorwertige Funktionen erhalten wir mit Hilfe der partiellen Integration. Diese ergibt sich aus folgender Produktregel.

Lemma 15.6.2. *Seien $I = (a, b)$ ein offenes Intervall, Y ein Banachraum und $f : I \rightarrow Y$, sowie $\varphi : I \rightarrow \mathbb{R}$ differenzierbare Abbildungen. Dann gilt für $t \in I$:*

$$\widehat{f \cdot \varphi}(t) = f(t) \cdot \dot{\varphi}(t) + \dot{f}(t) \cdot \varphi(t).$$

Beweis. Sei $\alpha : Y \times \mathbb{R} \rightarrow Y$ die bilineare Abbildung mit $\alpha(y, t) = y \cdot t$. Wegen $\|\alpha(y, t)\| = \|y\| \cdot |t|$ ist α stetig. Betrachte die Abbildung $F : I \rightarrow Y \times \mathbb{R}$ mit $F(t) = (f(t), \varphi(t))$. Diese Abbildung ist differenzierbar und es gilt: $f(t) \cdot \varphi(t) = \alpha(F(t))$. Wegen der Kettenregel ist dann auch $t \mapsto f(t) \cdot \varphi(t)$ differenzierbar und aus Satz 15.4.4 folgt:

$$\begin{aligned} \widehat{f \cdot \varphi}(t) &= D\alpha(f(t), \varphi(t))(\dot{f}(t), \dot{\varphi}(t)) = \alpha(f(t), \dot{\varphi}(t)) + \alpha(\dot{f}(t), \varphi(t)) \\ &= f(t) \cdot \dot{\varphi}(t) + \dot{f}(t) \cdot \varphi(t). \end{aligned}$$

□

Korollar 15.6.3 (Partielle Integration). *Seien $[a, b] \subset I$, $I \subset \mathbb{R}$ ein offenes Intervall, Y ein Banachraum und $f : I \rightarrow Y$, sowie $\varphi : I \rightarrow \mathbb{R}$ stetig differenzierbare Abbildungen. Dann gilt:*

$$\int_a^b f(t) \dot{\varphi}(t) dt = f(t) \varphi(t) \Big|_a^b - \int_a^b \dot{f}(t) \varphi(t) dt.$$

Beweis. Da $t \mapsto \widehat{f \cdot \varphi}(t)$ stetig ist, folgt aus dem Hauptsatz der Differential- und Integralrechnung (für vektorwertige Funktionen) und Lemma 15.6.2:

$$f(t) \varphi(t) \Big|_a^b = f(b) \varphi(b) - f(a) \varphi(a) = \int_a^b \widehat{f \cdot \varphi}(t) dt = \int_a^b f(t) \dot{\varphi}(t) dt + \int_a^b \dot{f}(t) \varphi(t) dt.$$

□

Dies führt nun zu folgender Taylorformel für Kurven.

Satz 15.6.4 (Taylorformel für Kurven). *Sei $I \subset \mathbb{R}$ offen, Y Banachraum und $f : I \rightarrow Y$ eine $(n+1)$ -mal stetig differenzierbare Abbildung. Ist $[a, b] \subset I$, so gilt:*

$$f(b) = f(a) + \dot{f}(a)(b-a) + \frac{f^{(2)}(a)}{2!}(b-a)^2 + \dots + \frac{f^{(n)}(a)}{n!}(b-a)^n + \int_a^b f^{(n+1)}(t) \frac{(b-t)^n}{n!} dt$$

wobei $f^{(k)}(t) = \widehat{f^{(k-1)}}(t)$.

Beweis. Der Beweis folgt durch Induktion über \mathbb{N}_0 mit Hilfe der partiellen Integration. Für $n = 0$ ist die Aussage richtig, denn in diesem Fall reduziert sie sich auf den Hauptsatz der Integralrechnung

$$f(b) - f(a) = \int_a^b \dot{f}(t) dt.$$

Die Formel gelte für $n \in \mathbb{N}$. Wir wollen sie für $(n+1)$ zeigen.

Sei also f eine $(n+2)$ -mal stetig differenzierbare Abbildung. Nach Induktionsvoraussetzung gilt:

$$f(b) = f(a) + f'(a)(b-a) + \frac{f^{(2)}(a)}{2!}(b-a)^2 + \dots + \frac{f^{(n)}(a)}{n!}(b-a)^n + \int_a^b f^{(n+1)}(t) \frac{(b-t)^n}{n!} dt$$

Sei $g(t) = f^{(n+1)}(t)$ und $\varphi(t) = -\frac{(b-t)^{n+1}}{(n+1)!}$. Also ist $\dot{\varphi}(t) = \frac{(b-t)^n}{n!}$ und mittels partieller Integration folgt:

$$\begin{aligned} \int_a^b f^{(n+1)}(t) \frac{(b-t)^n}{n!} dt &= \int_a^b g(t) \cdot \dot{\varphi}(t) dt = g(t) \cdot \varphi(t) \Big|_a^b - \int_a^b \dot{g}(t) \cdot \varphi(t) dt \\ &= f^{(n+1)}(a) \cdot \frac{(b-a)^{n+1}}{(n+1)!} + \int_a^b f^{(n+2)}(t) \frac{(b-t)^{n+1}}{(n+1)!} dt, \end{aligned}$$

d.h. die Taylorformel gilt auch für $n+1$. □

Daraus ergibt sich mittels Kettenregel sofort die folgende Taylorformel für allgemeine Abbildungen zwischen Banachräumen.

Satz 15.6.5. *Seien X, Y Banachräume, $U \subset X$ offen und $f : U \rightarrow Y$ eine $(n+1)$ -mal stetig differenzierbare Abbildung. Es sei $x_0 \in U$ und $h \in X$, so dass die Verbindungsstrecke $\gamma(t) = x_0 + th$ für $t \in [0, 1]$ ganz in U verläuft. Dann gilt:*

$$\begin{aligned} f(x_0 + h) &= f(x_0) + Df(x_0)h + \frac{1}{2!}D^2f(x_0)(h, h) + \dots + \frac{1}{n!}D^n f(x_0)(h, \dots, h) \\ &\quad + \int_0^1 \frac{D^{n+1}f(x_0 + th)(h, \dots, h)}{n!} (1-t)^n dt. \end{aligned}$$

Beweis. Betrachte $F : (-\epsilon, 1 + \epsilon) \rightarrow Y$ mit $F(t) = f \circ \gamma(t)$, wobei $\epsilon > 0$ genügend klein gewählt ist. Wenden wir den Satz 15.6.4 auf die Kurve $F : (-\epsilon, 1 + \epsilon) \rightarrow Y$ und das Intervall $[0, 1]$ an, so erhalten wir:

$$\begin{aligned} f(x_0 + h) - f(x_0) &= F(1) - F(0) = \dot{F}(0) + \frac{F^{(2)}(0)}{2!} + \dots + \frac{F^{(n)}(0)}{n!} + \int_0^1 F^{(n+1)}(t) \frac{(1-t)^n}{n!} dt \\ &= Df(x_0)h + \frac{D^2f(x_0)(h, h)}{2!} + \dots + \frac{D^n f(x_0)(h, \dots, h)}{n!} \\ &\quad + \int_0^1 D^{n+1}f(x_0 + th)(h, \dots, h) \frac{(1-t)^n}{n!} dt \end{aligned}$$

□

Bemerkung. Ist also $f : U \rightarrow Y$ eine $(n + 1)$ -mal differenzierbare Abbildung von einer konvexen Menge U nach Y , so lässt sich f als Summe von multilinearen symmetrischen Abbildungen vom Grade $\leq n$ approximieren.

15.7 Extremwerte

Wir wollen nun die Taylorformel auf das Studium von Extremwerten anwenden.

Definition 15.7.1. Sei X ein normierter Vektorraum, $D \subset X$ und $f : D \rightarrow \mathbb{R}$. Dann heißt $x_0 \in D$ ein *lokales Maximum bzw. Minimum*, falls es ein $\epsilon > 0$ gibt mit $B(x_0, \epsilon) \subset D$, so dass

$$f(x_0) \geq f(x) \text{ bzw. } f(x_0) \leq f(x)$$

für alle $x \in B(x_0, \epsilon)$ gilt.

Bemerkung. Ein lokales Maximum bzw. Minimum x_0 nennen wir auch *lokales Extremum (Extremalstelle)*.

Definition 15.7.2. Sei $U \subset X$ offen und $f : U \rightarrow \mathbb{R}$ differenzierbar. Dann heißt $x_0 \in U$ *kritischer Punkt*, falls $Df(x_0) = 0$ gilt.

Satz 15.7.3. Sei $f : U \rightarrow \mathbb{R}$ differenzierbar, $U \subset X$ offen und $x_0 \in U$ ein lokales Extremum. Dann ist $Df(x_0) = 0$.

Beweis. Da die Abbildung $f : U \rightarrow \mathbb{R}$ auf U differenzierbar und U offen ist, existiert für jedes $h \in X$ ein $\epsilon > 0$, so dass die Funktion $g : (-\epsilon, \epsilon) \rightarrow \mathbb{R}$ mit $g(t) = f(x_0 + th)$ definiert und differenzierbar ist. Aus der Annahme folgt, dass $0 \in \mathbb{R}$ ein lokales Extremum von g ist. Also ist wegen Satz 4.4.3 $Df(x_0)h = g'(0) = 0$. \square

Um zu entscheiden, ob es sich bei einem Extremum x_0 um ein (lokales) Maximum bzw. Minimum handelt, spielt die zweite Ableitung eine zentrale Rolle. Wie wir im letzten Abschnitt gesehen haben ist die zweite Ableitung durch eine bilineare symmetrische reellwertige Form gegeben. Bilineare symmetrische Formen haben wir im Rahmen der linearen Algebra in Abschnitt 13.7 untersucht. Wir möchten an die folgende Begriffsbildung erinnern.

Definition 15.7.4. Sei $\alpha \in L^2(X, \mathbb{R})$ eine symmetrische bilineare reellwertige Abbildung (solche Abbildungen heißen auch *quadratische Formen*). Dann nennen wir α

- (a) *positiv definit (positiv semidefinit)* falls $\alpha(h, h) > 0$ ($\alpha(h, h) \geq 0$) für alle $h \in X$ mit $h \neq 0$.
- (b) *negativ definit (negativ semidefinit)* falls $\alpha(h, h) < 0$ ($\alpha(h, h) \leq 0$) für alle $h \neq 0$.
- (c) *indefinit* falls $h \mapsto \alpha(h, h)$ sowohl positive als auch negative Werte annimmt.

Bemerkung. Ist X ein endlich dimensionaler reeller Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$, so existiert ein symmetrischer Endomorphismus $A : X \rightarrow X$ mit $\alpha(h, h) = \langle A(h), h \rangle$. Dann folgt:

- (a) α ist genau dann positiv definit bzw. positiv semidefinit, wenn alle Eigenwerte von A größer bzw. größer oder gleich null sind.

(b) α ist genau dann negativ definit bzw. negativ semidefinit, wenn alle Eigenwerte von A kleiner bzw. kleiner oder gleich null sind.

(c) α ist genau dann indefinit, falls A sowohl negative als auch positive Eigenwerte besitzt.

Satz 15.7.5. *Es sei X ein Banachraum, $U \subset X$ offen $f : U \rightarrow \mathbb{R}$ eine zweimal stetig differenzierbare Funktion. Sei $x_0 \in U$ ein kritischer Punkt von f , so gilt:*

(a) *Ist $D^2f(x_0)$ indefinit, so ist x_0 kein Extremum.*

(b) *Ist X endlich dimensional und $D^2f(x_0)$ positiv definit (negativ definit), so ist x_0 ein lokales Minimum (Maximum).*

Beweis. Sei $x_0 \in U$ kritischer Punkt und $D^2f(x_0)$ indefinit, d.h. es existieren $h_1, h_2 \in X$ mit

$$D^2f(x_0)(h_1, h_1) > 0 \text{ und } D^2f(x_0)(h_2, h_2) < 0.$$

Da die Abbildung $f : U \rightarrow \mathbb{R}$ auf U zweimal stetig differenzierbar und U offen ist, existiert ein $\epsilon > 0$, so dass die Funktionen $g_{h_1} : (-\epsilon, \epsilon) \rightarrow \mathbb{R}$ und $g_{h_2} : (-\epsilon, \epsilon) \rightarrow \mathbb{R}$ mit $g_{h_1}(t) = f(x_0 + th_1)$ und $g_{h_2}(t) = f(x_0 + th_2)$ definiert und zweimal stetig differenzierbar sind. Da $g'_{h_1}(0) = Df(x_0)h_1 = 0$ und $g''_{h_1}(0) = D^2f(x_0)(h_1, h_1) > 0$, so ist 0 wegen Satz 4.5.5 ein (striktes) lokales Minimum für g_{h_1} , d.h.

$$g_{h_1}(t) = f(x_0 + th_1) < f(x_0) = g_{h_1}(0)$$

für $|t| < \delta_1, t \neq 0$. Genauso folgt die Existenz von $\delta_2 > 0$, so dass $g_{h_2}(t) = f(x_0 + th_2) > f(x_0)$ für $|t| < \delta_2, t \neq 0$. Somit existieren in jeder Umgebung von x_0 Punkte x mit $f(x) < f(x_0)$ sowie solche mit $f(x) > f(x_0)$.

Sei nun $D^2f(x_0)$ positiv definit. Aus der Taylorformel für reellwertige Funktionen folgt:

$$f(x_0 + h) = f(x_0) + \frac{1}{2}D^2f(x_0 + \delta h)(h, h)$$

für ein $\delta \in (0, 1)$. Ist X endlich dimensional, so ist die Sphäre

$$S(0, 1) = \{h \in X \mid \|h\| = 1\}$$

als eine abgeschlossene und beschränkte Teilmenge von X kompakt. Da die quadratische Form $h \mapsto D^2f(x_0)(h, h)$ stetig und positiv ist, existiert ein $c > 0$ mit

$$D^2f(x_0)(h, h) \geq c$$

für alle $h \in S(0, 1)$. Daher gilt auch $D^2f(x_0)(h, h) \geq c\|h\|^2$ für alle $h \in X$. Wegen der Stetigkeit von $x \mapsto D^2f(x)$ existiert ein $\epsilon > 0$ mit $D^2f(x)(h, h) \geq \frac{c}{2}\|h\|^2$ für alle $x \in B(x_0, \epsilon)$. Dies erhalten wir, wenn $\epsilon > 0$ so wählen, dass

$$|D^2f(x)(h, h) - D^2f(x_0)(h, h)| \leq \|D^2f(x) - D^2f(x_0)\| \|h\|^2 \leq \frac{c}{2}\|h\|^2$$

für alle $x \in B(x_0, \epsilon)$ gilt. Denn dann erhalten wir:

$$\begin{aligned} D^2f(x)(h, h) &\geq D^2f(x_0)(h, h) - |D^2f(x)(h, h) - D^2f(x_0)(h, h)| \\ &\geq c\|h\|^2 - \frac{c}{2}\|h\|^2 = \frac{c}{2}\|h\|^2. \end{aligned}$$

Ist $\|h\| < \epsilon$, so folgt aus der Taylorformel für reellwertige Funktionen

$$f(x_0 + h) - f(x_0) = \frac{1}{2}D^2f(x_0 + \delta h)(h, h) \geq \frac{1}{4}c\|h\|^2.$$

Der andere Fall folgt analog. □

Bemerkungen.

- (a) Die Aussage (b) des Satzes lässt sich auch im Fall X unendlich dimensional beweisen, falls ein $c > 0$ existiert mit $D^2f(x_0)(h, h) \geq c\|h\|^2$ oder $D^2f(x_0)(h, h) \leq -c\|h\|^2$, für alle $h \in X$.
- (b) Ist $X = \mathbb{R}^n$ und $\langle \cdot, \cdot \rangle$ das kanonische Skalarprodukt, so gilt:

$$D^2f(x_0)(h, h) = \langle H_f(x_0)h, h \rangle,$$

wobei $H_f(x_0)$ die Hessematrix 15.5.11 von f in x_0 bezeichnet. Ist x_0 ein kritischer Punkt, so liefern also die Eigenwerte von $H_f(x_0)$ hinreichende Kriterien dafür, ob ein (lokales) Maximum oder Minimum vorliegt.

Ist x_0 ein kritischer Punkt und besitzt $H_f(x_0)$ sowohl negative als auch positive Eigenwerte, so heißt x_0 *Sattelpunkt*, falls $H_f(x_0)$ invertierbar ist.

- (c) Ist x_0 kritischer Punkt und $D^2f(x_0)$ positiv semidefinit (negativ semidefinit), so kann man keine allgemeinen Aussagen machen. Betrachte z.B. $f(x, y) = x^2 + y^3$, so ist $(0, 0)$ kritischer Punkt und

$$H_f(x, y) = \begin{pmatrix} 2 & 0 \\ 0 & 6y \end{pmatrix}.$$

Daher ist

$$D^2f(0, 0)(h, h) = \langle H_f(0, 0)h, h \rangle = \left\langle \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} h, h \right\rangle$$

positiv semidefinit. Auf der anderen Seite ist $(0, 0)$ kein lokales Extremum, denn f nimmt in jeder Umgebung positive wie auch negative Werte an.

Umgekehrt ist $(0, 0)$ ein Minimum für $f(x, y) = x^2 + y^4$, und $D^2f(0, 0)$ ist (nur) positiv-semidefinit.

Kapitel 16

Der Banachsche Fixpunktsatz und seine Anwendungen

16.1 Banachscher Fixpunktsatz

Definition 16.1.1. Sei (M, d) ein metrischer Raum. Eine Abbildung

$$\phi : M \rightarrow M$$

heißt *kontrahierend*, wenn es ein $c \in [0, 1)$ gibt mit

$$d(\phi(x), \phi(y)) \leq c d(x, y)$$

für alle $x, y \in M$.

Satz 16.1.2 (Banachscher Fixpunktsatz). *Sei (M, d) ein vollständiger metrischer Raum und $\phi : M \rightarrow M$ eine kontrahierende Abbildung. Dann hat M genau einen Fixpunkt, d.h. es gibt genau ein $x \in M$ mit*

$$\phi(x) = x.$$

Beweis.

(1) Existenz:

Sei $x_0 \in M$ beliebig. Betrachte die Folge $(x_n)_{n \in \mathbb{N}_0}$, die induktiv definiert ist durch $x_{n+1} = \phi(x_n)$. Dann gilt für alle $n \geq 1$ die Abschätzung

$$d(x_{n+1}, x_n) = d(\phi(x_n), \phi(x_{n-1})) \leq c d(x_n, x_{n-1}).$$

Daraus folgt durch Induktion für $k \geq 1$

$$d(x_{n+k}, x_{n+k-1}) \leq c^{k-1} d(x_{n+1}, x_n).$$

Aus der Anwendung der Dreiecksungleichung erhalten wir

$$\begin{aligned} d(x_{n+k}, x_n) &\leq d(x_{n+k}, x_{n+k-1}) + d(x_{n+k-1}, x_{n+k-2}) + \dots + d(x_{n+1}, x_n) \\ &\leq (c^{k-1} + c^{k-2} + \dots + 1) d(x_{n+1}, x_n) \\ &\leq \frac{1 - c^k}{1 - c} c^n d(x_1, x_0) \leq \frac{c^n}{1 - c} d(x_1, x_0). \end{aligned}$$

Da c^n eine Nullfolge ist, ist (x_n) eine Cauchyfolge. Wegen der Vollständigkeit von M konvergiert diese Folge, d.h. der Grenzwert $\lim_{n \rightarrow \infty} x_n = x$ existiert. Aus der Stetigkeit von ϕ folgt:

$$\phi(x) = \phi\left(\lim_{n \rightarrow \infty} x_n\right) = \lim_{n \rightarrow \infty} \phi(x_n) = \lim_{n \rightarrow \infty} x_{n+1} = x$$

und somit ist $x \in M$ ein Fixpunkt.

(2) Eindeutigkeit: Ist $y \in M$ ein weiterer Fixpunkt von ϕ , so ist

$$d(y, x) = d(\phi(y), \phi(x)) \leq c d(y, x)$$

und somit $d(y, x)(1 - c) \leq 0$. Da $1 - c > 0$ und $d(y, x) \geq 0$ ist $d(y, x) = 0$, d.h. $y = x$.

□

□

Bemerkung. Sei (M, d) ein vollständiger metrischer Raum und $A \subset M$ abgeschlossen, so ist auch (A, d) ein vollständiger metrischer Raum. Denn ist (x_n) eine Cauchyfolge in A , so hat sie einen Grenzwert in M . Dieser Grenzwert liegt wegen der Abgeschlossenheit von A in A . Damit folgt aus dem Banachschen Fixpunktsatz: Ist $f : M \rightarrow M$ eine Abbildung, die eine abgeschlossene Teilmenge A invariant lässt, d.h. es gelte $f(A) \subset A$, und ist $f : A \rightarrow A$ eine Kontraktion, so existiert genau ein $x \in A$ mit $f(x) = x$.

16.2 Lokaler Umkehrsatz

Ist $f : (a, b) \rightarrow \mathbb{R}$ eine stetig differenzierbare Funktion und ist $f'(x_0) \neq 0$, so existiert ein $\epsilon > 0$ mit $f'(x) \neq 0$, für alle $x \in (-\epsilon + x_0, \epsilon + x_0)$. Wegen Korollar 4.4.6 ist dann die Abbildung $f : (-\epsilon + x_0, \epsilon + x_0) \rightarrow f(-\epsilon + x_0, \epsilon + x_0)$ streng monoton und somit bijektiv. Wegen Satz 4.3.3 ist dann ihre Umkehrfunktion ebenfalls differenzierbar.

Wir wollen diesen Satz auf differenzierbare Abbildungen zwischen Vektorräumen verallgemeinern. Dazu werden wir in entscheidender Weise den Banachschen Fixpunktsatz verwenden. Die Ableitung der Umkehrabbildung lässt sich aus der Kettenregel berechnen, falls die Umkehrabbildung differenzierbar ist.

Seien $U \subset X, V \subset Y$ offene Mengen in Banachräumen und $f : U \rightarrow V$ eine bijektive Abbildung. Ist f in $a \in U$ und f^{-1} in $b = f(a)$ differenzierbar, so folgt aus der Kettenregel

$$\begin{aligned} \text{id}_X &= D(f^{-1} \circ f)(a) = Df^{-1}(f(a))Df(a) \\ \text{id}_Y &= D(f \circ f^{-1})(f(a)) = Df(a)Df^{-1}(f(a)) \end{aligned}$$

und somit

$$Df^{-1}(f(a)) = (Df(a))^{-1} \in \mathcal{L}(Y, X).$$

Der nächste Satz zeigt, dass eine differenzierbare bijektive Abbildung unter gewissen Voraussetzungen eine differenzierbare Inverse besitzt. Dies ist die Verallgemeinerung des Satzes 4.3.3.

Satz 16.2.1. Seien X, Y Banachräume, $U \subset X, V \subset Y$ offen und $f : U \rightarrow V$ bijektiv. Sei f in $a \in U$ differenzierbar und $Df(a)$ invertierbar mit

$$(Df(a))^{-1} \in \mathcal{L}(Y, X).$$

Ist $f^{-1} : V \rightarrow U$ in $b = f(a)$ stetig, so ist f^{-1} auch in b differenzierbar und es gilt:

$$Df^{-1}(b) = (Df(a))^{-1}.$$

Beweis. Zu zeigen ist:

$$\lim_{h \rightarrow 0} \frac{\|f^{-1}(b+h) - f^{-1}(b) - (Df(a))^{-1}(h)\|}{\|h\|} = 0.$$

Setze

$$r(h) = f^{-1}(b+h) - f^{-1}(b) = f^{-1}(b+h) - a.$$

Dann folgt:

$$h = f(a+r(h)) - f(a) = Df(a)(r(h)) + \varphi(r(h))\|r(h)\|,$$

wobei $\lim_{x \rightarrow 0} \varphi(x) = 0$, da f in a differenzierbar ist. Insbesondere ist $q(h) = \varphi(r(h))$ stetig in 0 mit $q(0) = 0$. Wenden wir auf beiden Seiten $Df(a)^{-1}$ an, so erhalten wir

$$Df(a)^{-1}(h) = r(h) + Df(a)^{-1}(q(h))\|r(h)\| \quad (16.1)$$

und somit

$$f^{-1}(b+h) - f^{-1}(b) - Df(a)^{-1}(h) = r(h) - Df(a)^{-1}(h) = -Df(a)^{-1}(q(h))\|r(h)\| = \psi(h)\|r(h)\|.$$

Dabei ist $\psi(h) = -Df(a)^{-1}(q(h))$ stetig in 0 mit $\psi(0) = 0$. Damit erhalten wir für $h \neq 0$:

$$\frac{\|f^{-1}(b+h) - f^{-1}(b) - (Df(a))^{-1}(h)\|}{\|h\|} = \psi(h) \frac{\|r(h)\|}{\|h\|}.$$

Für den Beweis des Satzes genügt es nun die Beschränktheit des Quotienten $\frac{\|r(h)\|}{\|h\|}$ zu zeigen. Aus (16.1) folgt

$$\|r(h)\| = \|Df(a)^{-1}(h) + \psi(h)\|r(h)\| \leq \|Df(a)^{-1}(h)\| + \|\psi(h)\|\|r(h)\|.$$

Da $(Df(a))^{-1} \in \mathcal{L}(Y, X)$, existiert eine Konstante $c > 0$ mit $\|Df(a)^{-1}h\| \leq c\|h\|$. Dies impliziert:

$$\|r(h)\| \leq c\|h\| + \|\psi(h)\|\|r(h)\|$$

Da $\psi(h)$ stetig in 0 ist mit $\psi(0) = 0$, existiert ein $\delta > 0$ mit

$$\|\psi(h)\| \leq \frac{1}{2}$$

für alle $\|h\| \leq \delta$. Somit erhalten wir die Abschätzung

$$\|r(h)\| \leq 2c\|h\|$$

für $\|h\| \leq \delta$. Damit erhalten wir:

$$\lim_{h \rightarrow 0} \frac{\|f^{-1}(b+h) - f^{-1}(b) - (Df(a))^{-1}(h)\|}{\|h\|} \leq \lim_{h \rightarrow 0} \psi(h) \frac{\|r(h)\|}{\|h\|} = 0.$$

□

□

Nun wollen wir den lokalen Umkehrsatz mit Hilfe des Banachschen Fixpunktsatzes beweisen.

Satz 16.2.2 (Inverse Funktionen - Theorem). *Seien X, Y Banachräume, $W \subset X$ offen und $f : W \rightarrow Y$ stetig differenzierbar. Für $a \in W$ sei die lineare Abbildung $Df(a)$ invertierbar, und $Df(a)^{-1} \in \mathcal{L}(Y, X)$. Dann existiert eine offene Umgebung $U \subset W$ von a , so dass gilt:*

- (1) $f : U \rightarrow f(U) =: V$ ist bijektiv.
- (2) V ist offen.
- (3) $f^{-1} : V \rightarrow U$ ist stetig differenzierbar.

Beweis.

- (a) Wir dürfen annehmen, dass $X = Y$ und $Df(a) = \text{id}$ gilt, denn sonst betrachte die Abbildung $F : W \rightarrow X$ mit $F(x) = (Df(a))^{-1}(f(x))$. Die Kettenregel impliziert dann $DF(a) = (Df(a))^{-1}Df(a) = \text{id}$. Falls F die Eigenschaften (1), (2), (3) erfüllt, so auch f , da $Df(a)$ ein Banachraum-Isomorphismus ist.
- (b) Desweiteren nehmen wir $a = 0$ und $f(a) = 0$ an, denn sonst betrachte die Abbildung

$$F(x) = f(a + x) - f(a).$$

Falls F die Eigenschaften (1), (2), (3) erfüllt, so auch f .

Sei also $X = Y, f(0) = 0$ und $Df(0) = \text{id}$.

- (i) Wir wollen zunächst für y "nahe" bei 0 die Gleichung $f(x) = y$ nach x auflösen, d.h. für jedes gegebene $y \in X$ nahe bei 0 wollen wir ein x finden mit $f(x) = y$. Die Lösungen dieser Gleichung sind die Fixpunkte der Abbildung $\phi_y : W \rightarrow X$ mit

$$\phi_y(x) := y + x - f(x).$$

Da $D\phi_y(x) = \text{id} - Df(x)$ stetig von x abhängt und $\text{id} - Df(0) = 0$ ist, existiert ein $r > 0$ mit $K(0, 2r) \subset U$ und

$$\|\text{id} - Df(x)\| \leq \frac{1}{2} \tag{16.2}$$

für $x \in K(0, 2r)$. Daraus folgt für alle $x \in K(0, 2r)$: Die lineare Abbildung $Df(x) \in \mathcal{L}(X, X)$ ist invertierbar mit stetiger Inversen $Df(x)^{-1}$. Denn ist $A \in \mathcal{L}(X, X)$ und $\|A\| < 1$, so ist wegen Aufgabenblatt 5 Aufgabe 4 die Abbildung $\text{id} - A$ invertierbar.

Die Inverse ist stetig und durch die Reihe $\sum_{k=1}^{\infty} A^k$ gegeben. Aus

$$Df(x) = \text{id} - (\text{id} - Df(x))$$

folgt somit die Existenz und Stetigkeit von $Df(x)^{-1}$. Der Mittelwertsatz, angewandt auf ϕ_y , liefert wegen (16.2) die Abschätzung

$$\|\phi_y(x_1) - \phi_y(x_2)\| \leq \frac{1}{2}\|x_1 - x_2\| \tag{16.3}$$

für alle $x_1, x_2 \in K(0, 2r)$. Da $\phi_y(0) = y$, folgt somit für $\|y\| < r$ und $\|x\| \leq 2r$

$$\|\phi_y(x)\| \leq \|\phi_y(x) - \phi_y(0)\| + \|y\| < \frac{1}{2}\|x\| + r = 2r, \quad (16.4)$$

d.h. für jedes $y \in B(0, r)$ ist

$$\phi_y : K(0, 2r) \rightarrow K(0, 2r)$$

eine Kontraktion. Da $K(0, 2r)$ als abgeschlossene Teilmenge von X vollständig ist, existiert für jedes $y \in B(0, r)$ ein $x \in K(0, 2r)$ mit $\phi_y(x) = x$. Wegen (16.4) ist $\|x\| < 2r$, d.h. $x \in B(0, 2r)$. Daher existiert zu jedem $y \in B(0, r)$ genau ein $x = x(y) \in B(0, 2r)$ mit $f(x) = y$.

Setze nun $V = B(0, r)$ und $U = f^{-1}(V) \cap B(0, 2r)$. Dann ist $U \subset X$ wegen der Stetigkeit von f offen und $f : U \rightarrow V$ ist bijektiv ($f(B(0, 2r)) \subset B(0, r)$ muss natürlich nicht gelten).

- (ii) Wir zeigen nun, dass $f^{-1} : V \rightarrow U$ stetig ist. Seien $y_1, y_2 \in V$ und $x_1 = f^{-1}(y_1), x_2 = f^{-1}(y_2)$. Dann folgt wegen (16.3) mit Hilfe der Dreiecksungleichung

$$\begin{aligned} \|x_1 - x_2\| &= \|\phi_0(x_1) - \phi_0(x_2) + f(x_1) - f(x_2)\| \leq \\ &\frac{1}{2}\|x_1 - x_2\| + \|f(x_2) - f(x_1)\| \end{aligned}$$

und damit

$$\|f^{-1}(y_1) - f^{-1}(y_2)\| = \|x_1 - x_2\| \leq 2\|f(x_2) - f(x_1)\| = 2\|y_1 - y_2\|.$$

Also ist f^{-1} auf V Lipschitz-stetig und somit insbesondere stetig.

- (iii) Aus Satz 2.1 folgt aus der Bijektivität von f und der Existenz der stetigen Inversen $Df(x)^{-1}$ schließlich die Differenzierbarkeit von f^{-1} .

□

Definition 16.2.3. Seien X, Y Banachräume und $U \subset X, V \subset Y$ offen. Eine bijektive Abbildung $f : U \rightarrow V$ heißt *Diffeomorphismus*, falls f und f^{-1} stetig differenzierbare Abbildungen sind.

Bemerkung. Wir haben also gezeigt: Ist $f : W \rightarrow Y$ für $W \subset X$ offen eine C^1 -Abbildung und $Df(a) \in \mathcal{L}(X, Y)$ für ein $a \in W$ invertierbar mit stetiger Inverser, so existieren Umgebungen U von a und V von $f(a)$, so dass $f : U \rightarrow V$ ein Diffeomorphismus ist. Man sagt auch, dass $f : W \rightarrow Y$ in a ein *lokaler Diffeomorphismus* ist.

Übung: Betrachte die Abbildung $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ mit

$$f(x_1, x_2, x_3) = (x_1 + x_2 + x_3, x_2x_3 + x_3x_1 + x_1x_2, x_1 \cdot x_2 \cdot x_3)$$

(die Komponenten heißen elementarsymmetrische Funktionen aus \mathbb{R}^2). Zeige:

$$\det(Df(x)) = (x_1 - x_2)(x_1 - x_3)(x_2 - x_3).$$

Sei $a = (a_1, a_2, a_3) \in \mathbb{R}^3$ mit a_i paarweise verschieden, so lässt sich lokal um a das Gleichungssystem

$$\begin{aligned} y_1 &= x_1 + x_2 + x_3 \\ y_2 &= x_2x_3 + x_3x_1 + x_1x_2 \\ y_3 &= x_1x_2x_3 \end{aligned}$$

lösen, d.h. es existiert eine Umgebung V von $f(a)$ und eine Umgebung U von a , so dass für jedes $y \in V$ genau eine Lösung $x \in U$ existiert. Diese Lösungen hängen differenzierbar von y ab.

16.3 Implizite Abbildungen und das Lösen von nicht linearen Gleichungen.

Wir beginnen mit einem einfachen Beispiel, das die Situation erläutern soll. Betrachte die Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ mit $f(x, y) = x^2 + y^2 - 1$ und die nicht lineare Gleichung

$$f(x, y) = x^2 + y^2 - 1 = 0. \quad (16.5)$$

Diese Menge beschreibt den Kreis K mit Mittelpunkt $(0, 0) \in \mathbb{R}^2$ und Radius 1.

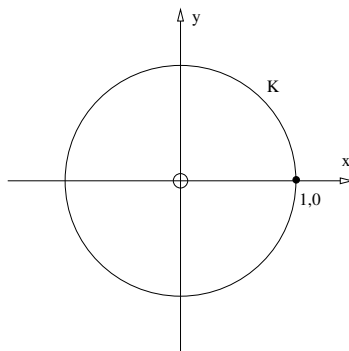


Abbildung 16.1: Kreis

Betrachte einen festen Punkt $(x_0, y_0) \in K$. Wann können wir die Gleichung (16.5) in der Nähe von (x_0, y_0) nach y auflösen, d.h. wann existiert eine Umgebung $V = U_1 \times U_2$ und eine Funktion $g : U_1 \rightarrow U_2$ mit $f(x, g(x)) = 0$, so dass der Graph von g mit $V \cap K$ übereinstimmt? Aus dem Bild ist offensichtlich, dass dies für $(1, 0)$ nicht möglich ist.

Wir wissen, dass $\text{grad } f(x, y) = (\frac{\partial f}{\partial x}(x, y), \frac{\partial f}{\partial y}(x, y))$ für alle $(x, y) \in K$ senkrecht auf K steht. Insbesondere ist die y -Komponente $\frac{\partial f}{\partial y}(x, y)$ des Gradienten an der Stelle $(1, 0)$ gleich null. Es ist aus obigem Bild anschaulich klar, dass es eine Funktion g geben muss, die die Gleichung (16.5) um (x_0, y_0) lokal löst, falls $\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$.

Analog können wir das Gleichungssystem nach x auflösen, falls $\frac{\partial f}{\partial x}(x_0, y_0) \neq 0$ ist.

Im allgemeinen werden unsere Variablen x, y Vektoren sein. Wir betrachten also folgende Situation: Es seien X, Y, Z Banachräume, $U \subset X \times Y$ offen und $f : U \rightarrow Z$ eine stetig differenzierbare Funktion. Sei $f(x_0, y_0) = 0$, so wollen wir das Gleichungssystem $f(x, y) = 0$

lokal nach y auflösen. Dies bedeutet, dass wir Umgebungen $U_1 \subset X$ von x_0 und $U_2 \subset Y$ von y_0 mit $U_1 \times U_2 \subset U$ sowie eine Funktion $g : U_1 \rightarrow U_2$ finden müssen mit $f(x, g(x)) = 0$ und

$$f^{-1}(0) \cap U_1 \times U_2 = \{(x, g(x)) \mid x \in U_1\} = \text{Graph}(g).$$

Betrachte zunächst das Differential

$$Df(x_0, y_0) : X \times Y \rightarrow Z$$

und schreibe es in der Form

$$Df(x_0, y_0)(h, k) = Df(x_0, y_0)(h, 0) + Df(x_0, y_0)(0, k).$$

Definiere die partiellen Differentiale $D_X f(x_0, y_0) : X \rightarrow Z$ und $D_Y f(x_0, y_0) : Y \rightarrow Z$ durch

$$D_X f(x_0, y_0)(h) := Df(x_0, y_0)(h, 0) \quad \text{und} \quad D_Y f(x_0, y_0)(k) := Df(x_0, y_0)(0, k).$$

Die Differentiale $D_X f(x_0, y_0)$ und $D_Y f(x_0, y_0)$ sind nichts anderes als die Ableitungen der partiellen Abbildungen $x \mapsto f(x, y_0)$ bzw. $y \mapsto f(x_0, y)$ im Punkte x_0 bzw. y_0 . Damit hat $Df(x_0, y_0)$ also die Darstellung

$$Df(x_0, y_0)(h, k) = D_X f(x_0, y_0)(h) + D_Y f(x_0, y_0)(k).$$

Ist $D_Y f(x_0, y_0)$ invertierbar, so ist $Df(x_0, y_0)$ surjektiv. Insbesondere kann man im Fall endlichdimensionaler Vektorräume \tilde{X}, Z zu $f : \tilde{X} \rightarrow Z$ für $Df(\tilde{x}_0)$ surjektiv auch umgekehrt einen Untervektorraum $Y \subset \tilde{X}$ mit $\dim Y = \dim Z$ so wählen, dass $D_Y(\tilde{x}_0)$ invertierbar ist.

Satz 16.3.1 (Satz über implizite Funktionen). *Es seien X, Y, Z Banachräume und $W \subset X \times Y$ offen. Sei $f : W \rightarrow Z$ eine C^1 -Abbildung und $f(x_0, y_0) = 0$. Ist das partielle Differential*

$$D_Y f(x_0, y_0) \in \mathcal{L}(Y, Z)$$

invertierbar, so existieren Umgebungen $U_1 \subset X$ von x_0 und $U_2 \subset Y$ von y_0 mit $U_1 \times U_2 \subset W$ und eine C^1 -Funktion $g : U_1 \rightarrow U_2$, so dass

$$f^{-1}(0) \cap (U_1 \times U_2) = \{(x, g(x)) \mid x \in U_1\}. \quad (16.6)$$

Bemerkung. Die Gleichung (16.6) lässt sich auch wie folgt interpretieren:

$$f(x, y) = 0 \quad \text{mit} \quad (x, y) \in U_1 \times U_2 \Leftrightarrow y = g(x) \quad \text{mit} \quad x \in U_1.$$

Man sagt: Die Abbildung g ist durch die Gleichung $f(x, y) = 0$ in der Umgebung von (x_0, y_0) implizit definiert.

Beweis. Der Beweis wird auf das Inverse Funktionen - Theorem zurückgeführt. Wir betrachten die Abbildung $\varphi : W \rightarrow X \times Z$ mit $\varphi(x, y) = (x, f(x, y))$. Das Differential $D\varphi(x_0, y_0) : X \times Y \rightarrow X \times Z$ ist dann eine bijektive lineare Abbildung mit stetiger Inverser. Denn

$$\begin{aligned} D\varphi(x_0, y_0)(h, k) &= (h, D_X f(x_0, y_0)(h) + D_Y f(x_0, y_0)(k)) \\ &= (h, D_Y f(x_0, y_0)(k + (D_Y f(x_0, y_0))^{-1} D_X f(x_0, y_0)(h))) \end{aligned}$$

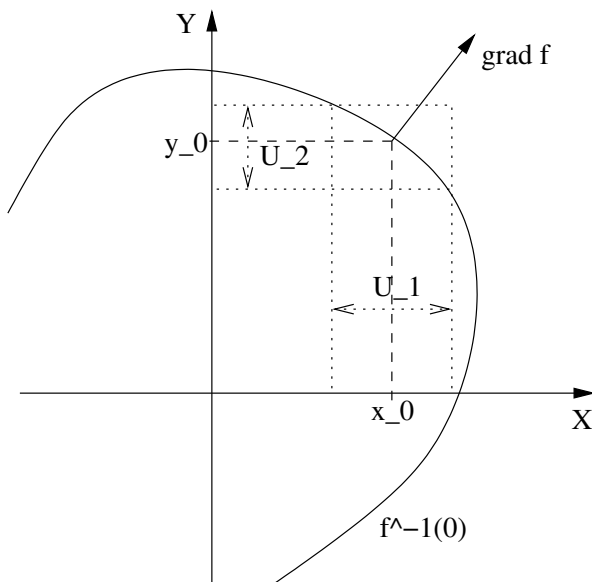


Abbildung 16.2: Implizit definierte Funktionen

für $h \in X, k \in Y$. Die Inverse $D\varphi(x_0, y_0)^{-1} : X \times Z \rightarrow X \times X$ ist somit durch

$$D\varphi(x_0, y_0)^{-1}(h', k') = (h', (D_Y f(x_0, y_0))^{-1}k' - (D_Y f(x_0, y_0))^{-1}D_X f(x_0, y_0)h')$$

gegeben. Diese Abbildung ist als Verknüpfung von stetigen Abbildungen stetig.

Wir wenden nun auf φ in (x_0, y_0) den Umkehrsatz an. Daher existiert eine offene Umgebung $U \subset W$ von (x_0, y_0) und V von $\varphi(x_0, y_0) = (x_0, 0)$, so dass $\varphi : U \rightarrow V$ ein Diffeomorphismus ist. Dabei ist $\varphi^{-1} : V \rightarrow U$ für $(\xi, \mu) \in V \subset X \times Z$ von der Form:

$$\varphi^{-1}(\xi, \mu) = (\xi, \psi(\xi, \mu)),$$

wobei $\psi : V \rightarrow Y$ eine C^1 -Abbildung ist. Für $(x, y) \in U$ lässt sich die Gleichung $f(x, y) = 0$ wie folgt umformen:

$$f(x, y) = 0 \Leftrightarrow \varphi(x, y) = (x, 0) \Leftrightarrow (x, y) = \varphi^{-1}(x, 0) \Leftrightarrow (x, y) = (x, \psi(x, 0)).$$

Insbesondere ist $\psi(x_0, 0) = y_0$. Wegen der Stetigkeit von ψ existiert eine Umgebung U_1 von x_0 und U_2 von y_0 , so dass $\psi(x, 0) \in U_2$ für alle $x \in U_1$. Definiere nun die Funktion $g : U_1 \rightarrow U_2$ durch $g(x) := \psi(x, 0)$. Damit ist g eine C^1 -Abbildung mit $f(x, g(x)) = 0$. \square

Bemerkung. Das Differential von g lässt sich leicht aus der Kettenregel berechnen, denn

$$f \circ (\text{id}_X, g) = \text{konst.} \quad \Rightarrow \quad D(f \circ (\text{id}_X, g)) = Df(x, g(x)) \circ (\text{id}_X, Dg(x)) = 0$$

für alle $x \in U_1$. Da

$$Df(x, g(x))(h, k) = D_X f(x, g(x))(h) + D_Y f(x, g(x))(k),$$

folgt

$$0 = Df(x, g(x))(h, Dg(x)(h)) = D_X f(x, g(x))h + D_Y f(x, g(x))Dg(x)h$$

und damit erhalten wir für $x = x_0$ und $y_0 = g(x_0)$:

$$Dg(x_0) = -D_Y f(x_0, y_0)^{-1}(D_X f(x_0, y_0)).$$

Beispiele.

(1) Fall einer Gleichung:

Sei $f : \mathbb{R}^n \rightarrow \mathbb{R}$ und $f(a_1, \dots, a_n) = 0$ für ein $(a_1, \dots, a_n) \in \mathbb{R}^n$. Ist

$$\frac{\partial f}{\partial x_n}(a_1, \dots, a_n) \neq 0,$$

so existiert eine Umgebung $U_1 \subset \mathbb{R}^{n-1}$ von (a_1, \dots, a_{n-1}) und eine Umgebung $U_2 \subset \mathbb{R}$ von a_n , so dass $g : U_1 \rightarrow U_2 \subset \mathbb{R}$ die Gleichung $f(x_1, \dots, x_n) = 0$ nach x_n auflöst, d.h.

$$f(x_1, \dots, x_{n-1}, g(x_1, \dots, x_{n-1})) = 0.$$

Insbesondere gilt:

$$Dg(a_1, \dots, a_{n-1}) = \frac{-1}{\frac{\partial f}{\partial x_n}(a_1, \dots, a_{n-1})} \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_{n-1}} \right).$$

(2) Gleichungssystem bestehend aus zwei Gleichungen $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$:

$$f(x, y_1, y_2) = (x^3 + y_1^3 + y_2^3 - 7, xy_1 + y_1y_2 + y_2x + 2)$$

$$f(2, -1, 0) = (8 - 1 - 7, -2 + 2) = 0$$

$$D_y f(x, y_1, y_2) = \begin{pmatrix} 3y_1^2 & 3y_2^2 \\ x + y_2 & y_1 + x \end{pmatrix}$$

$$D_y f(2, -1, 0) = \begin{pmatrix} 3 & 0 \\ 2 & 1 \end{pmatrix}$$

$$\det D_y f(2, -1, 0) = 3$$

$\Rightarrow D_y f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ist invertierbar. Aus dem implizite Funktionen - Theorem folgt damit die Existenz eines offenen Intervalls $I \ni 2$ und einer stetig differenzierbaren Funktion $g : I \rightarrow \mathbb{R}^2$, so dass $f(x, g_1(x), g_2(x)) = 0$.

16.4 Differenzierbare Untermannigfaltigkeiten des \mathbb{R}^n

Bisher haben wir die Differentialrechnung nur auf Vektorräumen betrieben. Dieser Rahmen ist aus vielen Gründen viel zu eng. Zum Beispiel ist es oft erforderlich, Extremwerte von Funktionen unter Nebenbedingungen zu bestimmen. Diese Nebenbedingungen sind oft durch Gleichungen bestimmt, deren Lösungsmenge keine Vektorraumstruktur, jedoch in vielen Fällen die Struktur einer Untermannigfaltigkeit hat.

Beispiel. Sei $\sum_{i=1}^n x_i^2 = 1$, hier ist die Lösungsmenge die Sphäre mit Radius 1 um den Ursprung des \mathbb{R}^n .

Definition 16.4.1. Eine Teilmenge $M^k \subset \mathbb{R}^n$ heißt k -dimensionale *differenzierbare Untermannigfaltigkeit* des \mathbb{R}^n , falls für jedes $p \in M^k$ eine offene Umgebung $W \subset \mathbb{R}^n$ von p und ein Diffeomorphismus $\varphi : W \rightarrow V \subset \mathbb{R}^n$ existiert mit

$$\varphi(W \cap M^k) = (\mathbb{R}^k \times \{0\}) \cap V$$

wobei $0 \in \mathbb{R}^{n-k}$. Die Einschränkung von φ auf $W \cap M^k$ heißt auch *Karte*.

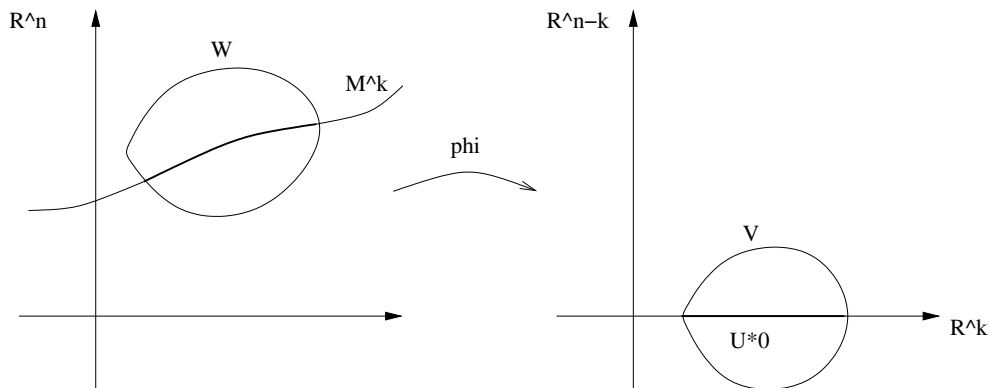


Abbildung 16.3: Untermannigfaltigkeit

Bemerkungen.

- Diese Definition lässt sich auch wie folgt interpretieren. Eine k -dimensionale Untermannigfaltigkeit des \mathbb{R}^n liegt lokal im \mathbb{R}^n wie der \mathbb{R}^k im \mathbb{R}^n .
- Die Einschränkung von φ auf $U \cap M^k$, $W \cap M^k \rightarrow \mathbb{R}^k \times \{0\} \cap V$, heißt auch die durch φ *induzierte Karte* von M^k . Die Menge $(\mathbb{R}^k \times \{0\}) \cap V$ ist von der Form $U \times \{0\}$, wobei U eine offene Teilmenge des \mathbb{R}^k ist. Daher können wir die Karte φ als Abbildung von $W \cap M^k$ nach U auffassen. Die Inverse einer Karte $\varphi : W \cap M^k \rightarrow U$ heißt auch *lokale Parametrisierung* von M^k .
- Seien $W_1, W_2 \subset \mathbb{R}^n$ zwei offene Umgebungen von $p \in M^k$ und seien $\varphi_1 : W_1 \rightarrow V_1 \subset \mathbb{R}^n$ sowie $\varphi_2 : W_2 \rightarrow V_2 \subset \mathbb{R}^n$ zwei Diffeomorphismen mit

$$\varphi_1(W_1 \cap M^k) = (\mathbb{R}^k \times \{0\}) \cap V_1 \quad \text{sowie} \quad \varphi_2(W_2 \cap M^k) = (\mathbb{R}^k \times \{0\}) \cap V_2,$$

so ist für $W = W_1 \cap W_2$

$$\varphi_2 \circ \varphi_1^{-1} : \varphi_1(W \cap M^k) \rightarrow \varphi_2(W \cap M^k)$$

ein Diffeomorphismus. Dieser Diffeomorphismus heißt auch *Kartenwechsel*. Bei der Definition “abstrakter Mannigfaltigkeiten” (d.h. solcher, die nicht von einem umgebenden Raum Gebrauch machen), ist die Differenzierbarkeit des Kartenwechsels Bestandteil der Definition (z.B. in der Relativitätstheorie).

Beispiele.

- (a) Graphen von stetig differenzierbaren Abbildungen sind Untermannigfaltigkeiten. Sei $U \subset \mathbb{R}^n$ offen und $f : U \rightarrow \mathbb{R}^k$ stetig differenzierbar. Dann ist der Graph von f

$$G_f = \{(x, f(x)) \mid x \in U\}$$

eine differenzierbare Untermannigfaltigkeit der Dimension n von \mathbb{R}^{n+k} . Denn $W := U \times \mathbb{R}^k \subset \mathbb{R}^n \times \mathbb{R}^k$ ist eine offene Teilmenge von \mathbb{R}^{n+k} und die Abbildung

$$\varphi : W \rightarrow W \quad \text{mit} \quad \varphi(x, y) = (x, f(x) - y).$$

ist stetig differenzierbar. Außerdem ist $\varphi : W \rightarrow W$ eine Involution, d.h. es gilt $\varphi \circ \varphi(x, y) = \varphi(x, f(x) - y) = (x, y)$ und somit ist $\varphi^{-1} = \varphi$ ebenfalls stetig differenzierbar und somit ein Diffeomorphismus. Ist insbesondere $(x, y) \in W \cap G_f$, d.h. $y = f(x)$, so folgt $\varphi(x, y) = (x, 0) \in \mathbb{R}^n \times \mathbb{R}^k$.

- (b)

$$S^n = \left\{ x \in \mathbb{R}^{n+1} \mid \sum_{i=1}^{n+1} x_i^2 = \|x\|^2 = 1 \right\}$$

ist eine n -dimensionale Untermannigfaltigkeit des \mathbb{R}^{n+1} . Seien $N = (0, \dots, 0, 1)$ der Nordpol und $S = (0, \dots, 0, -1)$ der Südpol von S^n . Betrachte die Abbildung $\varphi_N : \mathbb{R}^{n+1} \setminus \{N\} \rightarrow \mathbb{R}^{n+1} \setminus \{N\}$ gegeben durch

$$\varphi_N(x) = N + \frac{2}{\langle x - N, x - N \rangle} (x - N)$$

Diese Abbildung ist offensichtlich stetig differenzierbar. Diese Abbildung ist auch eine Involution, denn

$$\varphi_N \circ \varphi_N(x) = N + \frac{2}{\left\langle \frac{2(x-N)}{\|x-N\|^2}, \frac{2(x-N)}{\|x-N\|^2} \right\rangle} \left(\frac{2(x-N)}{\|x-N\|^2} \right) = x.$$

Damit ist insbesondere $\varphi_N : \mathbb{R}^{n+1} \setminus \{N\} \rightarrow \mathbb{R}^{n+1} \setminus \{N\}$ ein Diffeomorphismus. Außerdem ist $\varphi_N(x) \in \mathbb{R}^n \times \{0\}$, falls $x \in S^n$. Dazu genügt zu zeigen: $\langle \varphi_N(x), N \rangle = 0$. Dies folgt aus

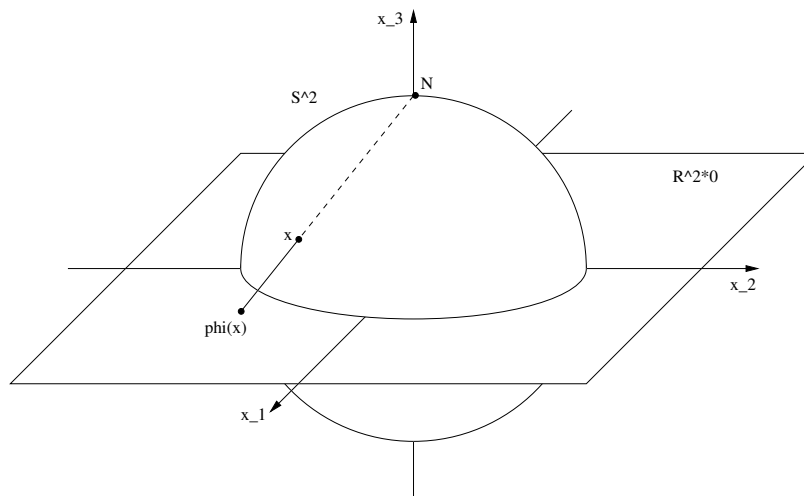
$$\begin{aligned} \langle \varphi_N(x), N \rangle &= \langle N, N \rangle + \frac{2\langle x - N, N \rangle}{\langle x - N, x - N \rangle} \\ &= 1 + \frac{2\langle x, N \rangle - 2}{\|x\|^2 - 2\langle x, N \rangle + \|N\|^2} = 0 \end{aligned}$$

. Genauso ist die Abbildung $\varphi_S : \mathbb{R}^{n+1} \setminus \{S\} \rightarrow \mathbb{R}^{n+1} \setminus \{S\}$ mit

$$\varphi_S(x) = S + \frac{2}{\langle x - S, x - S \rangle} (x - S)$$

eine Involution mit $\varphi_S(S^n) \subset \mathbb{R}^n \times \{0\}$. Da S^n in der Vereinigung der offenen Mengen $\mathbb{R}^{n+1} \setminus \{N\}$ und $\mathbb{R}^{n+1} \setminus \{S\}$ enthalten ist, ist S^n eine n -dimensionale Untermannigfaltigkeit von \mathbb{R}^{n+1} .

Die zu φ_N bzw. φ_S zugehörigen Karten $\varphi_N : S^n \setminus \{N\} \rightarrow \mathbb{R}^n \times \{0\}$ bzw. $\varphi_S : S^n \setminus \{S\} \rightarrow \mathbb{R}^n \times \{0\}$ werden *stereographische Projektionen* von Nordpol $N = (0, \dots, 0, 1)$ bzw. Südpol $S = (0, \dots, 0, -1)$ genannt.

Abbildung 16.4: Stereographische Projektion zu N

Oft sind Untermannigfaltigkeiten durch Gleichungssysteme gegeben.

Satz 16.4.2. Sei $f : U \rightarrow \mathbb{R}^k$ stetig differenzierbar und $U \subset \mathbb{R}^n$ offen. Sei $a \in \mathbb{R}^k$ ein regulärer Wert von f , d.h. $Df(x)$ sei surjektiv für alle $x \in f^{-1}(a)$. Dann ist $f^{-1}(a)$ eine $(n - k)$ -dimensionale Untermannigfaltigkeit von \mathbb{R}^n .

Bemerkung. $n \geq k$ folgt aus der Surjektivität von $Df(x)$.

Beweis. Setze $u = (x, y)$ mit $x \in \mathbb{R}^{n-k}$, $y \in \mathbb{R}^k$. Es sei $(x_0, y_0) \in f^{-1}(a)$. Nach eventueller Permutation der Variablen nehme an, dass

$$D_Y f(x_0, y_0) = \left(\frac{\partial f}{\partial y_1}(x_0, y_0), \dots, \frac{\partial f}{\partial y_k}(x_0, y_0) \right)$$

invertierbar ist. Nach dem Implizite Funktionen - Satz können wir die Gleichung $f(x, y) = a$ lokal nach y auflösen. Dies heißt, dass offene Umgebungen V_1 von $x_0 \in \mathbb{R}^{n-k}$ und V_2 von $y_0 \in \mathbb{R}^k$ und eine Abbildung $g : V_1 \rightarrow V_2$ existieren mit

$$f^{-1}(a) \cap (V_1 \times V_2) = \{(x, g(x)) \mid x \in V_1\} = G_g.$$

Wir haben also damit gezeigt, dass zu jedem $(x_0, y_0) \in f^{-1}(a)$ eine offene Umgebung W des \mathbb{R}^n der Form $W = V_1 \times V_2 \subset \mathbb{R}^{n-k} \times \mathbb{R}^k$ existiert, wobei $f^{-1}(a) \cap W$ Graph einer stetig differenzierbaren Funktion $\varphi : V_1 \rightarrow \mathbb{R}^k$ ist. Wegen des obigen Beispiels (a) ist somit $f^{-1}(a)$ eine $(n - k)$ - dimensionale Untermannigfaltigkeit des \mathbb{R}^n . \square

Beispiele.

(a) Betrachte $f : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ mit $f(x_1, \dots, x_{n+1}) = \sum_{i=1}^{n+1} x_i^2$. Insbesondere ist $f^{-1}(1) = S^n$.

Da $Df(x) = 2(x_1, \dots, x_n) \neq 0$ für alle $x \in f^{-1}(1)$ ist $Df(x) : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ surjektiv und damit $f^{-1}(1)$ eine n -dimensionale Untermannigfaltigkeit von \mathbb{R}^{n+1} .

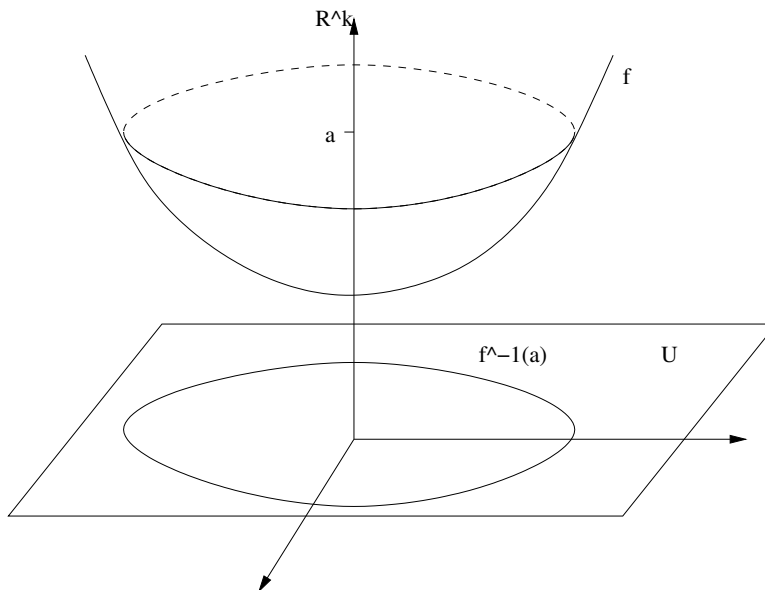


Abbildung 16.5: Das Urbild eines regulären Wertes ist eine Untermannigfaltigkeit

(b) Betrachte den Vektorraum

$$M(n, \mathbb{R}) = \{A \mid A \text{ reelle } n \times n \text{ Matrix}\} \cong \mathbb{R}^n \times \dots \times \mathbb{R}^n = \mathbb{R}^{n^2}$$

der reellen $n \times n$ Matrizen.

- (1) $SL(n, \mathbb{R}) = \{A \in M(n, \mathbb{R}) \mid \det A = 1\}$ heißt spezielle lineare Gruppe (vgl. 11.2.5). Sie ist eine $n^2 - 1$ -dimensionale Untermannigfaltigkeit von \mathbb{R}^{n^2} , denn $\det : \mathbb{R}^{n^2} \rightarrow \mathbb{R}$ ist stetig differenzierbar und

$$D \det(a_1, \dots, a_n)(h_1, \dots, h_n) = \det(h_1, a_2, \dots, a_n) + \det(a_1, h_2, a_3, \dots, a_n) + \dots + \det(a_1, \dots, a_{n-1}, h_n).$$

Damit ist $D \det(a_1, \dots, a_n) : \mathbb{R}^{n^2} \rightarrow \mathbb{R}$ surjektiv für alle $A = (a_1, \dots, a_n) \in \det^{-1}(1)$, denn z.B. ist

$$D \det(a_1, \dots, a_n)(a_1, 0, \dots, 0) = \det(a_1, \dots, a_n) = 1 \neq 0.$$

$SL(n, \mathbb{R})$ ist eine Untermannigfaltigkeit mit einer Gruppenstruktur. Solche Mannigfaltigkeiten heißen *Liegruppen*.

- (2) $O(n, \mathbb{R}) = \{A \in M(n, \mathbb{R}) \mid A A^t = \text{id}\}$ heißt orthogonale Gruppe (vgl. 13.5.13). Betrachte den Vektorraum

$$\text{Sym}(n, \mathbb{R}) = \{B \in M(n, \mathbb{R}) \mid B = B^t\}$$

der *symmetrischen Matrizen*. Dann gilt $\dim \text{Sym}(n, \mathbb{R}) = \frac{n}{2}(n+1)$ und somit kann man $\text{Sym}(n, \mathbb{R})$ mit $\mathbb{R}^{\frac{n}{2}(n+1)}$ identifizieren. Betrachte die Abbildung

$$F : M(n, \mathbb{R}) \rightarrow \text{Sym}(n, \mathbb{R})$$

mit $F(A) = AA^t$. Dann ist $F^{-1}(\text{id}) = O(n, \mathbb{R})$. Da die Abbildung

$$G : M(n, \mathbb{R}) \times M(n, \mathbb{R}) \rightarrow M(n, \mathbb{R})$$

mit $G(A, B) = AB^t$ bilinear – und wegen endlicher Dimension auch stetig – ist, ist nach Satz 15.4.4 G und somit auch F stetig differenzierbar, wobei die Ableitung $DF(A) : M(n, \mathbb{R}) \rightarrow \text{Sym}(n, \mathbb{R})$ durch

$$DF(A)H = HA^t + AH^t$$

gegeben ist. Das Differential $DF(A)$ ist für alle $A \in F^{-1}(\text{id})$ surjektiv, denn für jedes $S \in \text{Sym}(n, \mathbb{R})$ ist $H = \frac{1}{2}SA$ wegen

$$DF(A)H = HA^t + AH^t = \frac{1}{2}S + \frac{1}{2}S^t = S$$

ein Urbild von S unter $DF(A)$. Es gilt:

$$\dim O(n, \mathbb{R}) = \dim M(n, \mathbb{R}) - \dim \text{Sym}(n, \mathbb{R}) = n^2 - \frac{n^2}{2} - \frac{n}{2} = \frac{n^2}{2} - \frac{n}{2} = \frac{n}{2}(n-1).$$

Da $O(n, \mathbb{R})$ als Teilmenge von $\mathbb{R}^{n^2} = M(n, \mathbb{R})$ abgeschlossen und beschränkt ist (Satz v. Heine-Borel), ist $O(n, \mathbb{R})$ im Gegensatz zu $SL(n, \mathbb{R})$ kompakt. Diese Menge ist nicht zusammenhängend. Sie besteht aus zwei Zusammenhangskomponenten, nämlich solchen Matrizen mit Determinante 1 und solchen mit Determinante -1 . Denn wäre $O(n, \mathbb{R})$ wegzusammenhängend, so gäbe es eine stetige Kurve $c : [0, 1] \rightarrow O(n, \mathbb{R})$ mit $\det c(0) = 1, \det c(1) = -1$. Dann ist aber $\det \circ c : [0, 1] \rightarrow \mathbb{R}$ stetig. Dies widerspricht dem Zwischenwertsatz.

Eine weitere Möglichkeit besteht darin, Untermannigfaltigkeiten mittels Parametrisierungen darzustellen.

Definition 16.4.3. Sei $U \subset \mathbb{R}^k$ offen und $f : U \rightarrow \mathbb{R}^n$ eine stetig differenzierbare Abbildung. Dann heißt f *Immersion (Parametrisierung)*, falls $Df(x) \in \mathcal{L}(\mathbb{R}^k, \mathbb{R}^n)$ für alle $x \in U$ injektiv ist.

Bemerkungen.

- (a) Da $Df(x)$ injektiv ist, folgt $k \leq n$.
- (b) Unter $f(U)$ stellen wir uns ein k -dimensionales Flächenstück vor. Diese Flächenstücke können Selbstdurchdringungen aufweisen.
- (c) Ist $M^k \subset \mathbb{R}^n$ eine k -dimensionale Untermannigfaltigkeit des \mathbb{R}^n , so existiert nach Definition für jedes $p \in M$ eine offene Umgebung $W \subset \mathbb{R}^n$ von p und ein Diffeomorphismus $\varphi : W \rightarrow V \subset \mathbb{R}^n$ mit

$$\varphi(W \cap M^k) = (\mathbb{R}^k \times \{0\}) \cap V := U \times \{0\}$$

Die Umkehrung $\psi : U \rightarrow W \cap M^k$ mit $\psi(x) = \varphi^{-1}(x, 0)$ ist eine Immersion. Sie stellt also eine lokale Parametrisierung von M^k dar.

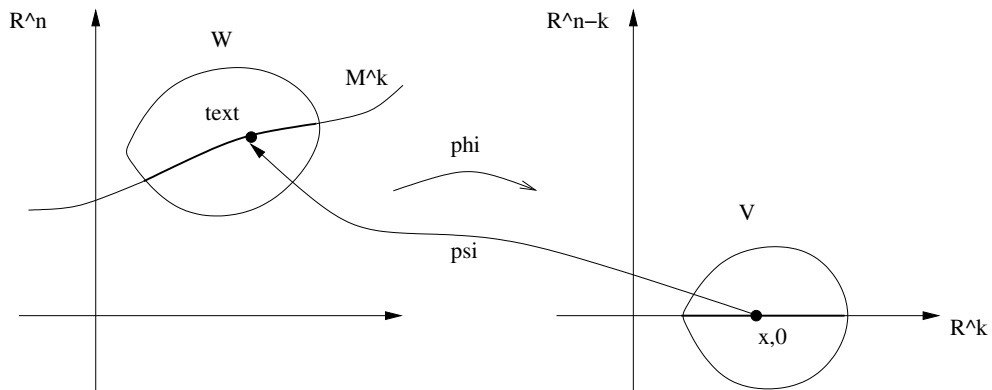


Abbildung 16.6: Immersion

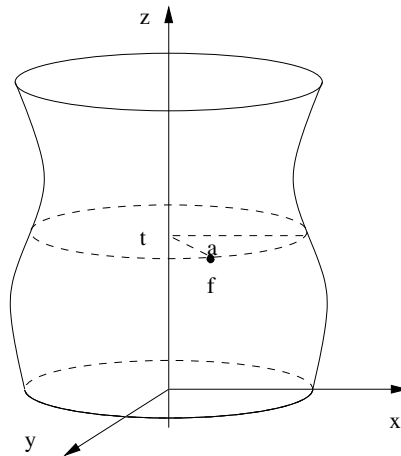


Abbildung 16.7: Rotationsfläche

Beispiele.

(a) *Rotationsflächen*

Betrachte die stetig differenzierbare Kurve $c : (a, b) \rightarrow \mathbb{R}^3$ mit $c(t) = (x(t), 0, z(t))$ in der (x, z) -Ebene, wobei $x(t) > 0$ und $x'(t)^2 + z'(t)^2 > 0$. Die *Rotationsfläche* mit Rotationsachse z ist dann durch

$$f(t, \alpha) = (x(t) \cos \alpha, x(t) \sin \alpha, z(t))$$

gegeben. Das Differential an der Stelle (t, α) ist

$$Df(t, \alpha) = \begin{pmatrix} x'(t) \cos \alpha & -x(t) \sin \alpha \\ x'(t) \sin \alpha & x(t) \cos \alpha \\ z'(t) & 0 \end{pmatrix}.$$

Da die Spaltenvektoren senkrecht aufeinander stehen und von null verschieden sind, sind sie linear unabhängig. Damit ist $Df(t, \alpha)$ injektiv.

(b) *Graphen*

Sei $U \subset \mathbb{R}^2$ offen und $g : U \rightarrow \mathbb{R}$ stetig differenzierbar. Betrachte $f : U \rightarrow \mathbb{R}^3$ mit $f(u, v) = (u, v, g(u, v))$.

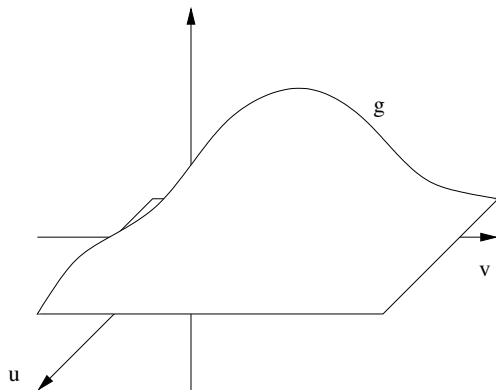


Abbildung 16.8: Graph

Dann gilt:

$$Df(u, v) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ \frac{\partial g}{\partial u}(u, v) & \frac{\partial g}{\partial v}(u, v) \end{pmatrix}$$

Die Spaltenvektoren sind linear unabhängig.

Satz 16.4.4. Sei $U \subset \mathbb{R}^k$ offen und $f : U \rightarrow \mathbb{R}^n$ eine Immersion. Dann existiert zu jedem $x_0 \in U$ eine offene Umgebung $U_0 \subset U$ von x_0 , so dass $f(U_0) \subset \mathbb{R}^n$ eine k -dimensionale Untermannigfaltigkeit des \mathbb{R}^n ist.

Beweis. Da

$$Df(x_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x_0) & \dots & \frac{\partial f_1}{\partial x_k}(x_0) \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1}(x_0) & \dots & \frac{\partial f_n}{\partial x_k}(x_0) \end{pmatrix}$$

injektiv ist, sind k -Zeilenvektoren linear unabhängig. Nehme an, dass die ersten k -Zeilen linear unabhängig sind, d.h. dass die Matrix

$$\left(\frac{\partial f_i}{\partial x_j}(x_0) \right)_{1 \leq i, j \leq k} =: A$$

invertierbar ist. Sonst betrachte die Permutation $P : \mathbb{R}^n \rightarrow \mathbb{R}^n$ der Koordinaten, so dass $\tilde{f} = P \circ f$ die verlangte Eigenschaft besitzt. Betrachte nun die Abbildung $F : U \times \mathbb{R}^{n-k} \rightarrow \mathbb{R}^n$ mit $F(x, y) = f(x) + (0, y)$. Dann gilt

$$DF(x_0, 0) = \begin{pmatrix} A & 0 \\ * & \text{id} \end{pmatrix} : \mathbb{R}^n \longrightarrow \mathbb{R}^n$$

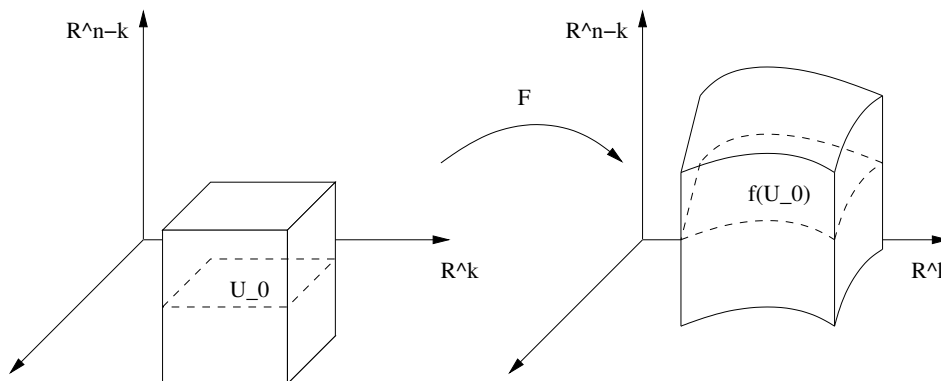


Abbildung 16.9: Zu Satz 16.4.4

ist invertierbar, da die Spalten linear unabhängig sind. Nach dem Inverse Funktionen - Theorem besitzt F eine lokale stetig differenzierbare Umkehrung, d.h. es gibt eine offene Umgebung

$$V = U_0 \times U' \subset U \times \mathbb{R}^{n-k}$$

von $(x_0, 0) \in \mathbb{R}^n$ und $W \subset \mathbb{R}^n$ eine offene Umgebung von $f(x_0) = F(x_0, 0)$, so dass $F : V \rightarrow W$ bijektiv ist mit der stetig differenzierbaren Inversen $\varphi := F^{-1} : W \rightarrow V$. Dann gilt: $f(U_0) \subset W$ und $\varphi(W \cap f(U_0)) =: \mathbb{R}^k \times \{0\} \cap V$ \square

Bemerkung. Ist $f : U \rightarrow \mathbb{R}^n$ eine Immersion, so muss $f(U) \subset \mathbb{R}^n$ selbst dann keine Untermannigfaltigkeit sein, falls $f : U \rightarrow f(U)$ bijektiv ist. Existiert jedoch eine *stetige* Umkehrung, so ist $f(U)$ eine Untermannigfaltigkeit von \mathbb{R}^n . In Aufgabe 2(b) von Aufgabenblatt 6 ist ein Beispiel einer stetigen bijektiven Abbildung diskutiert worden, deren Inverse nicht stetig ist.

Bijektive stetige Abbildungen, deren Inverse ebenfalls stetig sind, heißen Homöomorphismen. Allgemein definiert man

Definition 16.4.5. Seien (M_1, d_1) und (M_2, d_2) metrische Räume und $f : M_1 \rightarrow M_2$ eine stetige Abbildung mit stetiger Inverser $f^{-1} : M_2 \rightarrow M_1$. Dann heißt $f : M_1 \rightarrow M_2$ *Homöomorphismus*.

Bemerkung. Ist $f : M_1 \rightarrow M_2$ ein Homöomorphismus, so ist nicht nur das Urbild, sondern auch das Bild einer offenen Menge offen.

Ist A Teilmenge eines metrischen Raumes (M, d) , so ist (A, d) wieder ein metrischer Raum. Eine Teilmenge $U \subset A$ ist genau dann offen im metrischen Raum (A, d) , falls eine offene Teilmenge $W \subset M$ existiert mit $U = A \cap W$.

Definition 16.4.6. Eine Immersion $f : U \rightarrow \mathbb{R}^n$ heißt *Einbettung*, falls $f : U \rightarrow f(U)$ ein Homöomorphismus ist, d.h. $f : U \rightarrow f(U)$ ist bijektiv und $f : U \rightarrow f(U)$ sowie $f^{-1} : f(U) \rightarrow U$ sind stetig.

Satz 16.4.7. Ist $U \subset \mathbb{R}^k$ offen und $f : U \rightarrow \mathbb{R}^n$ eine Einbettung, so ist $f(U)$ eine k -dimensionale Untermannigfaltigkeit von \mathbb{R}^n .

Beweis. Da $f : U \rightarrow \mathbb{R}^n$ insbesondere eine Immersion ist, existiert zu jedem $x_0 \in U$ eine offene Umgebung $U_0 \subset U$ von x_0 , so dass $f(U_0) \subset \mathbb{R}^n$ eine k -dimensionale Untermannigfaltigkeit des \mathbb{R}^n ist. Daher existiert zu jedem $y \in f(U_0)$ eine offene Umgebung $W_1 \subset \mathbb{R}^n$ und ein Diffeomorphismus $\varphi : W_1 \rightarrow V$ auf eine offene Menge $V \subset \mathbb{R}^n$ mit

$$\varphi(W_1 \cap f(U_0)) = (\mathbb{R}^k \times \{0\}) \cap V,$$

wobei $0 \in \mathbb{R}^{n-k}$. Da aber f auch eine Einbettung ist, ist $f(U_0)$ offen als Teilmenge von $f(U)$. Daraus folgt (siehe obige Bemerkung) die Existenz einer offenen Menge $W_2 \subset \mathbb{R}^n$ mit $f(U_0) = f(U) \cap W_2$. Also gilt für $W := W_1 \cap W_2$:

$$\varphi(W \cap f(U)) = \varphi(W_1 \cap f(U_0)) = (\mathbb{R}^k \times \{0\}) \cap V$$

und somit ist auch $f(U) \subset \mathbb{R}^n$ eine Untermannigfaltigkeit. □

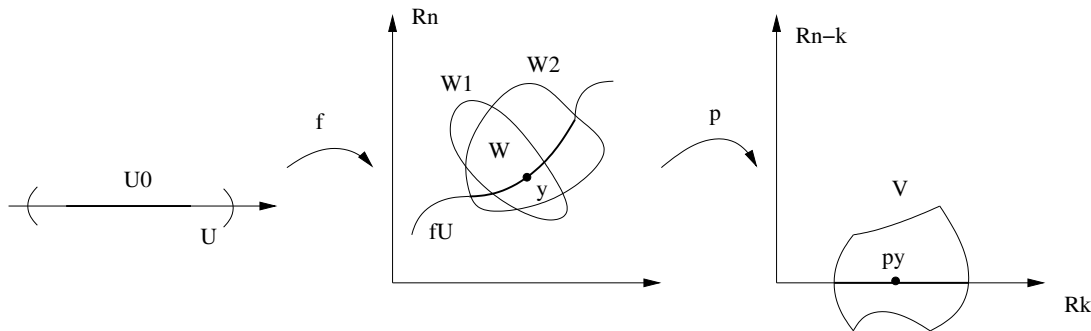


Abbildung 16.10: Zu Satz 16.4.7

16.5 Tangentialraum und Normalraum

Definition 16.5.1. (Tangentialraum)

Sei $M \subset \mathbb{R}^n$ eine k -dimensionale Untermannigfaltigkeit und $p \in M$. Dann heißt

$$T_p M = \{ \dot{c}(0) \mid c : (-\epsilon, +\epsilon) \rightarrow M \text{ differenzierbare Kurve mit } c(0) = p, \epsilon > 0 \}$$

Tangentialraum von $p \in M$. Die Elemente von $T_p M$ heißen *Tangentialvektoren*.

Satz 16.5.2. Sei $\psi : U \rightarrow \mathbb{R}^n$ eine lokale Parametrisierung der k -dimensionalen Untermannigfaltigkeit M mit $\psi(x) = p$. Dann ist $T_p M = D\psi(x)(\mathbb{R}^k)$. Insbesondere ist $T_p M$ ein k -dimensionaler linearer Untervektorraum von \mathbb{R}^n .

Beweis.

(i) Wir zeigen zunächst:

$$T_p M \subset D\psi(x)(\mathbb{R}^k).$$

Sei $c : (-\epsilon, +\epsilon) \rightarrow M$ eine differenzierbare Kurve mit $c(0) = p$, so dass $c(-\epsilon, +\epsilon) \subset \psi(U)$. Sei $\alpha = \psi^{-1} \circ c : (-\epsilon, +\epsilon) \rightarrow U$ die Urbildkurve von c unter ψ . Dann gilt: $c = \psi \circ \alpha$ und $\dot{c}(0) = D\psi(\alpha(0))\dot{\alpha}(0) = D\psi(x)\dot{\alpha}(0)$, wobei $\dot{\alpha}(0) \in \mathbb{R}^k$.

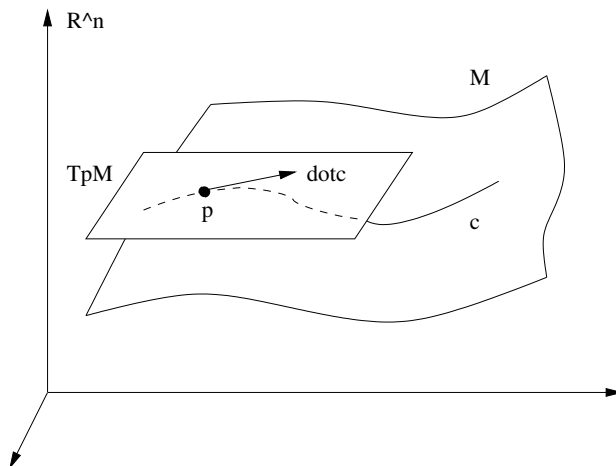


Abbildung 16.11: Tangentialraum

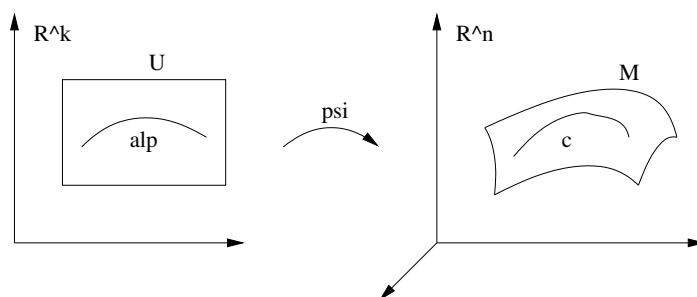


Abbildung 16.12: Zu Satz 16.5.2

(ii) Wir zeigen nun

$$D\psi(x)(\mathbb{R}^k) \subset T_pM.$$

Sei $w = D\psi(x)(v) \in D\psi(x)(\mathbb{R}^k)$, so betrachte die Kurve $\alpha : (-\epsilon, \epsilon) \rightarrow M$ mit $\alpha(t) = x + tv \in U$ für $t \in (-\epsilon, \epsilon)$. Sei $c(t) = \psi \circ \alpha(t)$, so folgt aus der Kettenregel: $\dot{c}(0) = D\psi(x)(v) \in T_pM$.

Da $D\psi(x) : \mathbb{R}^k \rightarrow \mathbb{R}^n$ eine injektive lineare Abbildung ist, ist somit $D\psi(\mathbb{R}^k) = T_pM$ ein k -dimensionaler linearer Untervektorraum von \mathbb{R}^n . □

Bemerkung. Ist $e_1, \dots, e_k \in \mathbb{R}^k$ die Standardbasis, so bilden die Vektoren

$$D\psi(x)(e_i) = \frac{\partial \psi}{\partial x_i}(x), \quad i \in \{1, \dots, k\}$$

eine Basis von T_pM .

Definition 16.5.3. Ist $M \subset \mathbb{R}^n$ eine Untermannigfaltigkeit und $p \in M$, so heißt

$$N_pM = \{v \in \mathbb{R}^n \mid \langle v, w \rangle = 0 \text{ für alle } w \in T_pM\} =: (T_pM)^\perp$$

der *Normalraum* von M in p . Dabei bezeichnet $\langle \cdot, \cdot \rangle$ das kanonische Skalarprodukt des \mathbb{R}^n .

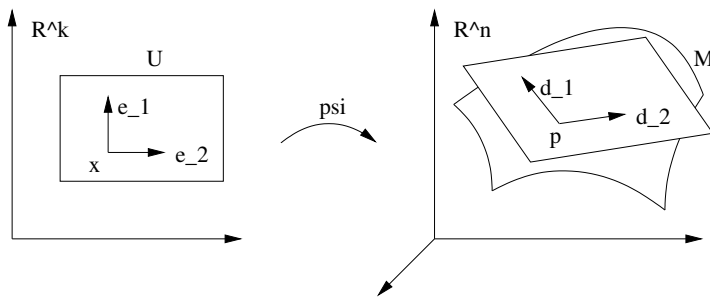


Abbildung 16.13: Basis des T_pM

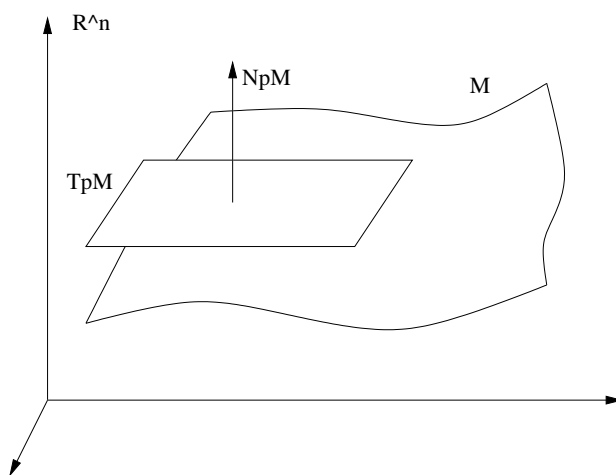


Abbildung 16.14: Normalenraum

Der Normalenraum lässt sich für Untermannigfaltigkeiten, die durch Gleichungssysteme gegeben sind, ohne Kenntnis des Tangentialraumes beschreiben.

Satz 16.5.4. Sei $U \subset \mathbb{R}^n$ offen, $f = (f_1, \dots, f_k) : U \rightarrow \mathbb{R}^k$ eine C^1 -Abbildung und $a \in \mathbb{R}^k$ ein regulärer Wert von f , d.h. für alle $p \in f^{-1}(a)$ ist das Differential $Df(p) : \mathbb{R}^n \rightarrow \mathbb{R}^k$ surjektiv. Dann sind die Vektoren

$$\text{grad } f_1(p), \dots, \text{grad } f_k(p)$$

linear unabhängig und bilden eine Basis des Normalraumes N_pM_a an die $(n - k)$ -dimensionale Untermannigfaltigkeit $M_a = f^{-1}(a)$ im Punkte p .

Beweis. Da $Df(p)$ surjektiv ist und

$$Df(p) = \begin{pmatrix} \text{grad } f_1(p) \\ \vdots \\ \text{grad } f_k(p) \end{pmatrix}$$

die Jacobimatrix darstellt, sind die Vektoren $\text{grad } f_1(p), \dots, \text{grad } f_k(p)$ linear unabhängig. Sei $c : (-\epsilon, \epsilon) \rightarrow M_a$ eine Kurve mit $c(0) = p$. Dann ist $f_i \circ c(t) = a_i$ die i -te Komponente von a

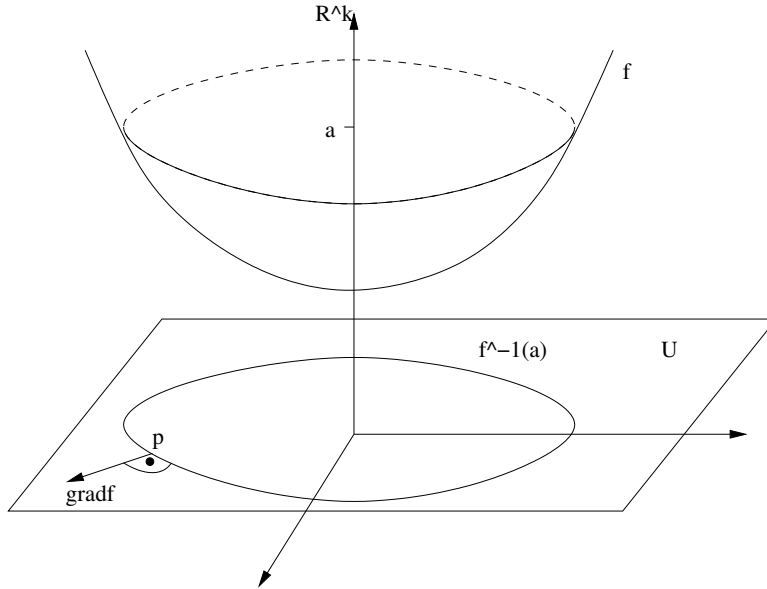


Abbildung 16.15: $\text{grad } f_i$ stehen senkrecht auf M_a .

und somit konstant. Dann gilt:

$$0 = \left. \frac{d}{dt} \right|_{t=0} f_i \circ c(t) = \langle \text{grad } f_i(p), \dot{c}(0) \rangle$$

und somit ist $\text{grad } f_i(p) \in N_p M_a$. Da $N_p M_a$ das orthogonale Komplement zu $T_p M_a$ ist, folgt

$$N_p M_a \oplus T_p M = \mathbb{R}^n$$

und somit $\dim N_p M_a = n - (n - k) = k$. Aus der linearen Unabhängigkeit der Vektoren $\text{grad } f_1(p), \dots, \text{grad } f_k(p)$ folgt somit die Behauptung. \square

16.6 Lokale Extrema und Lagrangesche Multiplikatoren

Genau wie bei Funktionen auf Vektorräumen lässt sich der Begriff des Extremums für Funktionen auf Untermannigfaltigkeiten erklären.

Definition 16.6.1. Sei $M \subset \mathbb{R}^n$ eine Untermannigfaltigkeit und $f : M \rightarrow \mathbb{R}^n$ eine Abbildung. Dann heißt f *differenzierbar* in $p \in M$, falls f eine differenzierbare Erweiterung in einer offenen Umgebung $U \subset \mathbb{R}^n$ von p besitzt, d.h. es existiert eine differenzierbare Abbildung $F : U \rightarrow \mathbb{R}$ mit $F(q) = f(q)$ für alle $q \in U \cap M$. Definiere das Differential $Df(p) : T_p M \rightarrow \mathbb{R}^n$ von f im Punkte p durch

$$Df(p)(v) = DF(p)(v)$$

für alle $v \in T_p M$.

Bemerkung. Die Abbildung $Df(p) : T_p M \rightarrow \mathbb{R}^n$ ist wohldefiniert, denn ist $v \in T_p M$ so existiert eine differenzierbare Kurve $c : (-\epsilon, +\epsilon) \rightarrow M$ mit $\dot{c}(0) = v$. Dann gilt:

$$DF(p)(v) = DF(p)(\dot{c}(0)) = \widehat{F \circ c}(0) = \lim_{t \rightarrow 0} \frac{f(c(t)) - f(c(0))}{t} = \widehat{f \circ c}(0).$$

Insbesondere hängt $Df(p)(v)$ nicht von der Wahl der Fortsetzung F ab. Wegen $Df(p) = DF(p)|_{T_p M}$ ist $Df : T_p M \rightarrow \mathbb{R}^n \in \mathcal{L}(T_p M, \mathbb{R}^n)$ eine lineare Abbildung.

Definition 16.6.2. Sei $M \subset \mathbb{R}^n$ eine Untermannigfaltigkeit und $f : M \rightarrow \mathbb{R}$ eine Abbildung. Dann heißt $p \in M$ *lokales Maximum (Minimum)*, falls eine offene Umgebung $U \subset \mathbb{R}^n$ von p existiert, so dass $f(p) \geq f(q)$ ($f(p) \leq f(q)$) für alle $q \in M \cap U$.

Bemerkung. Ein lokales Maximum bzw. Minimum p heißt auch *lokales Extremum (Extremalstelle)*.

Definition 16.6.3. Sei $f : M \rightarrow \mathbb{R}$ in $p \in M$ differenzierbar, so heißt p *kritischer Punkt*, falls $Df(p) = 0 \in \mathcal{L}(T_p M, \mathbb{R})$.

Bemerkung. Ist $F : U \rightarrow \mathbb{R}$ eine differenzierbare Erweiterung von f in einer Umgebung $U \subset \mathbb{R}^n$ von p , so ist $Df(p)v = DF(p)v = \langle \text{grad } F(p), v \rangle$ für alle $v \in T_p M$. Daher ist p genau dann ein kritischer Punkt, falls $\text{grad } F(p) \in N_p M$.

Satz 16.6.4. Ist $p \in M$ eine Extremalstelle von f , so ist p ein kritischer Punkt.

Beweis. Sei $c : (-\epsilon, \epsilon) \rightarrow M$ eine differenzierbare Kurve mit $c(0) = p$. Dann ist

$$0 = \left. \frac{d}{dt} \right|_{t=0} f \circ c(t) = Df(p)(\dot{c}(0)),$$

falls p eine Extremalstelle ist, denn 0 ist eine Extremalstelle von $t \rightarrow f \circ c(t)$. □

Bemerkung. Seien $U \subset \mathbb{R}^n$ offen, $f : U \rightarrow \mathbb{R}^n$ und $M \subset \mathbb{R}^n$ eine Untermannigfaltigkeit. Ist $p \in M \cap U$ eine Extremalstelle von $f : M \cap U \rightarrow \mathbb{R}$, so heißt p Extremalstelle von f unter der "Nebenbedingung" M . Diese Nebenbedingung ist oft durch ein Gleichungssystem gegeben.

Korollar 16.6.5. Seien $U \subset \mathbb{R}^n$ offen, $f : U \rightarrow \mathbb{R}$ und $g : U \rightarrow \mathbb{R}^k$ stetig differenzierbare Abbildungen und $a \in \mathbb{R}^k$ ein regulärer Wert von $g = (g_1, \dots, g_k)$. Sei $p \in g^{-1}(a)$ eine Extremalstelle von f unter der Nebenbedingung

$$M_a = g^{-1}(a) = \{x \in U \mid g(x) = a\}.$$

Dann ist $\text{grad } f(p) \in N_p M_a$ und es existieren Koeffizienten $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ mit

$$\text{grad } f(p) = \lambda_1 \text{grad } g_1(p) + \dots + \lambda_k \text{grad } g_k(p). \quad (16.7)$$

Die Koeffizienten $\lambda_1, \dots, \lambda_k$ heißen Lagrangesche Multiplikatoren.

Beweis. Ist p eine Extremalstelle von $f|_{M_a} : M_a \rightarrow \mathbb{R}$, so ist $Df|_{M_a}(p) : T_p M_a \rightarrow \mathbb{R}$ die Nullabbildung.

Da f eine differenzierbare Erweiterung von $f|_{M_a}$ ist, ist wegen obiger Bemerkung $\text{grad } f(p) \in N_p M_a$. Da die Vektoren $\text{grad } g_1(p), \dots, \text{grad } g_k(p)$ wegen Satz 16.5.4 den Normalraum aufspannen, folgt die Behauptung. □

Bemerkung. Ist (16.7) erfüllt, so muss p im Allgemeinen deshalb noch keine Extremalstelle sein. Um zu entscheiden, ob es sich um ein Maximum oder Minimum handelt, muss man analog zu Satz 15.7.5 die zweite Ableitung von f in p berechnen. Gilt für jede Kurve $c : (-\varepsilon, \varepsilon) \rightarrow M_a$ mit $c(0) = p, v := \dot{c}(0) \neq 0$

$$D^2 f(p)(v, v) = \left. \frac{d^2}{dt^2} \right|_{t=0} f \circ c(t) > 0 \quad \text{bzw.} \quad < 0,$$

so handelt es sich bei p um ein Minimum bzw. Maximum.

Beispiele.

(a) Betrachte die Abbildung $f : \mathbb{R}^n \rightarrow \mathbb{R}$ mit

$$f(x_1, \dots, x_n) = x_1 \cdot \dots \cdot x_n.$$

Sei $U = \{(x_1, \dots, x_n) \mid x_i > 0\}$ und $g : U \rightarrow \mathbb{R}$ die Abbildung mit

$$g(x_1, \dots, x_n) = x_1 + \dots + x_n.$$

Sei M die Untermannigfaltigkeit des \mathbb{R}^n mit

$$M = g^{-1}(1) = \{(x_1, \dots, x_n) \in U \mid x_1 + \dots + x_n = 1\}.$$

Dann ist der Abschluss \bar{M} kompakt und durch

$$\bar{M} = \{(x_1, \dots, x_n) \mid x_1 + \dots + x_n = 1, x_i \geq 0\}$$

gegeben. Daher besitzt $f|_{\bar{M}}$ ein globales Maximum. Dieses Maximum wird in M angenommen, denn ist $x \in \bar{M} \setminus M$, so ist $x_i = 0$ für wenigstens ein $i \in \{1, \dots, n\}$ und somit $f(x) = 0$. Nach dem obigen Korollar existiert dann ein $p = (p_1, \dots, p_n) \in M$ und $\lambda \in \mathbb{R}$ mit

$$\lambda \operatorname{grad} g(p) = \operatorname{grad} f(p)$$

oder in äquivalenter Formulierung:

$$\lambda \frac{\partial g}{\partial x_k}(p) = \lambda = \frac{\partial f}{\partial x_k}(p) = p_1 \cdot \dots \cdot p_{k-1} p_{k+1} \cdot \dots \cdot p_n.$$

Somit folgt $\lambda p_k = p_1 \cdot \dots \cdot p_n$ und damit erhalten wir durch aufsummieren:

$$\lambda = \lambda(p_1 + \dots + p_n) = n p_1 \cdot \dots \cdot p_n = n \lambda p_k.$$

Da $p_k \neq 0$ erhalten wir:

$$p_k = \frac{1}{n}.$$

Daraus folgt:

$$f(x_1, \dots, x_n) = x_1 \cdot \dots \cdot x_n \leq \frac{1}{n^n}$$

für alle $x_i > 0$ mit $\sum_{i=1}^n x_i = 1$. Da nur ein Extremwert auf M existiert, ist das globale Maximum eindeutig bestimmt und es gilt die Gleichheit dann, und nur dann, falls $x_i = \frac{1}{n}$

für alle $i \in \{1, \dots, n\}$.

Dieses Beispiel hat folgende Anwendung: Sei $(a_1, \dots, a_n) \in \mathbb{R}^n$ mit $a_i > 0$, so ist

$$x = \frac{(a_1, \dots, a_n)}{\sum_{i=1}^n a_i} \in M.$$

Also gilt:

$$\frac{\prod_{i=1}^n a_i}{\left(\sum_{i=1}^n a_i\right)^n} \leq \frac{1}{n^n}$$

oder

$$(a_1 \cdot \dots \cdot a_n)^{1/n} \leq \frac{1}{n} \cdot \sum_{i=1}^n a_i.$$

Dabei gilt die Gleichheit nur dann, falls $a_1 = \dots = a_n$ gilt. Man sagt, dass das geometrische Mittel kleiner als das arithmetische Mittel ist.

- (b) Hier wollen wir mit analytischen Mitteln die Diagonalisierbarkeit von symmetrischen Matrizen beweisen. Sei $A \in \text{Sym}(n, \mathbb{R})$ eine symmetrische $n \times n$ -Matrix. Betrachte die Abbildung $f: \mathbb{R}^n \rightarrow \mathbb{R}$ mit

$$f(x) = \langle Ax, x \rangle$$

und $g: \mathbb{R}^n \rightarrow \mathbb{R}$ die Abbildung mit

$$g(x) = \langle x, x \rangle = \sum_{i=1}^n x_i^2.$$

Dann ist $g^{-1}(1) = S^{n-1}$ und

$$\text{grad } g(x) = 2(x_1, \dots, x_n) \neq 0$$

für $(x_1, \dots, x_n) \in S^{n-1}$. Da S^{n-1} kompakt ist, nimmt f auf S^{n-1} an einer Stelle $p = (p_1, \dots, p_n) \in S^{n-1}$ ein Maximum an. Da A symmetrisch ist, folgt

$$Df(x)h = \langle Ah, x \rangle + \langle Ax, h \rangle = 2\langle h, Ax \rangle = \langle h, \text{grad } f(x) \rangle$$

für alle $h \in \mathbb{R}^n$. Also ist

$$2Ax = \text{grad } f(x).$$

Aus $\text{grad } f(p) = \lambda \text{grad } g(p)$ folgt $2A(p) = 2\lambda p$, d.h. $Ap = \lambda p$ und somit ist p Eigenvektor von A . Das orthogonale Komplement von $E_p = \text{Span}\{p\}$ ist der lineare $(n-1)$ -dimensionale Untervektorraum

$$E_p^\perp = \{x \in \mathbb{R}^n \mid \langle x, p \rangle = 0\}.$$

Da A symmetrisch ist, gilt: $A(E_p^\perp) \subset E_p^\perp$, wegen:

$$\langle Ax, p \rangle = \langle x, Ap \rangle = \langle x, \lambda p \rangle = 0,$$

falls $x \in E_p^\perp$. Mit obigem Argument hat dann aber auch $A|_{E_p^\perp}$ einen Eigenvektor. Die Wiederholung dieses Prozesses liefert eine direkte Summe orthogonaler Eigenräume von A .

16.7 Gewöhnliche Differentialgleichungen und der Satz von Picard-Lindelöf.

Nun wollen wir den Banachschen Fixpunktsatz auf gewöhnliche Differentialgleichungen anwenden. Gewisse lineare gewöhnliche Differentialgleichungen sind uns schon in Kapitel 9 begegnet und waren auch Gegenstand einer Reihe von Übungsaufgaben. Nun wollen wir den allgemeinen Fall untersuchen. Dazu müssen wir präzisieren, was wir unter einer gewöhnlichen Differentialgleichung verstehen. Zunächst erinnern wir an den Begriff des Vektorfeldes.

Sei U eine offene Teilmenge des \mathbb{R}^n und $k \in \mathbb{N} \cup \{0\}$. Unter einem C^k -Vektorfeld auf U verstehen wir eine k -mal stetig differenzierbare Abbildung $F : U \rightarrow \mathbb{R}^n$. Für $k = 0$ erhalten wir

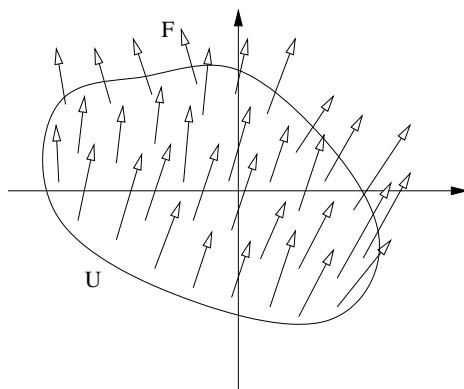


Abbildung 16.16: zeitunabhängiges Vektorfeld

die stetigen Vektorfelder. Oft hängen diese Vektorfelder noch von einem zusätzlichen reellen Parameter ab, der in der Physik meist als Zeit interpretiert wird. Genauer definiert man: Ist $I \subset \mathbb{R}$ ein Intervall, so heißt eine Abbildung

$$F : I \times U \rightarrow \mathbb{R}^n$$

auch zeitabhängiges Vektorfeld. Hier erlaubt man also, dass für jedes $x \in U$ der Vektor $F(t, x)$ sich mit $t \in I$ in einem Zeitintervall I ändert. Ist $J \subset I$ ein Intervall und $c : J \rightarrow U$ eine differenzierbare Kurve, so heißt c *Integralkurve* eines Vektorfeldes $F : I \times U \rightarrow \mathbb{R}^n$, falls

$$\dot{c}(t) = F(t, c(t)) \quad (16.8)$$

für alle $t \in J$ gilt. Die Gleichung (16.8) heißt auch *gewöhnliche, nicht autonome Differentialgleichung erster Ordnung*. Hängt das Vektorfeld $F : U \rightarrow \mathbb{R}^n$ nicht von der Zeit ab, so heißt

$$\dot{c}(t) = F(c(t)) \quad (16.9)$$

auch *gewöhnliche, autonome Differentialgleichung erster Ordnung*. Die Menge U wird in der Physik auch *Phasenraum* der Gleichung (16.8) bzw. (16.9) genannt. Ist $(t_0, x_0) \in I \times U$ und $c : J \rightarrow U$ eine Integralkurve von (16.8) mit $c(t_0) = x_0$, so sagen wir, dass c eine Lösung der gewöhnlichen Differentialgleichung (16.8) ist mit der Anfangsbedingung $c(t_0) = x_0$.

Beispiele.

- (a) *Lineare Differentialgleichungen*

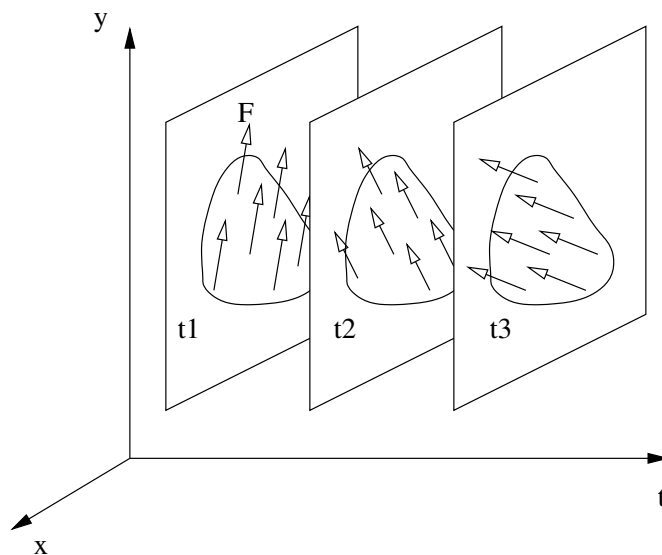


Abbildung 16.17: zeitabhängiges Vektorfeld

Sind $A : I \rightarrow M(n, \mathbb{R})$ und $b : I \rightarrow \mathbb{R}^n$ stetige Kurven und $F : I \times \mathbb{R}^n$ das Vektorfeld mit

$$F(t, x) = A(t)x + b(t),$$

so heißt die Differentialgleichung

$$\dot{c}(t) = F(t, c(t)) = A(t)c(t) + b(t) \quad (16.10)$$

lineare Differentialgleichung 1. Ordnung. Ist $b = 0$, so heißt (16.10) auch *homogene lineare Differentialgleichung 1. Ordnung.*

(b) *Hamiltonsche Differentialgleichungen*

Seien $U \subset \mathbb{R}^n$ eine offene Teilmenge und $H : U \times \mathbb{R}^n \rightarrow \mathbb{R}$ eine C^2 -Abbildung mit $(q, p) \mapsto H(q, p)$. Seien

$$\frac{\partial H}{\partial q} := \left(\frac{\partial H}{\partial q_1}, \dots, \frac{\partial H}{\partial q_n} \right) \quad \text{und} \quad \frac{\partial H}{\partial p} := \left(\frac{\partial H}{\partial p_1}, \dots, \frac{\partial H}{\partial p_n} \right)$$

die partiellen Ableitungen von H nach q und nach p . Dann heißt

$$X_H(q, p) = \left(-\frac{\partial H}{\partial p}, \frac{\partial H}{\partial q} \right)$$

das zu H gehörige *Hamiltonsche Vektorfeld*. Die Gleichung

$$(\dot{q}(t), \dot{p}(t)) = X_H(q(t), p(t)) \quad (16.11)$$

heißt *Hamiltonsche Differentialgleichung*. Viele Probleme, die in der Physik auftreten, lassen sich durch Differentialgleichungen dieses Typs beschreiben. Dabei wird H als Energie des Systems gedeutet. Als Beispiel betrachte die Funktionen $T : \mathbb{R}^n \rightarrow \mathbb{R}$ mit $T(p) = \frac{1}{2}\langle p, p \rangle$ (*kinetische Energie*) und $V \in C^2(U, \mathbb{R})$ (*potentielle Energie*). Dann ist $H(q, p) = T(p) + V(q)$ die *Gesamtenergie* des Systems.

Definition 16.7.1. Es seien $U \subset \mathbb{R}^n$, I ein Intervall und $F : I \times U \rightarrow \mathbb{R}^n$ ein stetiges Vektorfeld. Dann heißt F *Lipschitz-stetig* auf U , falls eine Konstante (sog. *Lipschitzkonstante*) $L \geq 0$ existiert, mit

$$\|F(t, x) - F(t, y)\| \leq L\|x - y\|$$

für alle $(t, x), (t, y) \in I \times U$. F heißt *lokal Lipschitz-stetig* auf U , falls zu jedem $(t_0, x_0) \in I \times U$ eine Umgebung $I_0 \times U_0 \subset I \times U$ von (t_0, x_0) existiert, so dass $F : I_0 \times U_0 \rightarrow \mathbb{R}^n$ Lipschitz-stetig auf U_0 ist.

Bemerkung. Ist $F : I \times U \rightarrow \mathbb{R}^n$ ein stetig differenzierbares Vektorfeld, so ist F lokal Lipschitz-stetig auf U . Denn wegen der Stetigkeit von DF existiert zu jeder kompakten Teilmenge $K \subset I \times U$ ein $L \geq 0$ mit $\|DF((t, x))\| \leq L$ für alle $(t, x) \in K$. Ist nun K konvex, so folgt aus dem Mittelwertsatz (siehe Korollar 15.3.16)

$$\|F(t, x) - F(t, y)\| \leq L\|(t, x) - (t, y)\| = L\|x - y\|.$$

Insbesondere sind Mengen der Form $Q = [a, b] \times K(x_0, r)$ kompakte konvexe Mengen.

Satz 16.7.2 (Picard-Lindelöf). *Seien $I \subset \mathbb{R}$ ein offenes Intervall, $U \subset \mathbb{R}^n$ eine offene Teilmenge des \mathbb{R}^n und $F : I \times U \rightarrow \mathbb{R}^n$ ein stetiges Vektorfeld, das lokal Lipschitz-stetig auf U ist. Dann existiert zu jedem Punkt (t_0, x_0) ein $\delta > 0$, so dass die Differentialgleichung*

$$\dot{c}(t) = F(t, c(t)) \tag{16.12}$$

genau eine Lösung $c : [t_0 - \delta, t_0 + \delta] \rightarrow U$ mit Anfangswert $c(t_0) = x_0$ besitzt.

Beweis. Es gilt: $c : [t_0 - \delta, t_0 + \delta] \rightarrow U$ ist genau dann eine Lösung von (16.12) mit Anfangswert $c(t_0) = x_0$, falls

$$c(t) = x_0 + \int_{t_0}^t F(s, c(s)) ds$$

für alle $t \in [t_0 - \delta, t_0 + \delta]$ gilt. Wir wollen diese Lösung durch Anwendung des Banachschen Fixpunktsatzes erhalten. Zunächst ist der Vektorraum $C^0([t_0 - \delta, t_0 + \delta], \mathbb{R}^n)$ bezüglich der Maximumsnorm

$$\|\varphi\|_0 := \max\{\|\varphi(t)\| \mid t \in [t_0 - \delta, t_0 + \delta]\}$$

vollständig. Außerdem ist für jede abgeschlossene Teilmenge $A \subset \mathbb{R}^n$ die Menge

$$C^0([t_0 - \delta, t_0 + \delta], A) = \{\varphi \in C^0([t_0 - \delta, t_0 + \delta], \mathbb{R}^n) \mid \varphi(t) \in A \text{ für alle } t \in [t_0 - \delta, t_0 + \delta]\}$$

abgeschlossen und somit als metrischer Raum vollständig. Betrachte die Abbildung

$$T : C^0([t_0 - \delta, t_0 + \delta], U) \rightarrow C^1([t_0 - \delta, t_0 + \delta], \mathbb{R}^n) \subset C^0([t_0 - \delta, t_0 + \delta], \mathbb{R}^n)$$

mit

$$T(\varphi)(t) = x_0 + \int_{t_0}^t F(s, \varphi(s)) ds.$$

Sei $Q := [t_0 - a, t_0 + a] \times K(x_0, r)$ eine kompakte Teilmenge von $I \times U$, so dass

$$F : [t_0 - a, t_0 + a] \times K(x_0, r) \rightarrow \mathbb{R}^n$$

auf $K(x_0, r)$ Lipschitz-stetig mit Lipschitzkonstante L ist. Da F insbesondere stetig ist, existiert eine Konstante $M \geq 0$ mit $\|F(t, x)\| \leq M$. Wähle nun ein $\delta > 0$ mit $a \geq \delta$, $\delta M \leq r$ und $\delta L < 1$. Betrachte die abgeschlossene Teilmenge von $C^0([t_0 - \delta, t_0 + \delta], \mathbb{R}^n)$

$$A_{\delta, r} := \{\varphi : [t_0 - \delta, t_0 + \delta] \rightarrow K(x_0, r) \mid \varphi \text{ stetig}\}.$$

Sie ist invariant unter T (d.h. $T(A_{\delta, r}) \subset A_{\delta, r}$), denn

$$\|T(\varphi)(t) - x_0\| \leq \int_{t_0}^t \|F(s, \varphi(s))\| ds \leq \delta M \leq r.$$

Außerdem ist die Einschränkung $T : A_{(\delta, r)} \rightarrow A_{(\delta, r)}$ eine Kontraktion, denn sind $\varphi, \psi \in A_{(\delta, r)}$, so folgt für alle $t \in [t_0 - \delta, t_0 + \delta]$:

$$\begin{aligned} \|T(\varphi)(t) - T(\psi)(t)\| &= \left\| \int_{t_0}^t (F(s, \varphi(s)) - F(s, \psi(s))) ds \right\| \leq \int_{t_0}^t \|F(s, \varphi(s)) - F(s, \psi(s))\| ds \\ &\leq \int_{t_0}^t L \|\varphi(s) - \psi(s)\| ds \leq \delta L \|\varphi - \psi\|_0 \end{aligned}$$

und somit

$$\|T(\varphi) - T(\psi)\|_0 \leq \delta L \|\varphi - \psi\|_0.$$

Da $\delta L < 1$, ist $T : A_{(\delta, r)} \rightarrow A_{(\delta, r)}$ eine Kontraktion und besitzt somit genau einen Fixpunkt $c \in A_{(\delta, r)}$. \square

Bemerkung. Ist $J \subset I$ ein beliebiges Intervall und sind $c : J \rightarrow U$ und $\tilde{c} : J \rightarrow U$ Integralkurven von (16.12) mit $c(t_0) = \tilde{c}(t_0)$ für ein $t_0 \in J$, so gilt: $c(t) = \tilde{c}(t)$ für alle $t \in J$. Denn die Menge

$$\{t \in J \mid c(t) = \tilde{c}(t)\}$$

ist wegen der Stetigkeit der Kurven c und \tilde{c} abgeschlossen in J . Wegen des Satzes von Picard-Lindelöf ist diese Menge aber auch offen in J , denn ist $t_1 \in J$, so existiert ein $\delta > 0$ mit $c(t) = \tilde{c}(t)$ für alle $t \in J \cap [t_1 - \delta, t_1 + \delta]$. Da J zusammenhängend ist, folgt die Behauptung.

Definition 16.7.3. Seien $I \subset \mathbb{R}$ ein offenes Intervall, $U \subset \mathbb{R}^n$ eine offene Teilmenge des \mathbb{R}^n und $F : I \times U \rightarrow \mathbb{R}^n$ ein stetiges Vektorfeld. Sei $J \subset I$ ein Intervall mit $t_0 \in J$ und $c : J \rightarrow \mathbb{R}^n$ eine Integralkurve mit Anfangswert (t_0, x_0) . Dann heißt $c : J \rightarrow \mathbb{R}^n$ *maximal*, falls für jede weitere Integralkurve $\tilde{c} : \tilde{J} \rightarrow \mathbb{R}^n$ mit Anfangswert (t_0, x_0) gilt: $\tilde{J} \subset J$ und $c(t) = \tilde{c}(t)$ für alle $t \in \tilde{J}$.

Satz 16.7.4. Sei $F : I \times U \rightarrow \mathbb{R}^n$ ein stetiges Vektorfeld, das lokal Lipschitz-stetig auf der offenen Menge $U \subset \mathbb{R}^n$ ist. Dann existiert zu jedem $(t_0, x_0) \in I \times U$ genau eine maximale Integralkurve $c : J_{(t_0, x_0)} \rightarrow \mathbb{R}^n$. Außerdem ist das maximale Intervall $J_{(t_0, x_0)}$ offen.

Beweis. Sei

$$J_{(t_0, x_0)} := \bigcup_{J \subset I} J \mid \text{es existiert eine Integralkurve } c_J : J \rightarrow \mathbb{R}^n \text{ mit } c_J(t_0) = x_0$$

Dann ist $J_{(t_0, x_0)}$ als Vereinigung von Intervallen, welche t_0 enthalten, wieder ein Intervall. Definiere eine Kurve $c : J_{(t_0, x_0)} \rightarrow \mathbb{R}^n$ mit

$$c(t) = c_J(t),$$

falls $t \in J$. Diese Definition hängt nicht von der Wahl von J ab. Denn sind $c_{J_1} : J_1 \rightarrow \mathbb{R}^n$ und $c_{J_2} : J_2 \rightarrow \mathbb{R}^n$ zwei Integralkurven mit $c_{J_1}(t_0) = c_{J_2}(t_0)$ und $t \in J_1 \cap J_2$, so folgt $c_{J_1}(t) = c_{J_2}(t)$ aus obiger Bemerkung. Damit ist $c : J_{(t_0, x_0)} \rightarrow \mathbb{R}^n$ eine Integralkurve und nach Konstruktion von $J_{(t_0, x_0)}$ ist sie auch maximal. Außerdem ist das maximale Intervall $J_{(t_0, x_0)}$ offen, denn wäre der rechte Randpunkt t_+ von $J_{(t_0, x_0)}$ in $J_{(t_0, x_0)}$ enthalten, so gäbe es wegen des Satzes von Picard-Lindelöf ein $\delta > 0$ und eine Integralkurve $\tilde{c} : [-\delta + t_+, \delta + t_+] \rightarrow \mathbb{R}^n$ mit $\tilde{c}(t_+) = c(t_+)$. Somit lässt sich die Lösung im Widerspruch zur Definition von $J_{(t_0, x_0)}$ fortsetzen. Also gehört der rechte Randpunkt t_+ nicht zu $J_{(t_0, x_0)}$. Dies gilt natürlich auch für den linken Randpunkt. \square

Beispiel. Betrachte das folgende zeitunabhängige Vektorfeld $F : \mathbb{R} \rightarrow \mathbb{R}$ mit $F(x) = x^2 + 1$. Die zugehörige Differentialgleichung lautet

$$\dot{c}(t) = c(t)^2 + 1.$$

Diese Differentialgleichung lässt sich elementar lösen, denn ist $c : I \rightarrow \mathbb{R}$ die maximale Lösung mit $c(t_0) = x_0$, so folgt aus

$$\arctan(c(s))' = \frac{\dot{c}(s)}{c(s)^2 + 1} = 1$$

die Gleichung

$$\arctan(c(t)) - \arctan(c(t_0)) = \int_{t_0}^t \frac{\dot{c}(s)}{c(s)^2 + 1} = t - t_0.$$

Also folgt

$$c(t) = \tan(t - t_0 + \arctan(x_0)).$$

Ist $y_0 = \arctan(x_0) \in (-\frac{\pi}{2}, +\frac{\pi}{2})$, so ist das maximale Lösungsintervall durch

$$I = \{t \in \mathbb{R} \mid t - t_0 + y_0 \in (-\frac{\pi}{2}, +\frac{\pi}{2})\} = \left(-\frac{\pi}{2} - y_0 + t_0, +\frac{\pi}{2} - y_0 + t_0\right)$$

gegeben.

Dieses Beispiel zeigt, dass die maximale Integralkurve des zeitunabhängigen Vektorfeldes F jeden kompakten Bereich des Phasenraumes verlässt. Der folgende Satz zeigt, dass dies immer richtig ist, falls die maximale Integralkurve nicht auf dem maximal möglichen Intervall definiert ist.

Satz 16.7.5. Sei $-\infty \leq a_0 < b_0 \leq \infty$ und $F : (a_0, b_0) \times U \rightarrow \mathbb{R}^n$ ein Vektorfeld, das lokal Lipschitz-stetig auf der offenen Menge $U \subset \mathbb{R}^n$ ist. Sei $c : (a, b) \rightarrow \mathbb{R}^n$ die maximale Integralkurve zu dem Anfangswert $c(t_0) = x_0$. Ist $b < b_0$, so existiert zu jeder kompakten

Teilmenge $K \subset U$ und jedem $\epsilon > 0$ ein $t \in (b - \epsilon, b)$ mit $c(t) \notin K$. Entsprechendes gilt für den linken Randpunkt.

Verläuft insbesondere die Integralkurve c ganz in einer kompakten Teilmenge von U , so ist c auf dem ganzen Intervall (a_0, b_0) erklärt.

Beweis. Sei $c : (a, b) \rightarrow \mathbb{R}^n$ eine maximale Integralkurve mit $b < b_0$. Angenommen $c(t)$ würde nicht bei Annäherung von t an den rechten Randpunkt jede kompakte Teilmenge von U verlassen. Dann existiert also eine kompakte Teilmenge $K \subset U$ und ein $\epsilon > 0$ mit $c(t) \in K$ für alle $t \in (b - \epsilon, b)$. Betrachte nun eine monoton steigende Folge $t_n \in (b - \epsilon, b)$, die gegen b konvergiert. Aus der Stetigkeit von F folgt die Existenz einer Konstanten $M > 0$ mit $\|F(t, x)\| \leq M$ für alle $(t, x) \in [b - \epsilon, b] \times K$. Dann folgt für $n \geq m$

$$\|c(t_n) - c(t_m)\| = \left\| \int_{t_m}^{t_n} \dot{c}(t) dt \right\| \leq \int_{t_m}^{t_n} \|F(t, c(t))\| dt \leq (t_n - t_m)M.$$

Daher ist die Folge $c(t_n)$ eine Cauchyfolge und hat damit einen Grenzwert x in K . Außerdem hängt der Grenzwert nicht von der Wahl der Folge ab und somit ist die Fortsetzung $\tilde{c} : (a, b] \rightarrow \mathbb{R}^n$ mit $\tilde{c}(t) = c(t)$ für $t \in (a, b)$ und $\tilde{c}(b) = x$ stetig. Nun gilt für alle $t \in (a, b)$:

$$\tilde{c}(t) = x_0 + \int_{t_0}^t F(s, \tilde{c}(s)) ds.$$

Da beide Seiten in b stetig sind, gilt diese Gleichung auch für $t = b$ und somit ist $\tilde{c} : (a, b] \rightarrow \mathbb{R}^n$ eine Integralkurve mit $\tilde{c}(t_0) = x_0$ im Widerspruch zur Maximalität von $c : (a, b) \rightarrow \mathbb{R}^n$. \square

Unter gewissen Bedingungen existieren die Integralkurven auf dem maximal möglichen Intervall. Dazu benötigen wir das folgende einfache Lemma von Gronwall.

Lemma 16.7.6 (Lemma von Gronwall). Sei $u : I \rightarrow \mathbb{R}$ eine stetige nicht-negative Funktion auf einem Intervall I und $t_0 \in I$. Existieren Konstanten $A \geq 0$ und $B \geq 0$ mit

$$u(t) \leq A + B \left| \int_{t_0}^t u(s) ds \right|$$

für alle $t \in I$, so gilt für alle $t \in I$:

$$u(t) \leq A e^{B|t-t_0|}.$$

Beweis. Sei $\epsilon > 0$ und $t \geq t_0$. Aus der Annahme folgt

$$u(t) \leq A + \epsilon + B \int_{t_0}^t u(s) ds =: g(t).$$

Dann gilt

$$g'(t) = Bu(t) \leq Bg(t)$$

und da $g(t) \geq A + \epsilon > 0$ erhalten wir

$$(\log g(t))' = \frac{g'(t)}{g(t)} \leq B.$$

Durch Integration dieser Abschätzung folgt

$$\log \left(\frac{g(t)}{g(t_0)} \right) = \log g(t) - \log g(t_0) \leq B(t - t_0)$$

und somit durch Anwendung der Exponentialabbildung

$$u(t) \leq g(t) \leq g(t_0)e^{B(t-t_0)} = (A + \epsilon)e^{B(t-t_0)}.$$

Da diese Abschätzung für alle $\epsilon > 0$ gilt, folgt die Behauptung des Lemmas für $t \geq t_0$. Der Fall $t < t_0$ folgt mit analogen Argumenten und sei daher als Übung überlassen. \square

Definition 16.7.7. Ein Vektorfeld $F : I \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ heißt *linear beschränkt*, wenn stetige Funktionen $\alpha, \beta : I \rightarrow \mathbb{R}$ existieren mit

$$\|F(t, x)\| = \alpha(t) + \beta(t)\|x\|.$$

Satz 16.7.8. Sei $F : (a_0, b_0) \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ ein linear beschränktes Vektorfeld, welches lokal Lipschitz-stetig auf \mathbb{R}^n ist. Dann ist jede maximale Integralkurve von F auf ganz (a_0, b_0) erklärt.

Beweis. Sei $c : (a, b) \rightarrow \mathbb{R}^n$ eine maximale Integralkurve und $b < b_0$. Dann wäre für ein gegebenes $t_0 \in (a, b)$ die Integralkurve c auf $[t_0, b)$ unbeschränkt nach Satz 16.7.5. Auf der anderen Seite erhalten wir aus

$$c(t) = c(t_0) + \int_{t_0}^t F(s, c(s)) ds$$

wegen

$$\|F(s, c(s))\| \leq \alpha(s) + \beta(s)\|c(s)\| \leq A + B\|c(s)\|$$

mit

$$A = \max\{\|\alpha(s)\| \mid s \in [t_0, b]\} \quad \text{und} \quad B = \max\{\|\beta(s)\| \mid s \in [t_0, b]\}$$

die Abschätzung

$$\|c(t)\| \leq \|c(t_0)\| + \int_{t_0}^t \|F(s, c(s))\| ds \leq \|c(t_0)\| + (b - t_0)A + B \int_{t_0}^t \|c(s)\| ds.$$

Aus dem Lemma von Gronwall erhalten wir

$$\|c(t)\| \leq (\|c(t_0)\| + (b - t_0)A)e^{B(b-t_0)}$$

und somit ist $\|c(t)\|$ beschränkt auf $[t_0, b)$. Dies ist ein Widerspruch. \square

In Physik und Mathematik treten häufig Differentialgleichungen höherer Ordnung auf. Diese sind wie folgt definiert.

Definition 16.7.9. Seien I ein Intervall,

$$U \subset \underbrace{\mathbb{R}^n \times \dots \times \mathbb{R}^n}_{k\text{-mal}} = \mathbb{R}^{nk}$$

eine offene Menge und $f : I \times U \rightarrow \mathbb{R}^n$ eine stetige Abbildung. Sei $(a, b) \subset I$. Dann heißt eine k -mal stetig differenzierbare Kurve $y : (a, b) \rightarrow \mathbb{R}^n$ mit $(y(t), \dot{y}(t), \dots, y^{(k-1)}(t)) \subset U$ und

$$y^{(k)}(t) = f(t, y(t), \dot{y}(t), \dots, y^{(k-1)}(t)) \quad (16.13)$$

Lösung der durch $f : I \times U \rightarrow \mathbb{R}^n$ induzierten *Differentialgleichung k -ter Ordnung* (16.13).

Jeder Differentialgleichung k -ter Ordnung entspricht eine Differentialgleichung erster Ordnung, denn es gilt:

Satz 16.7.10. Seien I ein Intervall, $U \subset \mathbb{R}^n \times \dots \times \mathbb{R}^n = \mathbb{R}^{nk}$ und $f : I \times U \rightarrow \mathbb{R}^n$ eine lokal Lipschitz-stetige Abbildung.

Eine k mal stetig differenzierbare Kurve $y : (a, b) \rightarrow \mathbb{R}^n$ mit $(a, b \subset I)$ ist genau dann Lösung der Differentialgleichung

$$y^{(k)}(t) = f(t, y(t), \dot{y}(t), \dots, y^{(k-1)}(t)),$$

falls $c : (a, b) \rightarrow U$ mit

$$c(t) = (c_1(t), c_2(t), \dots, c_k(t)) := (y(t), \dot{y}(t), \dots, y^{(k-1)}(t))$$

Lösung der folgenden Differentialgleichung erster Ordnung ist:

$$\begin{aligned} \dot{c}_1(t) &= c_2(t) \\ &\vdots \\ \dot{c}_{k-1}(t) &= c_k(t) \\ \dot{c}_k(t) &= f(t, c_1(t), \dots, c_k(t)). \end{aligned} \quad (16.14)$$

Definieren wir das Vektorfeld $F : I \times U \rightarrow \mathbb{R}^{nk}$ durch

$$F(t, x) = F(t, x_1, \dots, x_k) = (x_2, \dots, x_k, f(t, x_1, \dots, x_k)),$$

so lässt sich Gleichung (16.14) auch kompakt in der Form

$$\dot{c}(t) = F(t, c(t))$$

schreiben. Insbesondere gibt es wegen des Satzes von Picard-Lindelöf zu jedem Anfangswert $(t_0, y_1, \dots, y_k) \in I \times U$ genau eine Lösung

$$c(t) = (c_1(t), c_2(t), \dots, c_k(t)) := (y(t), \dot{y}(t), \dots, y^{(k-1)}(t))$$

mit

$$(y(t_0), \dot{y}(t_0), \dots, y^{(k-1)}(t_0)) = (y_1, y_2, \dots, y_k).$$

Beispiel. Sei $f : U \rightarrow \mathbb{R}^n$ ein Vektorfeld (*Kraftfeld*) auf einer offenen Menge $U \subset \mathbb{R}^n$. Dann heißen 2-mal stetig differenzierbare Kurven $y : I \rightarrow U$ mit

$$\ddot{y}(t) = f(y(t)) \quad (16.15)$$

Lösungen der *Newtonschen Bewegungsgleichungen* (16.15). Insbesondere entspricht jeder Lösung $y : I \rightarrow U$ die Lösung $c : I \rightarrow U \times \mathbb{R}^n$ mit $c(t) = (c_1(t), c_2(t)) = (y(t), \dot{y}(t))$ der Gleichung

$$\begin{aligned} \dot{c}_1(t) &= c_2(t) \\ \dot{c}_2(t) &= f(c_1(t)) \end{aligned}$$

In der Physik ist es üblich, U als *Konfigurationsraum* und $U \times \mathbb{R}^n$ als *Phasenraum* zu bezeichnen.

16.8 Lineare Differentialgleichungen II

Wir wollen nun das in Kapitel 9 begonnene Studium linearer Differentialgleichungen fortsetzen.

Satz 16.8.1. Seien $A : I \rightarrow M(n, \mathbb{R})$ und $b : I \rightarrow \mathbb{R}^n$ stetige Kurven und

$$\dot{c}(t) = F(t, c(t)) = A(t)c(t) + b(t) \quad (16.16)$$

die zugehörige lineare Differentialgleichung. Dann ist jede maximale Integralkurve von (16.16) auf ganz I erklärt.

Beweis. Das Vektorfeld $F(t, x) = A(t)x + b(t)$ ist wegen

$$\|F(t, x)\| \leq \|A(t)\| \|x\| + \|b(t)\|$$

linear beschränkt und wegen

$$\|F(t, x) - F(t, y)\| \leq \|A(t)\| \|x - y\|$$

auch lokal Lipschitz-stetig auf \mathbb{R}^n . □

Satz 16.8.2. Sei $A : I \rightarrow M(n, \mathbb{R})$ eine stetige Kurve und $C^1(I, \mathbb{R}^n)$ der Vektorraum der auf I stetig differenzierbaren Kurven. Betrachte den linearen Operator

$$L : C^1(I, \mathbb{R}^n) \rightarrow C^0(I, \mathbb{R}^n)$$

mit $L(c)(t) := \dot{c}(t) - A(t)c(t)$. Dann ist der Kern von L ein n -dimensionaler Untervektorraum von $C^1(I, \mathbb{R}^n)$. Für $c_1, \dots, c_n \in \text{Kern } L$ sind folgende Aussagen äquivalent:

- (a) c_1, \dots, c_n ist Basis von $\text{Kern } L$.
- (b) $c_1(t_0), \dots, c_n(t_0)$ ist Basis von \mathbb{R}^n für alle $t_0 \in I$.

(c) $c_1(t_0), \dots, c_n(t_0)$ ist Basis von \mathbb{R}^n für ein $t_0 \in I$.

Beweis. Da $L : C^1(I, \mathbb{R}^n) \rightarrow C^0(I, \mathbb{R}^n)$ ein linearer Operator ist, ist der Kern von L ein Untervektorraum von $C^1(I, \mathbb{R}^n)$. Zu $t_0 \in I$ betrachte die lineare Abbildung

$$\phi_{t_0} : \text{Kern } L \rightarrow \mathbb{R}^n$$

mit $\phi_{t_0}(c) = c(t_0)$. Da wegen des Satzes von Picard-Lindelöf und wegen des Satzes 16.8.1 zu jedem $x_0 \in \mathbb{R}^n$ genau ein $c \in \text{Kern } L$ mit $c(t_0) = x_0$ existiert, ist $\phi_{t_0} : \text{Kern } L \rightarrow \mathbb{R}^n$ ein Isomorphismus. Insbesondere ist $\dim \text{Kern } L = n$.

Da ϕ_{t_0} als Isomorphismus Basen in Basen überführt, folgt die Äquivalenz von (a) und (b) sowie die Äquivalenz von (a) und (c). Dann sind aber auch (b) und (c) äquivalent. \square

Definition 16.8.3. Sei $A : I \rightarrow M(n, \mathbb{R})$ eine stetige Kurve und

$$L : C^1(I, \mathbb{R}^n) \rightarrow C^0(I, \mathbb{R}^n)$$

der lineare Operator mit $L(c)(t) := \dot{c}(t) - A(t)c(t)$. Eine Basis $c_1, \dots, c_n \in \text{Kern } L$ heißt *Fundamentalsystem* der homogenen Differentialgleichung $\dot{c}(t) = A(t)c(t)$.

Bemerkungen.

(a) Es ist zweckmäßig, ein Fundamentalsystem c_1, \dots, c_n zu einer matrixwertigen Abbildung (sog. *Fundamentalmatrix*)

$$C = (c_1, \dots, c_n) : I \rightarrow M(n, \mathbb{R}), \text{ mit } t \mapsto (c_1(t), \dots, c_n(t))$$

zusammenzufassen. Aus den Regeln der Matrixmultiplikation folgt:

$$\dot{C}(t) = A(t)C(t).$$

Ist $c \in \text{Kern } L$ eine beliebige Lösung, so existiert ein $y = (y_1, \dots, y_n) \in \mathbb{R}^n$ mit $c(t) = y_1 c_1(t) + \dots + y_n c_n(t)$. Insbesondere ist $c(t) = C(t)y$.

(b) Ist $A(t) = A$, so ist

$$C(t) = e^{At}$$

ein Fundamentalsystem. Die zugehörige homogene lineare Differentialgleichung heißt auch *lineare homogene Differentialgleichung mit konstanten Koeffizienten*.

Satz 16.8.4 (Satz von Liouville). Sei $C : I \rightarrow M(n, \mathbb{R})$ ein Fundamentalsystem der homogenen Differentialgleichung

$$\dot{c}(t) = A(t)c(t).$$

Dann gilt

$$\frac{d}{dt} \det C(t) = \text{Spur}(A(t)) \det C(t)$$

und somit für $t_0, t \in I$:

$$\det C(t) = \det C(t_0) e^{\int_{t_0}^t \text{Spur}(A(s)) ds}$$

Beweis. Zunächst erinnern wir an die folgende Formel für die Ableitung einer Determinante (siehe Übungsblatt 13, MfP III). Sei $B : I \rightarrow M(n, \mathbb{R})$ eine differenzierbare Kurve mit $B(t_0) = \text{id}$, so folgt:

$$\left. \frac{d}{dt} \right|_{t=t_0} \det B(t) = \text{Spur } \dot{B}(t_0).$$

Ist nun $B : I \rightarrow M(n, \mathbb{R})$ eine differenzierbare Kurve und $B(t_0)$ invertierbar, so gilt daher:

$$\left. \frac{d}{dt} \right|_{t=t_0} \det B(t)B^{-1}(t_0) = \text{Spur}(\dot{B}(t_0)B^{-1}(t_0))$$

und somit

$$\left. \frac{d}{dt} \right|_{t=t_0} \det B(t) = \text{Spur}(\dot{B}(t_0)B^{-1}(t_0)) \det B(t_0).$$

Ist weiter $C : I \rightarrow M(n, \mathbb{R})$ ein Fundamentalsystem, so folgt aus $\dot{C}(t) = A(t)C(t)$ und der Invertierbarkeit von $C(t)$:

$$\dot{C}(t)C^{-1}(t) = A(t).$$

Also erhalten wir:

$$\frac{d}{dt} \det C(t) = \text{Spur}(\dot{C}(t)C^{-1}(t)) \det C(t) = \text{Spur}(A(t)) \det C(t).$$

Dies impliziert:

$$\frac{d}{dt} \log(\det C(t)) = \text{Spur}(A(t)).$$

Seien nun $t_0, t \in I$ so folgt:

$$\log(\det C(t)) - \log(\det C(t_0)) = \log \left(\frac{\det C(t)}{\det C(t_0)} \right) = \int_{t_0}^t \text{Spur}(A(s)) ds$$

und somit

$$\det C(t) = \det C(t_0) \exp \left(\int_{t_0}^t \text{Spur}(A(s)) ds \right).$$

□

Sei nun eine inhomogene lineare Differentialgleichung gegeben. Ist ein Fundamentalsystem für die homogene lineare Differentialgleichung gefunden, so lässt sich mit Hilfe der Variation der Konstanten eine spezielle Lösung der inhomogenen linearen Differentialgleichung bestimmen.

Satz 16.8.5 (Variation der Konstanten). *Seien $A : I \rightarrow M(n, \mathbb{R})$ und $b : I \rightarrow \mathbb{R}^n$ stetige Kurven und*

$$\dot{c}(t) = F(t, c(t)) = A(t)c(t) + b(t)$$

die zugehörige lineare Differentialgleichung. Sei $C : I \rightarrow M(n, \mathbb{R})$ ein Fundamentalsystem für die homogene lineare Differentialgleichung $\dot{c}(t) = A(t)c(t)$. Sei $t_0 \in I$, so ist $c_0 : I \rightarrow \mathbb{R}^n$ mit

$$c_0(t) = C(t)q(t) \quad \text{und} \quad q(t) = \int_{t_0}^t C(s)^{-1}b(s) ds$$

eine spezielle Lösung der inhomogenen Differentialgleichung.

Beweis. Es gilt:

$$\begin{aligned}\dot{c}_0(t) &= \dot{C}(t)q(t) + C(t)\dot{q}(t) = A(t)C(t)q(t) + C(t)C'(t)^{-1}b(t) \\ &= A(t)c_0(t) + b(t).\end{aligned}$$

□

Nun wollen wir den Zusammenhang zu den linearen Differentialgleichungen n -ter Ordnung herstellen, die wir in Kapitel 9 diskutiert haben. Ein lineare Abbildung $L : C^n(I, \mathbb{R}) \rightarrow C^0(I, \mathbb{R})$ mit

$$L(y)(t) = y^{(n)}(t) + a_{n-1}(t)y^{(n-1)}(t) + \dots + a_0(t)y(t)$$

und $a_i \in C^0(I, \mathbb{R})$ heißt linearer Differentialoperator n -ter Ordnung. Ist $b \in C^0(I, \mathbb{R})$, so heißt

$$L(y)(t) = y^{(n)}(t) + a_{n-1}(t)y^{(n-1)}(t) + \dots + a_0(t)y(t) = b(t) \quad (16.17)$$

lineare Differentialgleichung n -ter Ordnung. Betrachte zu $y \in C^n(I, \mathbb{R})$ die Kurve $c : I \rightarrow \mathbb{R}^n$ mit

$$c(t) := (c_1(t), c_2(t), \dots, c_n(t)) = (y(t), y'(t), \dots, y^{(n-1)}(t)).$$

Ist $y \in C^n(I, \mathbb{R})$ eine Lösung der linearen Differentialgleichung (16.17) n -ter Ordnung, so ist $c : I \rightarrow \mathbb{R}^n$ eine Lösung der linearen Differentialgleichung 1. Ordnung.

$$\begin{aligned}\dot{c}_1(t) &= c_2(t) \\ &\vdots \\ \dot{c}_{n-1}(t) &= c_n(t) \\ \dot{c}_n(t) &= -a_{n-1}(t)c_n(t) - \dots - a_0(t)c_1(t) + b(t).\end{aligned} \quad (16.18)$$

Dieses Gleichungssystem können wir auch wie folgt schreiben

$$c(t) = \begin{pmatrix} c'_1(t) \\ c'_2(t) \\ \vdots \\ \vdots \\ c'_n(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -a_0(t) & -a_1(t) & -a_2(t) & \dots & -a_{n-1}(t) \end{pmatrix} \begin{pmatrix} c_1(t) \\ c_2(t) \\ \vdots \\ \vdots \\ c_n(t) \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ b(t) \end{pmatrix}.$$

Ist umgekehrt $c : I \rightarrow \mathbb{R}^n$ eine Lösung von (16.18), so ist $y := c_1$ eine Lösung von (16.17).

16.9 Autonome Vektorfelder und dynamische Systeme.

Sei $U \subset \mathbb{R}^n$ offen und $f : U \rightarrow \mathbb{R}^n$ ein autonomes (d.h. zeitunabhängiges) lokal Lipschitz-stetiges Vektorfeld. Betrachte die zu f gehörige gewöhnliche Differentialgleichung

$$\dot{c}(t) = f(c(t)).$$

Wegen des Satzes von Picard-Lindelöf existiert für jedes $x \in U$ genau eine maximale Integralkurve $c_x : J_x \rightarrow U$ mit $c_x(0) = x$. Insbesondere ist J_x ein offenes Intervall mit $0 \in J_x$ und ist somit von der Form (a_x, b_x) mit $-\infty \leq a_x < 0 < b_x \leq \infty$. Sei nun

$$\Omega := \{(t, x) \mid t \in (a_x, b_x)\},$$

so ist die Abbildung $\phi : \Omega \rightarrow U$ mit

$$\phi(t, x) = c_x(t)$$

der Fluss des Vektorfeldes f . Manchmal schreibt man auch

$$\phi(t, x) = \phi^t(x).$$

Die Kurve $t \mapsto \phi(t, x)$ heißt auch *Flusslinie* (*Trajektorie*, *Bahnkurve*) des Vektorfeldes f durch $x \in U$. Man kann zeigen: $\Omega \subset \mathbb{R} \times \mathbb{R}^n$ ist offen und $\phi : \Omega \rightarrow U$ ist stetig (Hirsch, Smale: Differential Equations, Dynamical Systems and Linear Algebra, Seite 175).

Beispiel. Ist $A \in M(n, \mathbb{R})$ und $f(x) = A(x)$, so gilt (siehe Aufgabenblatt 8, MfP III):

$$\Omega = \mathbb{R} \times \mathbb{R}^n \quad \text{und} \quad \phi^t(x) = e^{tA}x.$$

Satz 16.9.1. Sei $U \subset \mathbb{R}^n$ offen und $f : U \rightarrow \mathbb{R}^n$ ein lokal Lipschitz-stetiges Vektorfeld. Sei $x \in U$ und $c_x : (a_x, b_x) \rightarrow U$ die maximale Integralkurve von f mit $c_x(0) = x$. Dann ist für jedes $s \in (a_x, b_x)$ die Kurve $c : (a_x - s, b_x - s) \rightarrow U$ mit $c(t) := c_x(t + s)$ die maximale Integralkurve mit $c(0) = c_x(s)$. Insbesondere folgt:

$$(a_{c_x(s)}, b_{c_x(s)}) = (a_x - s, b_x - s)$$

und

$$\phi^{t+s}(x) = c_x(t + s) = c(t) = c_{c_x(s)}(t) = \phi^t(c_x(s)) = \phi^t \circ \phi^s(x)$$

für alle $s \in (a_x, b_x)$ und $t \in (a_x - s, b_x - s)$.

Beweis. Sei $s \in (a(x), b(x))$ und $c(t) := c_x(t + s)$ mit $t \in (a_x - s, b_x - s)$. Dann ist $t + s \in (a_x, b_x)$. Da

$$\dot{c}(t) = \dot{c}_x(t + s) = f(c_x(t + s)) = f(c(t)), \tag{16.19}$$

ist $c : (a_x - s, b_x - s) \rightarrow U$ eine Integralkurve mit $c(0) = c_x(s)$. Aus der Maximalität von c_x folgt die Maximalität von c . Wegen der Eindeutigkeit der maximalen Integralkurven zu gegebenen Anfangswerten stimmt c mit der maximalen Integralkurve

$$c_{c_x(s)} : (a_{c_x(s)}, b_{c_x(s)}) \rightarrow U$$

überein, d.h. es gilt

$$c_x(t + s) = c(t) = c_{c_x(s)}(t)$$

für alle $t \in (a_{c_x(s)}, b_{c_x(s)}) = (a_x - s, b_x - s)$. □

Bemerkung. Gleichung (16.19) zeigt eine grundlegende Eigenschaft autonomer Vektorfelder: ist $t \mapsto c(t)$ eine Integralkurve, so auch die zeitverschobene Kurve $t \mapsto c(t + s)$. Dies ist für zeitabhängige Vektorfelder nicht der Fall.

Satz 16.9.2. Sei $U \subset \mathbb{R}^n$ offen und $f : U \rightarrow \mathbb{R}^n$ ein k -mal stetig differenzierbares Vektorfeld. Dann ist ϕ auf

$$\Omega := \{(t, x) \mid x \in U, t \in (a_x, b_x)\}$$

k -mal stetig differenzierbar. Außerdem existiert zu jedem $x \in U$ ein $\epsilon > 0$ und eine offene Umgebung $U_x \subset U$ von x mit $(-\epsilon, \epsilon) \times U_x \in \Omega$. Des Weiteren ist für jedes $t \in (-\epsilon, \epsilon)$ die Abbildung

$$\phi^t : U_x \rightarrow \phi^t(U_x)$$

ein Diffeomorphismus.

Beweis. Den Beweis der Differenzierbarkeit findet man z.B. in V.I. Arnold (Gewöhnliche Differentialgleichungen) oder Königsberger (Analysis 2). Da $\Omega \subset \mathbb{R} \times \mathbb{R}^n$ offen ist und da $(0, x) \in \Omega$ für alle $x \in U$, existiert zu jedem $x \in U$ ein $\epsilon > 0$ und eine offene Umgebung $U_x \subset U$ von x mit $(-\epsilon, \epsilon) \times U_x \in \Omega$. Insbesondere ist für jedes $t \in (-\epsilon, \epsilon)$ die Abbildung

$$\phi^t : U_x \rightarrow \phi^t(U_x)$$

definiert und differenzierbar. Wegen des Satzes 16.9.1 gilt:

$$\phi^{-t} \circ \phi^t(y) = \phi^0(y) = y$$

für alle $y \in U_x$ und somit ist $\phi^t : U_x \rightarrow \phi^t(U_x)$ ein Diffeomorphismus mit der inversen Abbildung $\phi^{-t} : \phi^t(U_x) \rightarrow U_x$. \square

Bemerkung. Ist $f : U \rightarrow \mathbb{R}^n$ ein k -mal stetig differenzierbares Vektorfeld und sind die Integralkurven auf ganz \mathbb{R} erklärt, so ist wegen Satz 16.9.2 $\phi^t : U \rightarrow \phi^t(U) = U$ ein Diffeomorphismus für alle $t \in \mathbb{R}$ und es gilt für alle $t, s \in \mathbb{R}$:

$$\phi^{t+s} = \phi^t \circ \phi^s.$$

Dabei ist $\phi^t(U) = U$, denn jeder Punkt $y \in U$ hat $\phi^{-t}(y) \in U$ als Urbild.

Als Beispiel betrachte das Vektorfeld $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ mit $A \in M(n, \mathbb{R})$. Dann gilt:

$$\phi^{t+s} = e^{(t+s)A} = e^{tA} e^{sA} = \phi^t \circ \phi^s.$$

Definition 16.9.3. Sei $U \subset \mathbb{R}^n$ offen. Ein *dynamisches System (globaler Fluss)* auf U ist eine differenzierbare Abbildung $\phi : \mathbb{R} \times U \rightarrow U$ mit $(t, x) \mapsto \phi^t(x)$ und

$$\phi^{t+s}(x) = \phi^t \circ \phi^s(x), \quad \phi^0(x) = x$$

für alle $x \in U$.

Bemerkungen.

- (a) Ist $\phi : \mathbb{R} \times U \rightarrow U$ ein dynamisches System so ist $\phi^t : U \rightarrow U$ für alle $t \in \mathbb{R}$ ein Diffeomorphismus, denn aus der Definition folgt, dass $\phi^{-t} : U \rightarrow U$ die differenzierbare Inverse ist.
- (b) Wir haben gesehen, dass jedes autonome Vektorfeld, dessen Integralkurven auf ganz \mathbb{R} existieren, ein dynamisches System erzeugt. Ist umgekehrt ein dynamisches System $\phi : \mathbb{R} \times U \rightarrow U$ gegeben, so erhalten wir durch

$$f(x) = \left. \frac{d}{dt} \right|_{t=0} \phi^t(x)$$

ein autonomes Vektorfeld. Man nennt daher dieses Vektorfeld auch den infinitesimalen Erzeuger des dynamischen Systemes.

- (c) Der Begriff des dynamischen Systemes lässt sich leicht auf Mannigfaltigkeiten ausdehnen.

Wir wollen nun das Verhalten von gewöhnlichen Differentialgleichungen unter Koordinatentransformationen untersuchen. Der Grund ist, dass die Wahl eines geeigneten Koordinatensystems die Lösung der Differentialgleichung stark vereinfachen kann.

Definition 16.9.4. Seien U und V offene Teilmengen des \mathbb{R}^n . Dann heißt ein Diffeomorphismus $\varphi : U \rightarrow V$ eine *Koordinatentransformation* von U nach V .

Satz 16.9.5. Sei $g : U \rightarrow \mathbb{R}^n$ ein lokal Lipschitz-stetiges Vektorfeld auf der offenen Menge $U \subset \mathbb{R}^n$ und

$$G : \Omega_g \rightarrow \mathbb{R}^n$$

der zu g gehörige Fluss. Sei $\varphi : U \rightarrow V$ eine Koordinatentransformation von U nach V und $f : U \rightarrow \mathbb{R}^n$ das Vektorfeld mit

$$f(\varphi(x)) = D\varphi(x)(g(x)).$$

Ist

$$F : \Omega_f \rightarrow \mathbb{R}^n$$

der zu f gehörige Fluss, so gilt:

$$(t, x) \in \Omega_g \Leftrightarrow (t, \varphi(x)) \in \Omega_f$$

und

$$F^t(\varphi(x)) = F(t, \varphi(x)) = \varphi(G(t, x)) = \varphi G^t(x)$$

für alle $(t, x) \in \Omega_g$.

Bemerkung. Daraus folgt also

$$F^t(y) = \varphi \circ G^t \circ \varphi^{-1}(y).$$

Zwei Flüsse F, G mit dieser Eigenschaft heißen auch *konjugiert*.

Beweis. Sei $c_x^g(t) : (a_x, b_x) \rightarrow U$ die maximale Integralkurve des Vektorfeldes g mit Anfangsbedingung $c_x^g(0) = x \in U$. Dann gilt für alle $t \in (a_x, b_x)$:

$$\dot{c}_x^g(t) = g(c_x^g(t))$$

Wenden wir das Differential $D\varphi(c_x^g(t))$ auf beide Seiten der Gleichung an, so erhalten wir aus der Definition von f :

$$D\varphi(c_x^g(t))(\dot{c}_x^g(t)) = D\varphi(c_x^g(t))(g(c_x^g(t))) = f(\varphi(c_x^g(t))).$$

Also folgt mit Hilfe der Kettenregel:

$$\frac{d}{dt}(\varphi(c_x^g(t))) = f(\varphi(c_x^g(t)))$$

und somit ist $\varphi \circ c_x^g : (a_x^g, b_x^g) \rightarrow V$ eine Integralkurve von f mit $\varphi \circ c_x^g(0) = \varphi(x)$. Diese Integralkurve ist auch maximal, denn ist c eine Integralkurve von f , so ist auch $\varphi^{-1} \circ c$ eine Integralkurve von g . Damit folgt

$$(a_x^g, b_x^g) = (a_{\varphi(x)}^f, b_{\varphi(x)}^f)$$

und

$$c_{\varphi(x)}^f(t) = \varphi \circ c_x^g(t)$$

für alle $t \in (a_x^g, b_x^g)$ und somit auch

$$F(t, \varphi(x)) = \varphi(G(t, x))$$

für alle $(t, x) \in \Omega_g$. □

16.10 Euler-Lagrange Gleichungen und Variationsrechnung.

In der Variationsrechnung geht es um die Berechnung kritischer Punkte von Funktionalen. Funktionale sind im Allgemeinen reellwertige Abbildungen, die auf Funktionenräumen definiert sind. Statt einer präzisen Definition werden wir nur das für die Physik wichtige Lagrangesche Funktional definieren.

Definition 16.10.1 (Lagrangesches Funktional). Es seien $q_1, q_2 \in \mathbb{R}^n$ zwei Punkte des \mathbb{R}^n und $t_1 < t_2$, $t_1, t_2 \in \mathbb{R}$. Sei

$$A = A_{q_1, q_2}^{t_1, t_2} = \{x : [t_1, t_2] \rightarrow \mathbb{R}^n \mid x \text{ 2-mal stetig differenzierbar mit } x(t_1) = q_1, x(t_2) = q_2\}$$

und $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^n \times [t_1, t_2] \rightarrow \mathbb{R}$ mit $(q, v, t) \mapsto \mathcal{L}(q, v, t)$ eine 2-mal stetig differenzierbare Funktion (sog. *Lagrangefunktion*). Dann heißt die Abbildung $\Phi : A \rightarrow \mathbb{R}$ mit

$$\Phi(x) := \int_{t_1}^{t_2} \mathcal{L}(x(t), \dot{x}(t), t) dt$$

Lagrangesches Funktional. Dieses Funktional wird auch *Wirkung* genannt.

Bemerkung. In der Physik wird $A_{q_1, q_2}^{t_1, t_2}$ als die Menge der Bahnkurven von ‘‘Teilchen’’ interpretiert, die sich zum Zeitpunkt t_1 in q_1 und zum Zeitpunkt t_2 in q_2 befinden. Die physikalisch relevanten Bahnen sind nun genau die kritischen Bahnen des Lagrangeschen Funktional mit $\mathcal{L} = T - V$. Häufig minimieren diese Bahnen \mathcal{L} , das ist aber nicht notwendig der Fall.

Beispiele.

(a) $\mathcal{L}(q, v, t) = \frac{1}{2} \langle v, v \rangle$. \mathcal{L} beschreibt die *kinetische* Energie und $\Phi(x) = \int_{t_1}^{t_2} \frac{1}{2} \langle \dot{x}(t), \dot{x}(t) \rangle dt$ die Wirkung von x .

(b) $\mathcal{L}(q, v, t) = \|v\|$. Dann beschreibt $\Phi(x) = \int_{t_1}^{t_2} \|\dot{x}(t)\| dt$ die Länge der Kurve x . (siehe Kap. 17.1)

Definition 16.10.2. Sei $x : [t_1, t_2] \rightarrow \mathbb{R}^n$ eine Kurve in A . Eine *Variation* von x ist eine 2-mal stetig differenzierbare Abbildung $\alpha : (-\epsilon, +\epsilon) \times [t_1, t_2] \rightarrow \mathbb{R}^n$ mit $(s, t) \mapsto \alpha_s(t) := \alpha(s, t)$ und $\alpha_s \in A \forall s$ sowie $\alpha_0(t) = x(t) \forall t$. Das Vektorfeld $Z(t) = \left. \frac{\partial}{\partial s} \right|_{s=0} \alpha_s(t)$ heißt *Variationsvektorfeld* der Variation α .

Bemerkung. Da $\alpha_s \in A$, gilt $Z(t_1) = Z(t_2) = 0$.

Definition 16.10.3. $x \in A$ heißt *kritischer Punkt* von Φ , falls

$$\left. \frac{d}{ds} \right|_{s=0} \Phi(\alpha_s) = 0$$

für alle Variationen $\alpha : (-\epsilon, +\epsilon) \times [t_1, t_2] \rightarrow \mathbb{R}^n$ von x gilt.

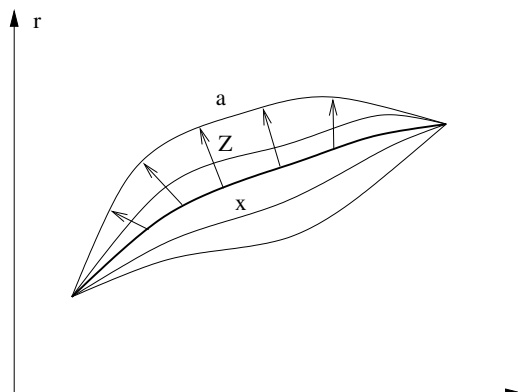


Abbildung 16.18: Variation einer Kurve

Satz 16.10.4. Sei Φ ein Lagrangesches Funktional auf A . $x \in A$ ist kritischer Punkt von Φ genau dann, falls die Gleichung (sog. Euler-Lagrange-Gleichung)

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial v}(x, \dot{x}, t) \right) = \frac{\partial \mathcal{L}}{\partial q}(x, \dot{x}, t)$$

erfüllt ist, wobei $\frac{\partial \mathcal{L}}{\partial v} = \left(\frac{\partial \mathcal{L}}{\partial v_1}, \dots, \frac{\partial \mathcal{L}}{\partial v_n} \right)$, $\frac{\partial \mathcal{L}}{\partial q}$ entsprechend (vgl. Implizite-Funktionen-Satz).

Beweis. Es gilt

$$\begin{aligned} \frac{d}{ds} \Big|_{s=0} \phi(\alpha_s) &= \frac{d}{ds} \Big|_{s=0} \int_{t_1}^{t_2} \mathcal{L}(\alpha_s(t), \dot{\alpha}_s(t), t) dt \\ &= \int_{t_1}^{t_2} \left\langle \frac{\partial \mathcal{L}}{\partial q}(\alpha_0(t), \dot{\alpha}_0(t), t), \frac{\partial}{\partial s} \Big|_{s=0} \alpha_s(t) \right\rangle + \left\langle \frac{\partial \mathcal{L}}{\partial p}(\alpha_0(t), \dot{\alpha}_0(t), t), \frac{\partial}{\partial s} \Big|_{s=0} \frac{\partial}{\partial t} \alpha_s(t) \right\rangle dt \\ &= \int_{t_1}^{t_2} \left\langle \frac{\partial \mathcal{L}}{\partial q}(x(t), \dot{x}(t), t), Z(t) \right\rangle + \left\langle \frac{\partial \mathcal{L}}{\partial p}(x(t), \dot{x}(t), t), \dot{Z}(t) \right\rangle dt. \end{aligned}$$

Durch partielle Integration erhalten wir:

$$\int_{t_1}^{t_2} \left\langle \frac{\partial \mathcal{L}}{\partial v}(x(t), \dot{x}(t), t), \dot{Z}(t) \right\rangle dt = \left\langle \frac{\partial \mathcal{L}}{\partial v}(x(t), \dot{x}(t), t), Z(t) \right\rangle \Big|_{t_1}^{t_2} - \int_{t_1}^{t_2} \left\langle \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial v}(x(t), \dot{x}(t), t), Z(t) \right\rangle dt.$$

Damit ist $x \in A$ genau dann ein kritischer Punkt, falls

$$\frac{d}{ds} \Big|_{s=0} \phi(x_s) = \int_{t_1}^{t_2} \left\langle \frac{\partial \mathcal{L}}{\partial q}(x(t), \dot{x}(t), t) - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial v}(x(t), \dot{x}(t), t), Z(t) \right\rangle dt = 0$$

für alle Variationsvektorfelder $Z(t)$ von x gilt. \square

Der Satz folgt nun aus dem folgenden Lemma.

Lemma 16.10.5. *Ist $f : [t_1, t_2] \rightarrow \mathbb{R}^n$ eine stetige Funktion und gilt*

$$\int_{t_1}^{t_2} \langle f(t), z(t) \rangle dt = 0$$

für alle stetig differenzierbaren Funktionen $z : [t_1, t_2] \rightarrow \mathbb{R}^n$ mit $z(t_1) = z(t_2) = 0$, dann ist f identisch null.

Beweis. Wir skizzieren den Beweis für $n = 1$. Ist $f(t_0) \neq 0$ für ein $t_0 \in (t_1, t_2)$, so gibt es wegen der Stetigkeit von f ein Intervall $[t_0 - \epsilon, t_0 + \epsilon]$, auf dem f keine Nullstellen besitzt. Wähle eine stetig differenzierbare Funktion $z : [t_1, t_2] \rightarrow \mathbb{R}$ mit $z(t_0) = f(t_0)$ so, dass z auf $(t_0 - \epsilon, t_0 + \epsilon)$ keine Nullstellen besitzt, aber außerhalb dieses Intervalls identisch null ist. Dann gilt

$$\int_{t_1}^{t_2} \langle f(t), z(t) \rangle dt > 0.$$

Der allgemeine Fall sei als Übung berlassen. □

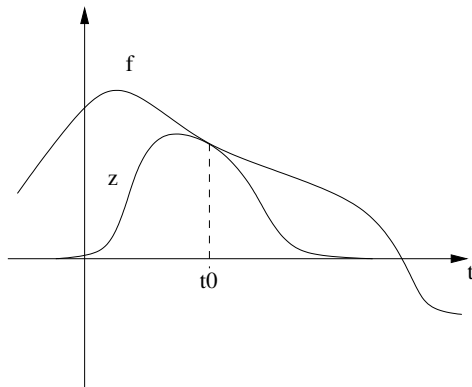


Abbildung 16.19: zu Lemma 16.10.5

Beispiele.

(a) Sei $\mathcal{L}(q, v, t) = \frac{1}{2} \langle v, v \rangle$. Dann gilt:

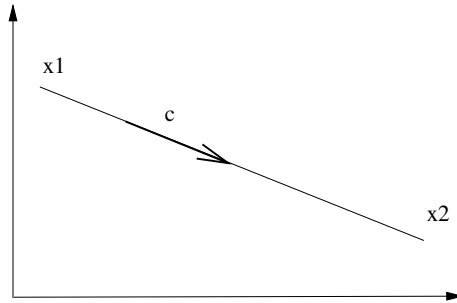
$$\frac{\partial \mathcal{L}}{\partial v}(q, v, t) = v \quad \text{und} \quad \frac{\partial \mathcal{L}}{\partial q}(q, v, t) = 0.$$

Ist x eine Lösung der Euler-Lagrange-Gleichung, so gilt

$$0 = \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial v}(x, \dot{x}, t) = \ddot{x}(t).$$

Also folgt

$$\dot{x}(t) = c \quad \text{und somit} \quad x(t) = c \cdot t + b \quad \text{mit Konstanten } b, c \in \mathbb{R}^n.$$

Abbildung 16.20: Gerade durch $x_1 = x(t_1)$ und $x_2 = x(t_2)$

(b) Sei $\mathcal{L}(q, v, t) = \|v\|$ mit $v \in \mathbb{R}^n \setminus \{0\}$. Dann gilt

$$\frac{\partial \mathcal{L}}{\partial v} = \frac{v}{\|v\|} \quad \text{und} \quad \frac{\partial \mathcal{L}}{\partial q}(q, v, t) = 0.$$

x ist eine Lösung der Euler-Lagrange-Gleichung, falls

$$\frac{\partial \mathcal{L}}{\partial v}(x, \dot{x}, t) = \frac{\dot{x}}{\|\dot{x}\|} = c \quad \text{für eine Konstante } c \in \mathbb{R}^n \text{ mit } \|c\| = 1,$$

d.h. falls die Richtung des Geschwindigkeitsvektors fest ist. Also sind die Lösungen von der Form

$$x(t) = c \cdot s(t) + b$$

für eine Parametrisierung $s : (t_1, t_2) \rightarrow \mathbb{R}$ mit $\dot{s} \neq 0$, d.h. die Lösungen sind Geraden, die x_1 mit x_2 verbinden.

Die Geschwindigkeit $\|\dot{x}(t)\| = \dot{s}(t)$, mit der die Gerade durchlaufen wird, bleibt unbestimmt. Der Grund dafür ist, dass das Längenfunktional

$$\Phi(x) = \int_{t_1}^{t_2} \|\dot{x}(t)\| dt =: L(x)$$

nicht von der Parametrisierung der Kurve abhängt, jedoch ist die Wirkung

$$\frac{1}{2} \int_{t_1}^{t_2} \langle \dot{x}(t), \dot{x}(t) \rangle dt$$

parameterabhängig.

Satz 16.10.6 (Hamiltons Prinzip der kleinsten Wirkung). Sei

$$\ddot{x}(t) = -\frac{dU}{dq}(x(t))$$

ein konservatives mechanisches System mit potentieller Energie $U : \mathbb{R}^n \rightarrow \mathbb{R}$. Ist $\mathcal{L}(q, v) = T(v) - U(q)$, wobei $T(v) = \frac{1}{2} \langle v, v \rangle$ die kinetische Energie bezeichnet, so stimmen die kritischen

Punkte des Lagrangeschen Funktionals

$$\Phi(x) = \int_{t_1}^{t_2} \mathcal{L}(x, \dot{x}) dt$$

mit den Lösungen der Newtonschen Bewegungsgleichungen überein.

Beweis. Aus der Definition von \mathcal{L} folgt

$$\frac{\partial \mathcal{L}}{\partial q}(q, v) = -\frac{\partial U}{\partial q} \quad \text{und} \quad \frac{\partial \mathcal{L}}{\partial v} = v.$$

Also folgt aus der Euler-Lagrange-Gleichung:

$$\ddot{x}(t) = \frac{d}{dt} \dot{x}(t) = \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial v}(x(t), \dot{x}(t)) = \frac{\partial \mathcal{L}}{\partial q}(x(t), \dot{x}(t)) = -\frac{dU}{dq}(x(t)).$$

□

Kapitel 17

Differentialformen 1. Grades (Pfaffsche Formen) und Kurvenintegrale

17.1 Kurven und ihre Länge

Definition 17.1.1. Sei $(X, \|\cdot\|)$ ein normierter Vektorraum, $\gamma : [a, b] \rightarrow X$ eine stetige Kurve und $Z := (t_0, t_1, \dots, t_n)$ eine Zerlegung von $[a, b]$, d.h. $a = t_0 < t_1 < \dots < t_n = b$. Dann heißt

$$L_Z(\gamma) := \sum_{j=1}^n \|\gamma(t_j) - \gamma(t_{j-1})\|$$

die Länge des Sehnepolygons von γ bezüglich der Zerlegung Z . Die Kurve $\gamma : [a, b] \rightarrow X$

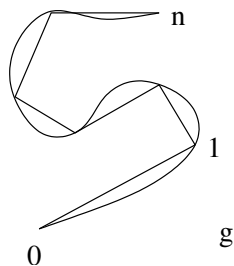


Abbildung 17.1: rektifizierbare Kurve und Sehnepolygon

heißt rektifizierbar, falls

$$L(\gamma, [a, b]) := \sup\{L_Z(\gamma) \mid Z \text{ Zerlegung von } [a, b]\}$$

endlich ist. Ist γ rektifizierbar, so heißt $L(\gamma) = L(\gamma, [a, b])$ die *Länge* von γ .

Bemerkung. Ist $\gamma : [a, b] \rightarrow X$ eine rektifizierbare Kurve und $c \in [a, b]$, so gilt:

$$L(\gamma, [a, b]) = L(\gamma, [a, c]) + L(\gamma, [c, b]).$$

Satz 17.1.2. Sei $\gamma : [a, b] \rightarrow X$ eine stetig differenzierbare Kurve. Dann ist γ rektifizierbar und es gilt:

$$L(\gamma) = \int_a^b \|\dot{\gamma}(t)\| dt.$$

Beweis. Wir zeigen zunächst die Rektifizierbarkeit. Sei $Z = (t_0, \dots, t_n)$ eine beliebige Zerlegung von $[a, b]$. Dann gilt:

$$\begin{aligned} L_Z(\gamma) &= \sum_{j=1}^n \|\gamma(t_j) - \gamma(t_{j-1})\| = \sum_{j=1}^n \left\| \int_{t_{j-1}}^{t_j} \dot{\gamma}(t) dt \right\| \\ &\leq \sum_{j=1}^n \int_{t_{j-1}}^{t_j} \|\dot{\gamma}(t)\| dt = \int_a^b \|\dot{\gamma}(t)\| dt. \end{aligned}$$

Also gilt:

$$L(\gamma) = L(\gamma, [a, b]) \leq \int_a^b \|\dot{\gamma}(t)\| dt.$$

Wir zeigen nun:

$$t \mapsto L(\gamma, [a, t]) =: l(t)$$

ist auf $[a, b]$ differenzierbar mit

$$l'(t) = \|\dot{\gamma}(t)\|.$$

Dann folgt aus dem Hauptsatz der Differential- und Integralrechnung

$$L(\gamma) = L(\gamma, [a, b]) = l(b) - l(a) = \int_a^b \|\dot{\gamma}(t)\| dt.$$

Sei also $t_0 \in [a, b]$ und $t_0 < t \leq b$. Dann gilt:

$$\left\| \frac{\gamma(t) - \gamma(t_0)}{t - t_0} \right\| \leq \frac{L(\gamma, [t_0, t])}{t - t_0} = \frac{L(\gamma, [a, t]) - L(\gamma, [a, t_0])}{t - t_0} \leq \frac{1}{t - t_0} \int_{t_0}^t \|\dot{\gamma}(s)\| ds.$$

Da

$$\|\dot{\gamma}(t_0)\| = \lim_{t \rightarrow t_0} \left\| \frac{\gamma(t) - \gamma(t_0)}{t - t_0} \right\| = \lim_{t \rightarrow t_0} \frac{1}{t - t_0} \int_{t_0}^t \|\dot{\gamma}(s)\| ds,$$

folgt die Differenzierbarkeit von $l(t) = L(\gamma, [a, t])$ mit $l'(t_0) = \|\dot{\gamma}(t_0)\|$. \square

Bemerkung. Sei $(\mathbb{R}^n, \|\cdot\|)$ der euklidische Raum mit $\|x\| = \sqrt{x_1^2 + \dots + x_n^2}$ und $\gamma = (\gamma_1, \dots, \gamma_n) : [a, b] \rightarrow \mathbb{R}^n$ eine stetig differenzierbare Kurve, so gilt:

$$L(\gamma) = \int_a^b \sqrt{\dot{\gamma}_1(t)^2 + \dots + \dot{\gamma}_n(t)^2} dt.$$

17.2 Differentialformen 1. Grades und Kurvenintegrale

Nun wollen wir uns mit einem wichtigen Integralbegriff beschäftigen, nämlich dem Kurvenintegral. Sei $U \subset \mathbb{R}^n$ offen und $f : U \rightarrow \mathbb{R}$ differenzierbar, so definiert ihr Differential

$$df : U \rightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R})$$

eine Abbildung von U in die Linearformen $\mathcal{L}(\mathbb{R}^n, \mathbb{R})$. (Es wird sich später als zweckmäßig erweisen, die Notation df statt wie bisher Df für das Differential zu verwenden). Ist e_1, \dots, e_n die Standardbasis des \mathbb{R}^n und $v = \sum_{i=1}^n v_i e_i \in \mathbb{R}^n$, so ist

$$df(x)v = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x)v_i = \langle \text{grad } f(x), v \rangle,$$

wobei \langle, \rangle das Standardskalarprodukt auf \mathbb{R}^n darstellt.

Definition 17.2.1. Eine *Differentialform 1. Grades* (auch *1-Form* oder *Pfaffsche Form*) auf $U \subset \mathbb{R}^n$ ist eine Abbildung

$$\omega : U \rightarrow L(\mathbb{R}^n, \mathbb{R}).$$

Bemerkungen.

- (a) Ist \langle, \rangle das Standard-Skalarprodukt auf \mathbb{R}^n und $v = \sum_{i=1}^n v_i e_i$, so existiert ein Vektorfeld $F : U \rightarrow \mathbb{R}^n$ mit

$$\omega(x)v = \langle F(x), v \rangle = \sum_{i=1}^m F_i(x)v_i.$$

- (b) Jede 1-Form lässt sich wie folgt als Linearkombination von Differentialen schreiben: Sei $x_i : \mathbb{R}^n \rightarrow \mathbb{R}$ mit $x_i(y_1, \dots, y_n) = y_i$ die Projektion auf die i -te Koordinate. Dann ist

$$dx_i(y)v = \sum_{j=1}^n \frac{\partial x_i}{\partial y_j} v_j = v_i.$$

Es gilt:

$$\omega(y)v = \sum_{i=1}^n F_i(y)v_i = \sum_{i=1}^n F_i(y)dx_i(y)v$$

oder kurz

$$\omega = \sum_{i=1}^n F_i dx_i.$$

- (c) Sei $U \subset \mathbb{R}^n$ offen. Eine 1-Form $\omega : U \rightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R})$ mit $\omega = \sum_{i=1}^n F_i dx_i$ heißt k -mal stetig differenzierbar, falls die Koeffizienten $F_i : U \rightarrow \mathbb{R}$ k -mal stetig differenzierbar sind.

Beispiele.

- (a) $\omega = \frac{-y}{x^2+y^2} dx + \frac{x}{x^2+y^2} dy$ ist eine 1-Form auf $\mathbb{R}^2 \setminus \{0\}$. ω heißt auch *Windungsform*.

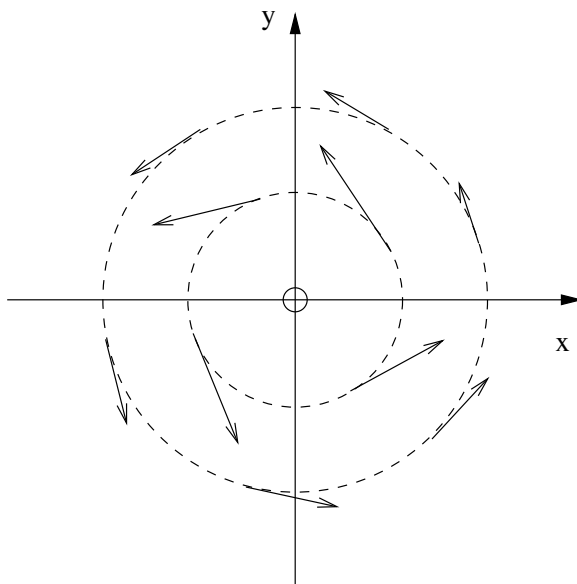


Abbildung 17.2: Windungsform

- (b) $\omega = -\frac{1}{r^3}(x dx + y dy + z dz)$ mit $r = \sqrt{x^2 + y^2 + z^2}$ heißt *Gravitationsform*. Das zugehörige Vektorfeld F ist gegeben durch $F(x, y, z) = -\frac{(x, y, z)}{r^3}$ und hat die Eigenschaft $\|F(x, y, z)\| = \frac{1}{r^2}$ (*Zentralfeld*).

Eine wichtige Frage ist, wann eine 1-Form $\omega : U \rightarrow L(\mathbb{R}^n, \mathbb{R})$ sich als Ableitung schreiben lässt, d.h. wann eine Abbildung $f : U \rightarrow \mathbb{R}$ existiert mit $df = \omega$. In der Physik nennt man ein solches f auch Potential.

Definition 17.2.2. Eine stetige 1-Form $\omega : U \rightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R})$ heißt *exakt*, falls eine stetig differenzierbare Funktion $f : U \rightarrow \mathbb{R}$ existiert mit $df = \omega$. Dann heißt f *Stammfunktion* oder *Potential* von ω .

Bemerkungen.

- (a) Wegen obiger Bemerkung hat jede stetige 1-Form $\omega : U \rightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R})$ die Form

$$\omega(x)v = \langle F(x), v \rangle,$$

wobei $F : U \rightarrow \mathbb{R}^n$ ein stetiges Vektorfeld auf U ist. Ist nun ω exakt, so existiert eine stetig differenzierbare Funktion $f : U \rightarrow \mathbb{R}$ mit $df = \omega$, d.h.

$$\omega(x)v = \langle F(x), v \rangle = df(x)v = \langle \text{grad } f(x), v \rangle.$$

Also ist ω genau dann exakt, falls das Vektorfeld $F : U \rightarrow \mathbb{R}^n$ ein Gradientenfeld ist.

- (b) Für $n = 1$ ist jede 1-Form exakt, denn jede stetige Funktion besitzt eine Stammfunktion.
 (c) Die Gravitationsform ω besitzt

$$f(x, y, z) = \frac{1}{\sqrt{x^2 + y^2 + z^2}}$$

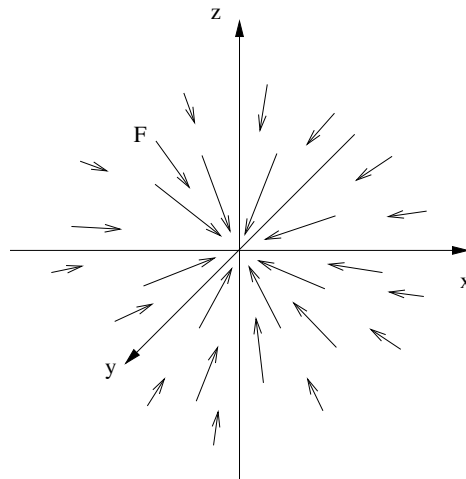


Abbildung 17.3: Gravitationsform

als Potential, denn

$$\frac{\partial f}{\partial x} = -\frac{1}{2} \frac{2x}{(x^2 + y^2 + z^2)^{3/2}}.$$

Die Windungsform besitzt, wie wir bald sehen werden, kein Potential.

Es ist naheliegend, ein Potential durch Integration zu konstruieren. Dazu müssen wir zunächst das Integral einer 1-Form erklären.

Definition 17.2.3 (Kurvenintegral). Sei $U \subset \mathbb{R}^n$, $\omega : U \rightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R})$ eine stetige 1-Form und $\gamma : [a, b] \rightarrow U$ eine stetig differenzierbare Kurve. Dann heißt

$$\int_{\gamma} \omega := \int_a^b \omega(\gamma(t)) \dot{\gamma}(t) dt$$

das Integral von ω längs der Kurve γ .

Bemerkungen.

- (a) Ist $\gamma : [a, b] \rightarrow U$ eine stetige Kurve, die stückweise stetig differenzierbar ist, d.h. existiert eine Zerlegung $a = t_0 < t_1 \dots < t_k = b$, so dass $\gamma : (t_i, t_{i+1}) \rightarrow U$ stetig differenzierbar ist, so ist

$$\int_{\gamma} \omega := \sum_{i=1}^k \int_{t_{i-1}}^{t_i} \omega(\gamma(t)) \dot{\gamma}(t) dt.$$

- (b) Ist $F : U \rightarrow \mathbb{R}^n$ ein Kraftfeld und $\omega(x)v = \langle F(x), v \rangle$ die zugehörige 1-Form, so beschreibt das Kurvenintegral

$$\int_{\gamma} \omega$$

die längs des Weges γ geleistete Arbeit.

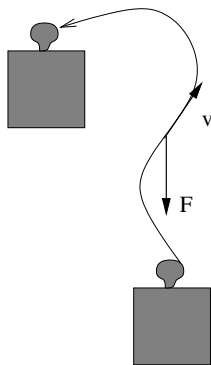


Abbildung 17.4: Arbeit längs einer Kurve

Wir werden nun sehen, dass das Kurvenintegral sich nicht unter orientierungserhaltenden Reparametrisierungen des Weges ändert.

Definition 17.2.4. Sei $\gamma : [a, b] \rightarrow \mathbb{R}^n$ eine Kurve und $\varphi : [c, d] \rightarrow [a, b]$ stetig differenzierbar und bijektiv. Dann heißt

$$\tilde{\gamma}(s) = \gamma \circ \varphi(s)$$

Reparametrisierung von γ . Die Reparametrisierung heißt *orientierungserhaltend* bzw. *orientierungsumkehrend*, falls $\varphi'(t) > 0$ bzw. $\varphi'(t) < 0$ für alle $t \in [c, d]$ gilt.

Bemerkung. Es gilt: Bild $\tilde{\gamma} =$ Bild γ , jedoch werden diese Kurven mit einer anderen Geschwindigkeit durchlaufen. Eine orientierungserhaltende Reparametrisierung gewährleistet, dass die Kurven in gleicher Richtung durchlaufen werden.

Satz 17.2.5. Sei $\omega : U \rightarrow L(\mathbb{R}^n, \mathbb{R})$ eine stetig differenzierbare 1-Form und $\tilde{\gamma} = \gamma \circ \varphi : [c, d] \rightarrow U$ eine orientierungserhaltende Reparametrisierung von $\gamma : [a, b] \rightarrow U$. Dann gilt:

$$\int_{\tilde{\gamma}} \omega = \int_{\gamma} \omega.$$

Beweis. Aus der Substitutionsregel folgt:

$$\int_{\tilde{\gamma}} \omega = \int_c^d \omega(\tilde{\gamma}(s)) \dot{\tilde{\gamma}}(s) ds = \int_c^d \omega(\gamma \circ \varphi(s)) \dot{\gamma}(\varphi(s)) \cdot \varphi'(s) ds = \int_a^b \omega(\gamma(t)) \dot{\gamma}(t) dt.$$

□

Bemerkung. Ist $\varphi : [c, d] \rightarrow [a, b]$ orientierungsumkehrend, d.h. $\varphi'(t) < 0$, so gilt

$$\int_{\tilde{\gamma}} \omega = - \int_{\gamma} \omega.$$

Satz 17.2.6. Sei $U \subset \mathbb{R}^n$ eine offene wegzusammenhängende Menge und $\omega : U \rightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R})$ eine stetige 1-Form. Dann ist ω genau dann exakt, falls

$$\int_{\gamma_1} \omega = \int_{\gamma_2} \omega$$

für alle stückweise stetig differenzierbaren Kurven $\gamma_1 : [a, b] \rightarrow U, \gamma_2 : [c, d] \rightarrow U$, deren Anfangspunkte und Endpunkte übereinstimmen, d.h.

$$\gamma_1(a) = \gamma_2(c) \text{ und } \gamma_1(b) = \gamma_2(d).$$

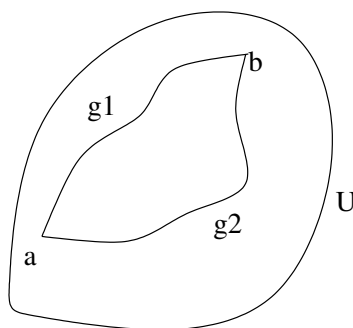


Abbildung 17.5: zu Satz 17.2.6

Beweis. Sei ω exakt. Dann existiert eine stetig differenzierbare Funktion $f : U \rightarrow \mathbb{R}$ mit $df = \omega$. Sei $\gamma : [a, b] \rightarrow U$ eine differenzierbare Kurve, so gilt:

$$\int_{\gamma} df = \int_a^b df(\gamma(s))\dot{\gamma}(s)ds = \int_a^b \frac{d}{ds} f \circ \gamma(s)ds = f(\gamma(b)) - f(\gamma(a)).$$

Ist γ für eine Zerlegung $a = t_0 < t_1 \dots < t_k = b$ stückweise differenzierbar, so folgt ebenfalls:

$$\int_{\gamma} df = \sum_{i=1}^k f(\gamma(t_i)) - f(\gamma(t_{i-1})) = f(\gamma(b)) - f(\gamma(a)).$$

Insbesondere hängt $\int_{\gamma} df$ nur von den Anfangs- und Endpunkten ab.

Sei nun umgekehrt

$$\int_{\gamma_1} \omega = \int_{\gamma_2} \omega$$

für alle stückweise differenzierbaren Kurven mit gleichen Anfangspunkten und Endpunkten. Sei $p \in U$ fest gewählt. Da U wegzusammenhängend ist, existiert für jedes $x \in U$ eine stetig differenzierbare Kurve $\gamma_x : [a, b] \rightarrow U$ mit $\gamma_x(a) = p$ und $\gamma_x(b) = x$. Setze

$$f(x) = \int_{\gamma_x} \omega.$$

Nach Voraussetzung ist $f : U \rightarrow \mathbb{R}$ wohldefiniert. Für jedes $x_0 \in U$ ist dann f in x_0 differenzierbar, was man wie folgt beweist.

Wähle $\delta > 0$, so dass für $\|h\| < \delta$ die Gerade $c(t) = x_0 + th$ für $0 \leq t \leq 1$ ganz in U verläuft. Dann gilt

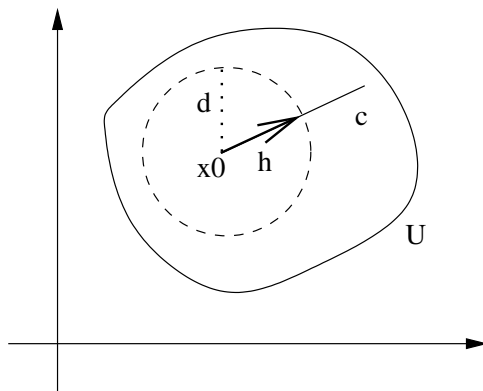


Abbildung 17.6: zum Beweis von Satz 17.2.6

$$f(x_0 + h) - f(x_0) = \int_c \omega = \int_0^1 \omega(c(t)) \dot{c}(t) dt = \int_0^1 \omega(x_0 + th) h dt$$

und somit

$$f(x_0 + h) - f(x_0) - \omega(x_0)h = \int_0^1 (\omega(x_0 + th) - \omega(x_0)) h dt.$$

Da ω stetig ist, existiert eine stetige Funktion $F : U \rightarrow \mathbb{R}^n$ mit

$$\langle F(x), h \rangle = \omega(x)h.$$

Also folgt:

$$\begin{aligned} |f(x_0 + h) - f(x_0) - \omega(x_0)h| &\leq \int_0^1 |\langle F(x_0 + th) - F(x_0), h \rangle| dt \\ &\leq \int_0^1 \|F(x_0 + th) - F(x_0)\| \|h\| dt \\ &= \|h\| \int_0^1 \|F(x_0 + th) - F(x_0)\| dt. \end{aligned}$$

Wegen der Stetigkeit von F folgt

$$\lim_{h \rightarrow 0} \int_0^1 \|F(x_0 + th) - F(x_0)\| dt = 0$$

und somit ist f differenzierbar mit $df(x_0) = \omega(x_0)$.

□

Bemerkung. Hängt $\int_{\gamma} \omega$ nur von den Anfangs- und Endpunkten der Kurve γ ab, so gilt für alle geschlossenen Kurven γ (d.h. $\gamma(a) = \gamma(b) = p$)

$$\int_{\gamma} \omega = 0.$$

Denn $\int_{\gamma} \omega = \int_{\gamma_p} \omega = 0$, wobei $\gamma_p(t) = p$ die konstante Kurve ist.

Ist umgekehrt $\int_{\gamma} \omega = 0$ für alle geschlossenen stückweise differenzierbaren Kurven, so gilt: $\int_{\gamma_1} \omega = \int_{\gamma_2} \omega$, falls γ_1, γ_2 gemeinsame Anfangs- und Endpunkte besitzen. Denn sei γ_2^- die umge-

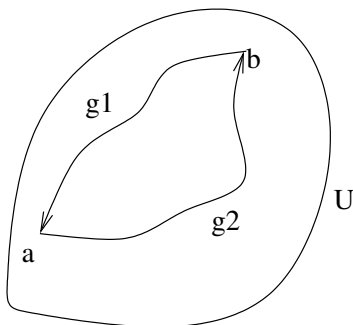


Abbildung 17.7: γ_2^- wird umgekehrt durchlaufen

kehrt durchlaufene Kurve γ_2 , so ist $\int_{\gamma_2^-} \omega = - \int_{\gamma_2} \omega$. Da $\int_{\gamma_1} \omega + \int_{\gamma_2^-} \omega = 0$, folgt die Behauptung.

Satz 17.2.7. Sei $\omega = \sum_{i=1}^n F_i(x)dx_i$ eine stetig differenzierbare exakte 1-Form. Dann folgt:

$$\frac{\partial F_j}{\partial x_i} = \frac{\partial F_i}{\partial x_j} \tag{17.1}$$

(Integrabilitätsbedingungen).

Beweis. Ist ω exakt, so existiert ein differenzierbares f mit $df(x) = \omega(x)$ und somit $\frac{\partial f}{\partial x_i} = F_i(x)$. Da die F_i stetig differenzierbar sind, ist daher f zweimal stetig differenzierbar. Insbesondere folgt mit dem Lemma von Schwarz

$$\frac{\partial F_j}{\partial x_i} = \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial F_i}{\partial x_j}.$$

□

Definition 17.2.8. Eine stetig differenzierbare 1-Form $\omega = \sum_{i=1}^n F_i(x)dx_i$ heißt *geschlossen*, falls sie die Integrabilitätsbedingungen (17.1) erfüllt.

Bemerkungen.

- (a) Sei $U \subset \mathbb{R}^3$ eine offene Teilmenge und $F : U \rightarrow \mathbb{R}^3$ ein differenzierbares Vektorfeld. Dann heißt

$$\operatorname{rot} F := \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \frac{\partial}{\partial x_3} \end{pmatrix} \times \begin{pmatrix} F_1 \\ F_2 \\ F_3 \end{pmatrix} = \begin{pmatrix} \frac{\partial F_3}{\partial x_2} - \frac{\partial F_2}{\partial x_3} \\ \frac{\partial F_1}{\partial x_3} - \frac{\partial F_3}{\partial x_1} \\ \frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \end{pmatrix}$$

die *Rotation* des Vektorfeldes F . Also ist eine 1-Form $\omega = \sum_{i=1}^3 F_i(x) dx_i$ genau dann geschlossen, falls die Rotation des Vektorfeldes $F = (F_1, F_2, F_3)$ verschwindet.

- (b) Wir haben also gesehen: Ist eine 1-Form exakt, so ist sie auch geschlossen. Die Umkehrung gilt nicht, wie das Beispiel der Windungsform $\omega : \mathbb{R}^2 \setminus \{0\} \rightarrow L(\mathbb{R}^2, \mathbb{R})$ mit $\omega = \frac{-y}{x^2+y^2} dx + \frac{x}{x^2+y^2} dy$ zeigt. Denn

$$\frac{\partial}{\partial y} \frac{-y}{x^2+y^2} = \frac{-(x^2+y^2) - 2y(-y)}{(x^2+y^2)^2} = \frac{y^2 - x^2}{(x^2+y^2)^2}$$

und

$$\frac{\partial}{\partial x} \frac{x}{x^2+y^2} = \frac{x^2+y^2 - 2x^2}{(x^2+y^2)^2} = \frac{y^2 - x^2}{(x^2+y^2)^2}.$$

Sei jedoch $c : [\sigma, 2\pi] \rightarrow \mathbb{R}^2 \setminus \{0\}$ der parametrisierte Einheitskreis mit $c(t) = (\cos t, \sin t)$, so gilt

$$\int_c \omega = \int_0^{2\pi} -\sin t(-\sin t) + \cos^2 t dt = 2\pi \neq 0,$$

d.h. ω ist nicht exakt nach Satz 17.2.6 und nachfolgender Bemerkung. Es stellt sich aber heraus, dass unter einer Zusatzannahme an U die notwendigen Integrierbarkeitsbedingungen auch hinreichend sind.

Definition 17.2.9. Eine Menge $A \subset \mathbb{R}^n$ heißt *sternförmig*, falls ein Punkt $p \in A$ existiert, so dass für jedes $x \in A$ die Verbindungsgerade $c(t) = p + t(x - p)$, $t \in [0, 1]$ ganz in A verläuft.

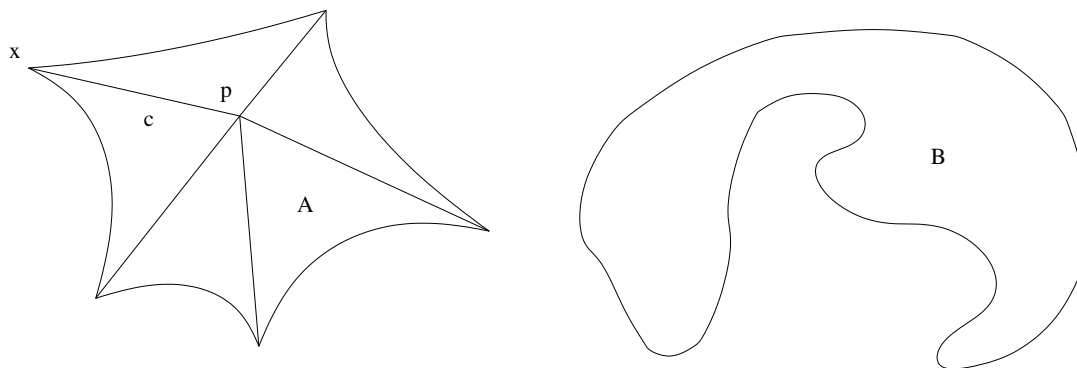


Abbildung 17.8: A ist sternförmig, B nicht

Satz 17.2.10 (Lemma von Poincaré). Sei U eine offene sternförmige Menge und $\omega = \sum_{i=1}^n F_i dx_i$ eine auf U definierte stetig-differenzierbare 1-Form, die die Integrabilitätsbedingung

$$\frac{\partial F_i}{\partial x_j} = \frac{\partial F_j}{\partial x_i}$$

erfüllt. Dann ist ω exakt.

Beweis. Wir nehmen an, dass U bezüglich $0 \in \mathbb{R}^n$ sternförmig ist, d.h. $c(t) = tx \in U$ für alle $x \in U$ und $t \in [0, 1]$. Definiere

$$f(x) = \int_c \omega = \int_0^1 \omega(c(t)) \dot{c}(t) dt = \int_0^1 \sum_{i=1}^n F_i(c(t)) \dot{c}_i(t) dt = \int_0^1 \sum_{i=1}^n F_i(tx) x_i dt.$$

Wir zeigen:

$$\frac{\partial f}{\partial x_j}(x) = F_j(x). \tag{17.2}$$

Wegen der Stetigkeit der partiellen Ableitung ist dann auch $f : U \rightarrow \mathbb{R}$ differenzierbar und es gilt: $df(x) = \sum_{i=1}^n \frac{\partial f}{\partial x_i} dx_i = \omega$.

Beweis von (17.2):

Da die Funktionen F_i stetig differenzierbar sind, ist

$$\frac{\partial}{\partial x_j} \sum_{i=1}^n F_i(tx) x_i = \sum_{i=1}^n \frac{\partial}{\partial x_j} F_i(tx) x_i t + F_j(tx)$$

stetig in t , also insbesondere auf $[0, 1]$ integrierbar.

Daher ist f nach x_j partiell differenzierbar (Differentiation unter dem Integralzeichen!) und es gilt:

$$\frac{\partial f}{\partial x_j}(x) = \int_0^1 \left(\sum_{i=1}^n \frac{\partial}{\partial x_j} F_i(tx) x_i \right) t dt + \int_0^1 F_j(tx) dt.$$

Wegen

$$g_j(t) := \sum_{i=1}^n \frac{\partial}{\partial x_j} F_i(tx) x_i = \sum_{i=1}^n \frac{\partial}{\partial x_i} F_j(tx) x_i = \frac{d}{dt} F_j(tx)$$

ist $G_j(t) = F_j(tx)$ eine Stammfunktion von g_j . Mittels partieller Integration erhalten wir:

$$\frac{\partial f}{\partial x_j}(x) = F_j(tx) \cdot t \Big|_0^1 - \int_0^1 F_j(tx) dt + \int_0^1 F_j(tx) dt = F_j(x).$$

□

Definition 17.2.11. Sei $A \subset \mathbb{R}^n$. Zwei Kurven $\gamma_0, \gamma_1 : [a, b] \rightarrow A$ mit gemeinsamem Anfangspunkt q_a und gemeinsamem Endpunkt q_b heißen *homotop*, falls eine stetige Abbildung $H : [0, 1] \times [a, b] \rightarrow A$, $(s, t) \mapsto H(s, t) =: H_s(t)$ existiert (*Homotopie*), mit

$$H_0 = \gamma_0(t) \text{ und } H_1(t) = \gamma_1(t)$$

für alle $t \in [a, b]$, sowie $H_s(a) = q_a$ und $H_s(b) = q_b$ für alle $s \in [0, 1]$.

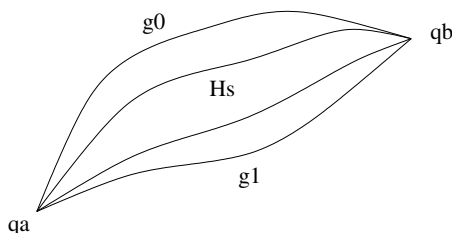


Abbildung 17.9: Homotopie

Beispiel. Seien $\gamma_0, \gamma_1 : [a, b] \rightarrow A$ zwei Kurven mit gemeinsamem Anfangspunkt und gemeinsamem Endpunkt. Liegt für jedes $t \in [a, b]$ die Verbindungsgerade von $\gamma_0(t)$ nach $\gamma_1(t)$ ganz in A , so definiert $H : [0, 1] \times [a, b] \rightarrow A$ mit

$$H(s, t) = s\gamma_1(t) + (1 - s)\gamma_0(t)$$

eine Homotopie.

Satz 17.2.12. Es sei $U \subset \mathbb{R}^n$ eine offene Menge und $\omega : U \rightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R})$ eine geschlossene 1-Form. Sind $\gamma_0, \gamma_1 : [a, b] \rightarrow U$ stetig differenzierbare homotope Kurven mit gemeinsamem Anfangspunkt q_a und gemeinsamem Endpunkt q_b , so gilt:

$$\int_{\gamma_0} \omega = \int_{\gamma_1} \omega$$

Beweis. Sei $H : [0, 1] \times [a, b] \rightarrow U$ eine Homotopie mit $H_0 = \gamma_0(t)$ und $H_1 = \gamma_1(t)$. Wegen der Stetigkeit von H kann man Zerlegungen

$$a = t_0 < t_1 < \dots < t_k = b$$

und

$$0 = s_0 < s_1 < \dots < s_l = 1$$

finden, so dass für jedes $(i, j) \in \{0, \dots, k-1\} \times \{0, \dots, l-1\}$ die Menge $H([s_i, s_{i+1}] \times [t_j, t_{j+1}])$ in einem Ball liegt, welcher ganz in U enthalten ist. Betrachte nun zu den Punkten $p_{ij} = H(s_i, t_j)$ die Geradenstücke

$$\alpha_{ij} := \overline{p_{ij}, p_{i,j+1}},$$

die $p_{i,j}$ mit $p_{i,j+1}$ verbinden, und die Geradenstücke

$$\sigma_{ij} := \overline{p_{ij}, p_{i+1,j}},$$

die $p_{i,j}$ mit $p_{i+1,j}$ verbinden. Diese Geradenstücke sind wegen der Konvexität von Bällen im gleichen Ball enthalten, der auch die Menge $H([s_j, s_{j+1}] \times [t_i, t_{i+1}])$ enthält. Da ω geschlossen ist, folgt:

$$\int_{\alpha_{ij}} \omega - \int_{\alpha_{i+1,j}} \omega = \int_{\sigma_{ij}} \omega - \int_{\sigma_{i,j+1}} \omega \quad (17.3)$$

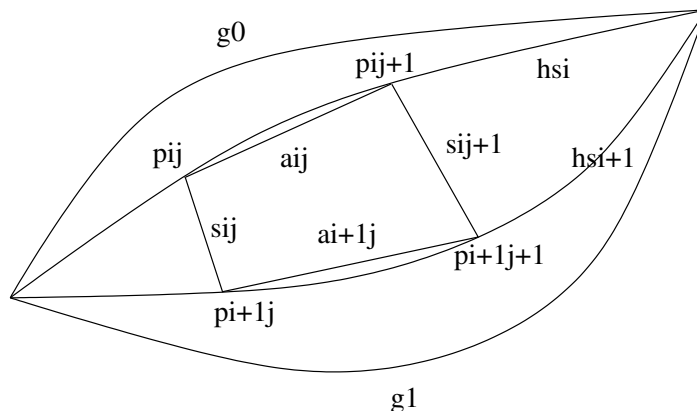


Abbildung 17.10: zum Beweis von Satz 17.2.12

Sei α_i die aus den Geradenstücken α_{ij} , $j \in \{0, l - 1\}$ zusammengesetzte Kurve. Sie sind stückweise differenzierbar und verbinden q_a mit q_b . Aus (17.3) folgt

$$\int_{\alpha_i} \omega = \int_{\alpha_{i+1}} \omega$$

und somit auch

$$\int_{\alpha_0} \omega = \int_{\alpha_k} \omega.$$

Da $\alpha_{0,j}$ und $\gamma_0|_{[t_j, t_{j+1}]}$ die gleichen Endpunkte besitzen und beide Kurven in einem Ball verlaufen, gilt:

$$\int_{\alpha_{0j}} \omega = \int_{\gamma_0|_{[t_j, t_{j+1}]}} \omega$$

und somit

$$\int_{\alpha_0} \omega = \int_{\gamma_0} \omega.$$

Genauso folgt:

$$\int_{\alpha_k} \omega = \int_{\gamma_1} \omega$$

und somit erhalten wir die Behauptung. □

Dieser Satz hat auch eine Version für geschlossene Kurven.

Definition 17.2.13. Sei $A \subset \mathbb{R}^n$. Zwei geschlossene Kurven $\gamma_0, \gamma_1 : [a, b] \rightarrow A$ heißen (*frei*) *homotop*, falls eine stetige Abbildung $H : [0, 1] \times [a, b] \rightarrow A$, $(s, t) \mapsto H(s, t) =: H_s(t)$ existiert (*freie Homotopie*) mit

$$H_0 = \gamma_0(t) \text{ und } H_1(t) = \gamma_1(t)$$

für alle $t \in [a, b]$, sowie $H_s(a) = H_s(b)$ für alle $s \in [0, 1]$.

Im Folgenden meint *homotop* für geschlossene Kurven stets *frei homotope* Kurven.

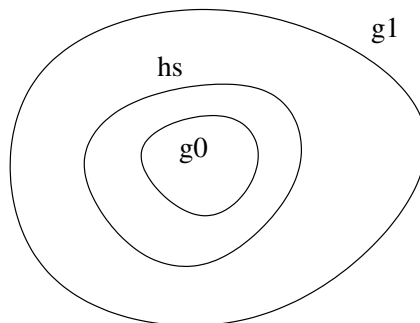
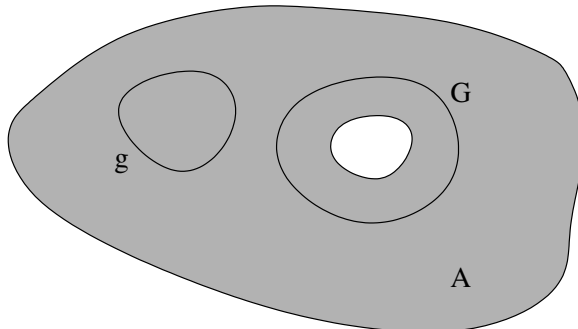


Abbildung 17.11: freie Homotopie

Satz 17.2.14. *Es sei $U \subset \mathbb{R}^n$ eine offene Menge und $\omega : U \rightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R})$ eine geschlossene 1-Form. Sind $\gamma_0, \gamma_1 : [a, b] \rightarrow U$ geschlossene, stückweise stetig differenzierbare, frei homotope Kurven, so gilt:*

$$\int_{\gamma_0} \omega = \int_{\gamma_1} \omega.$$

Definition 17.2.15. Eine Menge $A \subset \mathbb{R}^n$ heißt *einfach zusammenhängend*, wenn jede geschlossene Kurve in A *nullhomotop* ist, d.h. homotop zu einer Punktkurve ist.

Abbildung 17.12: γ ist nullhomotop, Γ nicht

Satz 17.2.16. *Es sei $U \subset \mathbb{R}^n$ eine offene einfach zusammenhängende Menge. Ist ω eine stetig differenzierbare geschlossene 1-Form auf U , so ist sie auch exakt.*

17.3 Holomorphe Funktionen

Nun wollen wir uns mit der komplexen Analysis einer Veränderlichen beschäftigen. Dieses Gebiet der Mathematik wird auch als Funktionentheorie bezeichnet. Zunächst lässt sich der Begriff der reellen Differenzierbarkeit (siehe Definition 4.1.1) unmittelbar auf komplexe Funktionen ausdehnen. Wir werden aber sehen, dass die Eigenschaften komplex differenzierbarer Funktionen erheblich von denen der reell-differenzierbaren Funktionen abweichen. Um sie zu studieren werden insbesondere auch die im letzten Abschnitt behandelten Kurvenintegrale eine wesentliche Rolle spielen.

Definition 17.3.1. Sei $U \subset \mathbb{C}$ eine offene Teilmenge und $f : U \rightarrow \mathbb{C}$ eine Funktion. Dann heißt f in $z_0 \in \mathbb{C}$ *komplex differenzierbar*, falls

$$\lim_{\substack{z \rightarrow z_0 \\ z \in U \setminus \{z_0\}}} \frac{f(z) - f(z_0)}{z - z_0} = c$$

existiert. Die Zahl c heißt die *Ableitung* von f in z_0 . Man schreibt: $f'(z_0) := c$. Ist f differenzierbar für alle $z_0 \in U$, so heißt f *holomorph*.

Bemerkungen.

- (a) Nach Bemerkung (a) nach Definition 3.4.3 ist f also genau dann komplex differenzierbar, falls der Differenzenquotient $q(z) = \frac{f(z) - f(z_0)}{z - z_0}$ sich stetig nach z_0 fortsetzen lässt, d.h. falls ein $c \in \mathbb{C}$ existiert, so dass

$$Q(z) = \begin{cases} q(z) & \text{für } z \in U \setminus \{z_0\} \\ c & \text{für } z = z_0 \end{cases}$$

stetig in z_0 ist.

- (b) f ist also auch genau dann in z_0 komplex differenzierbar, falls für jede konvergente Folge $z_n \in U \setminus \{z_0\}$ mit $\lim_{n \rightarrow \infty} z_n = z_0$ auch die Folge $q(z_n)$ konvergiert. Diese Bedingung impliziert, dass der Grenzwert der Folge $q(z_n)$ nicht von der Wahl der gegen z_0 konvergierenden Folge $z_n \in U$ abhängt.
- (c) f ist auch genau dann in z_0 komplex differenzierbar, falls eine Zahl $c \in \mathbb{C}$ und eine in z_0 stetige Funktion $r : D \rightarrow \mathbb{R}$ mit $r(z_0) = 0$ existiert, so dass

$$f(z) = f(z_0) + c(z - z_0) + r(z)(z - z_0).$$

- (d) Es ist offensichtlich, dass die Funktion $f(z) = z$ komplex differenzierbar ist. Hingegen ist auf den ersten Blick überraschend, dass die Funktion $f(z) = \bar{z}$ in keinem Punkt komplex differenzierbar ist, denn ist $z_0 \in \mathbb{C}$ und $h \in \mathbb{R}$, so gilt

$$q(z_0 + ih) = \frac{i\bar{h}}{ih} = -1$$

und

$$q(z_0 + h) = \frac{\bar{h}}{h} = 1$$

- (e) Man kann zeigen, dass die Ableitung einer holomorphen Funktion immer stetig ist (Lemma von Goursat (siehe z.B. Königsberger: Analysis 2, 2.Aufl.)). Die Stetigkeit der Ableitung wird an einigen Stellen wesentlich sein.

Satz 17.3.2. Sei $U \subset \mathbb{C}$ offen und $f : U \rightarrow \mathbb{C}$ eine Funktion mit

$$f(x + iy) = f_1(x, y) + if_2(x, y).$$

Dann sind folgende Aussagen äquivalent:

(a) f ist in $z_0 = x_0 + iy_0$ komplex differenzierbar,

(b) $F : U_{\mathbb{R}} \rightarrow \mathbb{R}^2$ mit $U_{\mathbb{R}} = \{(x, y) \in \mathbb{R}^2 \mid x + iy \in U\}$ und $F(x, y) = (f_1(x, y), f_2(x, y))$ ist in (x_0, y_0) reell differenzierbar und die Cauchy-Riemannschen Differentialgleichungen

$$\frac{\partial f_1}{\partial x}(x_0, y_0) = \frac{\partial f_2}{\partial y}(x_0, y_0) \quad \text{und} \quad \frac{\partial f_2}{\partial x}(x_0, y_0) = -\frac{\partial f_1}{\partial y}(x_0, y_0) \quad (17.4)$$

sind erfüllt.

Beweis. Zunächst ist $f : U \rightarrow \mathbb{C}$ genau dann komplex differenzierbar in z_0 , falls

$$\lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \frac{|f(z_0 + h) - f(z_0) - ch|}{|h|} = 0$$

für ein festes $c \in \mathbb{C}$. Ist $c = a + ib$ so und $h = h_1 + ih_2$ so folgt: $ch = ah_1 - bh_2 + i(bh_1 + ah_2)$. Außerdem stimmt für $z = x + iy \in \mathbb{C}$ der Betrag von z mit der euklidischen Norm von (x, y) überein, d.h. $|z| = \sqrt{x^2 + y^2} = \|(x, y)\|$. Also folgt:

$$\lim_{\substack{(h_1, h_2) \rightarrow 0 \\ (h_1, h_2) \neq 0}} \frac{\left\| F(x_0 + h_1, y_0 + h_2) - F(x_0, y_0) - \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} \right\|}{\|(h_1, h_2)\|} = 0.$$

Damit ist aber $F : U_{\mathbb{R}} \rightarrow \mathbb{R}^2$ in (x_0, y_0) differenzierbar mit

$$a = \frac{\partial f_1}{\partial x}(x_0, y_0) = \frac{\partial f_2}{\partial y}(x_0, y_0) \quad \text{und} \quad b = \frac{\partial f_2}{\partial x}(x_0, y_0) = -\frac{\partial f_1}{\partial y}(x_0, y_0).$$

Ist nun $F : U_{\mathbb{R}} \rightarrow \mathbb{R}^2$ in (x_0, y_0) differenzierbar mit $F(x, y) = (f_1(x, y), f_2(x, y))$ und gelten die Cauchy-Riemannschen Differentialgleichungen, so ist $f : U \rightarrow \mathbb{C}$ mit

$$f(x + iy) = f_1(x, y) + if_2(x, y)$$

in $x_0 + iy_0$ komplex differenzierbar und die Ableitung ist durch

$$f'(x_0 + iy_0) = \frac{\partial f_1}{\partial x}(x_0, y_0) + i \frac{\partial f_2}{\partial x}(x_0, y_0)$$

gegeben. □

Man kann wie in der reellen Analysis einer Variable die entsprechenden Regeln für komplex differenzierbare Funktionen beweisen.

Satz 17.3.3. Sei $U \subset \mathbb{C}$ offen und $f, g : U \rightarrow \mathbb{C}$ seien in z_0 komplex differenzierbar. Dann sind auch $f + g$, $f \cdot g$ und $\frac{1}{f}$ (falls $f(z_0) \neq 0$) in z_0 differenzierbar und es gilt:

- (1) $(f + g)'(z_0) = f'(z_0) + g'(z_0)$.
- (2) $(f \cdot g)'(z_0) = f'(z_0) \cdot g(z_0) + g'(z_0) \cdot f(z_0)$.
- (3) $(\frac{1}{f})'(z_0) = -\frac{f'(z_0)}{f^2(z_0)}$.

Beweis. Der Beweis ist vollkommen analog zum Beweis von Satz 4.3.1 □

Bemerkungen.

- (a) Aus der Produktregel (2) erhält man wie im reellen Fall mit Hilfe von (3) die Quotientenregel: Sind $f, g: U \rightarrow \mathbb{C}$ in z_0 komplex differenzierbar mit $g(z_0) \neq 0$, so gilt:

$$\left(\frac{f}{g}\right)'(z_0) = \left(f \cdot \frac{1}{g}\right)'(z_0) = f'(z_0) \cdot \frac{1}{g(z_0)} + f(z_0) \frac{-g'(z_0)}{g^2(z_0)} = \frac{f'(z_0)g(z_0) - f(z_0)g'(z_0)}{g^2(z_0)}.$$

- (b) Aus dem obigen Satz folgt unmittelbar, dass auch alle Polynome

$$p(z) = a_n z^n + \dots + a_1 z + a_0$$

mit $a_j \in \mathbb{C}$ holomorph sind. Außerdem stellen auch Potenzreihen holomorphe Funktionen dar.

Satz 17.3.4. Sei $\sum_{n=0}^{\infty} a_n (z - z_0)^n$ eine Potenzreihe mit $a_n \in \mathbb{C}$ und Konvergenzradius $R > 0$. Dann definiert

$$f(z) := \sum_{n=0}^{\infty} a_n (z - z_0)^n$$

eine auf $B(x_0, R)$ holomorphe Funktion. Ihre Ableitung ist durch

$$f'(z) = \sum_{n=1}^{\infty} n a_n (z - z_0)^{n-1}$$

gegeben. Außerdem hat $\sum_{n=1}^{\infty} n a_n (z - z_0)^{n-1}$ denselben Konvergenzradius wie $\sum_{n=0}^{\infty} a_n (z - z_0)^n$.

Beweis. Der Beweis ist vollkommen analog zu dem entsprechenden Satz 4.5.9 für reelle Potenzreihen. \square

Man definiert analog zum reellen Fall (Definition 4.5.10):

Definition 17.3.5. Eine auf einer offenen Menge $U \subset \mathbb{C}$ definierte Funktion f heißt *analytisch*, falls sich f in jedem Punkt in eine Potenzreihe entwickeln lässt, d.h. falls zu jedem $z_0 \in U$ ein $r > 0$ existiert mit $B(z_0, r) \subset U$ und

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n$$

für alle $z \in B(z_0, r)$.

Bemerkung. Insbesondere ist jede analytische Funktion holomorph. Überraschenderweise gilt auch die Umkehrung, wie wir bald beweisen werden.

Satz 17.3.6 (Kettenregel). Es seien $U, V \subset \mathbb{C}$ offen: Die Funktionen $f: U \rightarrow V$ bzw. $g: V \rightarrow \mathbb{C}$ seien in z_0 bzw. $f(z_0)$ komplex differenzierbar. Dann ist auch $g \circ f$ in z_0 komplex differenzierbar mit

$$(g \circ f)'(z_0) = g'(f(z_0)) \cdot f'(z_0).$$

Ist $\gamma: [a, b] \rightarrow U$ eine in $t_0 \in [a, b]$ differenzierbare Kurve mit $\gamma(t_0) = z_0$, so ist auch die Kurve $f \circ \gamma: [a, b] \rightarrow V$ differenzierbar in t_0 mit

$$(f \circ \gamma)'(t_0) = f'(\gamma(t_0))\gamma'(t_0)$$

Beweis. Der Beweis ist vollkommen analog zum Beweis von Satz 4.3.2. \square

Nun wollen wir das Kurvenintegral für komplexe Funktionen definieren.

Definition 17.3.7. Sei $A \subset \mathbb{C}$ und $f : A \rightarrow \mathbb{C}$ eine stetige Funktion. Ist $\gamma : [a, b] \rightarrow A$ eine stückweise stetig differenzierbare Kurve, so definiere

$$\int_{\gamma} f(z) dz := \int_a^b f(\gamma(t)) \gamma'(t) dt.$$

Bemerkung. Ist $U \subset \mathbb{C}$ offen und besitzt $f : U \rightarrow \mathbb{C}$ eine Stammfunktion F , so folgt aus der Kettenregel:

$$\int_{\gamma} f(z) dz = \int_a^b (F \circ \gamma)'(t) dt = F(\gamma(b)) - F(\gamma(a)).$$

Beispiel. Sei $z_0 \in \mathbb{C}$, $n \in \mathbb{Z} \setminus \{0\}$ und $\gamma : [0, 2\pi] \rightarrow \mathbb{C}$ der $|n|$ -fach durchlaufene Kreis mit $\gamma(t) = z_0 + re^{int}$. Ist $n > 0$ so wird γ *positiv orientiert* genannt, andernfalls *negativ orientiert*. Dann gilt:

$$\int_{\gamma} \frac{1}{z - z_0} dz = \int_0^{2\pi} \frac{1}{re^{int}} irne^{int} dt = n2\pi i.$$

Also misst

$$\frac{1}{2\pi i} \int_{\gamma} \frac{1}{z - z_0} dz,$$

wie oft sich γ um z_0 windet. Wie wir bald sehen werden, können wir allgemeinen geschlossenen Kurven auf diese Weise eine Windungszahl zuordnen. Hingegen ist

$$\int_{\gamma} (z - z_0)^n dz = 0$$

für alle $n \in \mathbb{Z} \setminus \{-1\}$ und alle geschlossenen Kurven $\gamma : [a, b] \rightarrow \mathbb{C} \setminus \{z_0\}$, denn der Integrand hat

$$\frac{1}{n+1} (z - z_0)^{n+1}$$

als Stammfunktion. Die auf $\mathbb{C} \setminus \{z_0\}$ holomorphe Funktion $\frac{1}{z - z_0}$ wird auch manchmal *Cauchy-kern* genannt. Wie wir bald sehen werden, spielt sie bei der Integraldarstellung von holomorphen Funktionen eine entscheidende Rolle.

Bemerkung. Aus der Definition des Kurvenintegrals ergibt sich die folgende Standardabschätzung:

$$\begin{aligned} \left| \int_{\gamma} f(z) dz \right| &\leq \int_a^b |f(\gamma(t))| \cdot |\gamma'(t)| dt \\ &\leq \sup\{|f(z)| \mid z \in \text{Bild } \gamma\} \cdot L(\gamma). \end{aligned}$$

Der Realteil und Imaginärteil eines komplexen Kurvenintegrals sind reelle Kurvenintegrale im Sinne des letzten Abschnittes. Genauer gilt:

Satz 17.3.8. Sei $A \subset \mathbb{C}$ und $f : A \rightarrow \mathbb{C}$ eine stetige Funktion mit

$$f(x + iy) = f_1(x, y) + if_2(x, y).$$

Ist $\gamma : [a, b] \rightarrow A$ eine stetig differenzierbare Kurve mit $\gamma(t) = \gamma_1(t) + i\gamma_2(t)$, so gilt:

$$\int_{\gamma} f(z)dz = \int_{\gamma} (f_1(x, y)dx - f_2(x, y)dy) + i \int_{\gamma} (f_2(x, y)dx + f_1(x, y)dy)$$

Beweis.

$$\begin{aligned} \int_{\gamma} f(z)dz &= \int_a^b f(\gamma(t))\gamma'(t)dt \\ &= \int_a^b (f_1(\gamma(t)) + if_2(\gamma(t)))(\gamma_1'(t) + i\gamma_2'(t))dt \\ &= \int_a^b f_1(\gamma(t))\gamma_1'(t) - f_2(\gamma(t))\gamma_2'(t)dt + i \int_a^b f_2(\gamma(t))\gamma_1'(t) + f_1(\gamma(t))\gamma_2'(t)dt \\ &= \int_{\gamma} (f_1dx - f_2dy) + i \int_{\gamma} (f_2dx + f_1dy) \end{aligned}$$

□

Daraus folgt sofort der Integralsatz von Cauchy:

Satz 17.3.9 (Cauchyscher Integralsatz). Sei $U \subset \mathbb{C}$ offen und $f : U \rightarrow \mathbb{C}$ eine holomorphe Funktion mit

$$f(x + iy) = f_1(x, y) + if_2(x, y).$$

Dann sind die 1-Formen $f_1dx - f_2dy$ und $f_2dx + f_1dy$ geschlossen. Insbesondere gilt:

(a) Sind γ_0 und γ_1 zwei homotope stückweise differenzierbare geschlossene Kurven so gilt:

$$\int_{\gamma_0} f(z)dz = \int_{\gamma_1} f(z)dz.$$

(b) Ist $U \subset \mathbb{C}$ offen und einfach zusammenhängend, so gilt für jede stückweise differenzierbare geschlossene Kurve γ :

$$\int_{\gamma} f(z)dz = 0$$

(c) Sei $U \subset \mathbb{C}$ offen und einfach zusammenhängend und $z, z_0 \in U$. Ist γ_z eine beliebige stückweise stetig differenzierbare Kurve in U , die z_0 mit z verbindet, so definiert

$$F(z) := \int_{\gamma_z} f(w)dw$$

eine holomorphe Funktion. Da U einfach zusammenhängend ist, ist $F(z)$ unabhängig von der Wahl der Kurve γ_z .

Bemerkung. Oft wird auch nur die Aussage in (b) als Cauchyscher Integralsatz bezeichnet.

Beweis. Die 1-Formen $f_2 dx + f_1 dy$ und $f_1 dx - f_2 dy$ sind genau dann geschlossen, falls die Gleichungen

$$\frac{\partial f_1}{\partial x} = \frac{\partial f_2}{\partial y} \quad \text{und} \quad \frac{\partial f_2}{\partial x} = -\frac{\partial f_1}{\partial y}$$

erfüllt sind. Dies sind aber genau die Cauchy-Riemannschen Differentialgleichungen. Damit ist die Aussage eine Konsequenz aus Satz 17.2.14. \square

Definition 17.3.10 (*Windungszahl, Umlaufzahl*). Sei $z_0 \in \mathbb{C}$ und $\gamma : [a, b] \rightarrow \mathbb{C} \setminus \{z_0\}$ eine geschlossene stückweise stetig differenzierbare Kurve. Dann heißt

$$w(\gamma, z_0) := \frac{1}{2\pi i} \int_{\gamma} \frac{1}{z - z_0} dz$$

die *Windungszahl* (*Umlaufzahl*) von γ bezüglich z_0 .

In Verallgemeinerung des obigen Beispiels zeigen wir nun, dass die Windungszahl eine ganze Zahl ist. Ihr Betrag misst, wie oft eine geschlossene Kurve $\gamma : [a, b] \rightarrow \mathbb{C} \setminus \{z_0\}$ den Punkt z_0 umläuft. Das Vorzeichen gibt die Orientierung der Umläufe an. Linksumläufe erhalten ein positives Vorzeichen und Rechtsumläufe ein negatives Vorzeichen.

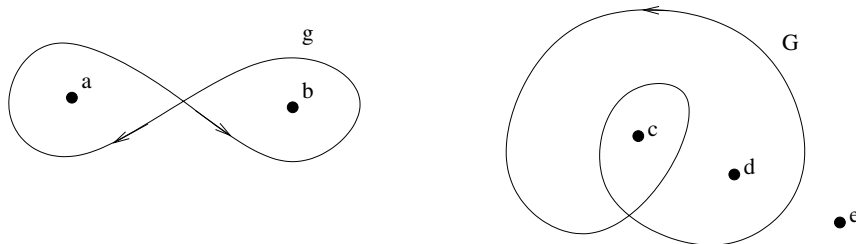


Abbildung 17.13: γ hat Windungszahl -1 bzgl. a und 1 bzgl. b . Γ hat Windungszahlen $2, 1, 0$ bzgl. c, d bzw. e .

Satz 17.3.11. Sei $z_0 \in \mathbb{C}$ und $\gamma : [a, b] \rightarrow \mathbb{C} \setminus \{z_0\}$ eine geschlossene stückweise stetig differenzierbare Kurve. Dann ist $w(\gamma, z_0)$ eine ganze Zahl. Darüberhinaus ist die Windungszahl homotopieinvariant, d.h. sind $\gamma_1 : [a, b] \rightarrow \mathbb{C} \setminus \{z_0\}$ und $\gamma_2 : [a, b] \rightarrow \mathbb{C} \setminus \{z_0\}$ homotope geschlossene stückweise stetig differenzierbare Kurven, so gilt:

$$w(\gamma_1, z_0) = w(\gamma_2, z_0).$$

Ist insbesondere $\gamma : [a, b] \rightarrow \mathbb{C} \setminus \{z_0\}$ nullhomotop in $\mathbb{C} \setminus \{z_0\}$, so gilt:

$$w(\gamma, z_0) = 0.$$

Beweis. Da der Cauchykernel $\frac{1}{z - z_0}$ auf $\mathbb{C} \setminus \{z_0\}$ holomorph ist, folgt die Homotopieinvarianz aus dem Cauchyschen Integralsatz. Es bleibt zu zeigen, dass die Windungszahl immer ganzzahlig ist. Sei also $\gamma : [a, b] \rightarrow \mathbb{C} \setminus \{z_0\}$ eine geschlossene stückweise stetig differenzierbare

Kurve. Der Einfachheit halber nehme man an, dass sie stetig differenzierbar ist. Da das Bild von γ den Punkt z_0 nicht trifft und γ geschlossen ist, besitzt γ die Darstellung (Polarkoordinaten um z_0)

$$\gamma(t) = r(t)e^{i\varphi(t)} + z_0$$

wobei $r : [a, b] \rightarrow \mathbb{R}_+$ eine positive differenzierbare Funktion mit $r(a) = r(b)$ ist und $\varphi : [a, b] \rightarrow \mathbb{R}$ eine differenzierbare Funktion ist mit

$$\varphi(b) - \varphi(a) = k2\pi$$

und $k \in \mathbb{Z}$. Damit erhalten wir

$$\begin{aligned} \frac{1}{2\pi i} \int_{\gamma} \frac{1}{z - z_0} &= \frac{1}{2\pi i} \int_a^b \frac{1}{r(t)e^{i\varphi(t)}} \left(r'(t)e^{i\varphi(t)} + r(t)i\varphi'(t)e^{i\varphi(t)} \right) dt \\ &= \frac{1}{2\pi i} \int_a^b \frac{r'(t)}{r(t)} + i\varphi'(t) dt \\ &= \frac{1}{2\pi i} \log r \Big|_a^b + \frac{1}{2\pi} (\varphi(b) - \varphi(a)) = k. \end{aligned}$$

17.4 Cauchysche Integralformel und Analytizität holomorpher Funktionen

Das folgende Lemma wird nützlich für den Beweis der Cauchyschen Integralformel sein.

Lemma 17.4.1. *Sei $U \subset \mathbb{C}$ offen und U enthalte die abgeschlossene Kreisscheibe $K(z_0, r)$ mit Radius $r > 0$ um z_0 . Sei $f : U \rightarrow \mathbb{C}$ eine stetige Funktion, die bis auf die Ausnahme eines Punktes w im Innern der Kreisscheibe $K(z_0, r)$ holomorph ist. Dann gilt:*

$$\int_{\gamma_r} f(z) dz = 0,$$

wobei $\gamma_r : [0, 2\pi] \rightarrow \mathbb{C}$ den parametrisierten Kreis mit $\gamma_r(t) = z_0 + re^{it}$ beschreibt.

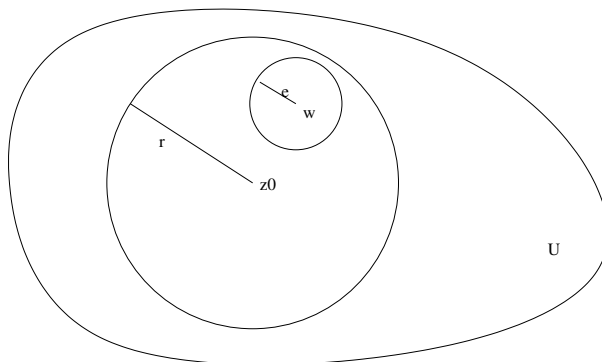


Abbildung 17.14: zu 17.4.1

Beweis. Sei also $f : U \rightarrow \mathbb{C}$ eine stetige Funktion, so dass f auf $U \setminus \{w\}$ mit $w \in B(z_0, r)$ holomorph ist. Für jedes $\epsilon > 0$ mit $K(w, \epsilon) \subset B(z_0, r)$, sind die Kreise $\gamma_r(t) = z_0 + re^{it}$ und $\alpha_\epsilon(t) = w + \epsilon e^{it}$ in $U \setminus \{w\}$ frei homotop. Damit folgt aus dem Cauchyschen Integralsatz und der obigen Abschätzung von Kurvenintegralen:

$$\left| \int_{\gamma_r} f(z) dz \right| = \left| \int_{\alpha_\epsilon} f(z) dz \right| \leq ML(\alpha_\epsilon) \leq M2\pi\epsilon,$$

wobei M das Maximum der stetigen Funktion f auf $K(z_0, r)$ bezeichnet. Da diese Abschätzung für alle $\epsilon > 0$ erfüllt ist, ist die Behauptung bewiesen. \square

Satz 17.4.2 (Cauchysche Integralformel). Sei $U \subset \mathbb{C}$ offen und $f : U \rightarrow \mathbb{C}$ eine holomorphe Funktion. Sei $z_0 \in U$ und $r > 0$ so gewählt, dass die abgeschlossene Kreisscheibe $K(z_0, r)$ in U enthalten ist. Sei $\gamma_r : [0, 2\pi] \rightarrow \mathbb{C}$ der parametrisierte Kreis mit $\gamma_r(t) = z_0 + re^{it}$. Dann gilt für jedes z in der offenen Kreisscheibe $B(z_0, r)$:

$$f(z) = \frac{1}{2\pi i} \int_{\gamma_r} \frac{f(w)}{w - z} dw$$

Beweis. Wegen der Holomorphie von f auf U ist auch der Differenzenquotient

$$q(w) = \frac{f(w) - f(z)}{w - z}$$

auf dem Definitionsbereich $U \setminus \{z\}$ holomorph und lässt sich nach z stetig fortsetzen. Damit folgt mit obigen Lemma:

$$0 = \int_{\gamma_r} \frac{f(w) - f(z)}{w - z} dw = \int_{\gamma_r} \frac{f(w)}{w - z} dw - f(z) \int_{\gamma_r} \frac{1}{w - z} dw.$$

Da wegen des obigen Beispiels

$$\int_{\gamma_r} \frac{1}{w - z} dw = 2\pi i$$

gilt, folgt die Behauptung. \square

Die Cauchysche Integralformel ist eine fundamentale Integraldarstellung holomorpher Funktionen. Sie besagt z.B. dass holomorphe Funktionen im Innern von Kreisscheiben durch die Werte auf dem Rand bestimmt sind. Die im Integranden der Integralformel vorkommende Funktion

$$CK(w, z) := \frac{1}{w - z}$$

wird manchmal auch *Cauchykernel* genannt. Basierend auf der geometrischen Reihe, werden wir nun den Cauchykernel $\frac{1}{w-z}$ für festes $w \in \mathbb{C}$ in eine Reihe um einen von w verschiedenen Punkt z_0 entwickeln. Damit werden wir dann die Potenzreihenentwicklung von f herleiten und somit die Analytizität holomorpher Funktionen beweisen.

Lemma 17.4.3. Seien $z_0, w \in \mathbb{C}$ mit $|w - z_0| = r$.

1. Ist $r > 0$, so gilt für alle $z \in B(z_0, r)$:

$$\frac{1}{w-z} = \frac{1}{w-z_0} \sum_{n=0}^{\infty} \left(\frac{z-z_0}{w-z_0} \right)^n = \sum_{n=0}^{\infty} \frac{1}{(w-z_0)^{n+1}} (z-z_0)^n.$$

2. Ist $r \geq 0$, so gilt für alle $z \in \mathbb{C} \setminus K(z_0, r)$:

$$\frac{1}{w-z} = \frac{-1}{z-z_0} \sum_{n=0}^{\infty} \left(\frac{w-z_0}{z-z_0} \right)^n = - \sum_{n=0}^{\infty} (w-z_0)^n (z-z_0)^{-(n+1)}.$$

Beide Reihen konvergieren absolut und gleichmäßig auf beschränkten Teilmengen.

Beweis. Sei $w \in \mathbb{C}$ mit $|w-z_0| = r$. Ist $r > 0$, so gilt für alle $z \in B(z_0, R)$:

$$\begin{aligned} \frac{1}{w-z} &= \frac{1}{w-z_0 - (z-z_0)} = \frac{1}{(w-z_0) \left(1 - \frac{z-z_0}{w-z_0} \right)} \\ &= \frac{1}{w-z_0} \sum_{n=0}^{\infty} \left(\frac{z-z_0}{w-z_0} \right)^n. \end{aligned}$$

Sei nun $r \geq 0$. Dann gilt für alle $z \in \mathbb{C} \setminus K(z_0, r)$:

$$\begin{aligned} \frac{1}{w-z} &= \frac{1}{w-z_0 - (z-z_0)} = \frac{1}{-(z-z_0) \left(1 - \frac{w-z_0}{z-z_0} \right)} \\ &= \frac{-1}{z-z_0} \sum_{n=0}^{\infty} \left(\frac{w-z_0}{z-z_0} \right)^n. \end{aligned}$$

□

Wir zeigen nun mit Hilfe der Cauchyschen Integralformel, dass sich jede holomorphe Funktion lokal in eine Potenzreihe entwickeln lässt. Genauer gilt:

Satz 17.4.4. Jede in einer offenen Menge $U \subset \mathbb{C}$ holomorphe Funktion f kann in einer Kreisscheibe $K(z_0, R) \subset U$ in eine Potenzreihe

$$f(z) = \sum_{n=0}^{\infty} a_n (z-z_0)^n$$

entwickelt werden. Dabei sind die Koeffizienten durch die Integrale

$$a_n := \frac{1}{2\pi i} \int_{\gamma_r} \frac{f(w)}{(w-z_0)^{n+1}} dw$$

gegeben, wobei $\gamma_r(t) = z_0 + re^{it}$ mit $t \in [0, 2\pi]$ einen parametrisierten Kreis um z_0 mit Radius $r < R$ bezeichnet.

Beweis. Aus der Cauchyschen Integralformel folgt für $|z - z_0| < r$:

$$f(z) = \frac{1}{2\pi i} \int_{\gamma_r} \frac{f(w)}{w - z} dw.$$

Wegen Teil 1 des obigen Lemmas 17.4.4 gilt für jedes $w \in \text{Bild}(\gamma_r)$ und alle $z \in B(z_0, r)$:

$$\frac{f(w)}{w - z} = \frac{f(w)}{w - z_0} \sum_{n=0}^{\infty} \left(\frac{z - z_0}{w - z_0} \right)^n$$

Da die Reihe gleichmäßig auf $\text{Bild} \gamma_r$ konvergiert, dürfen wir wegen des Satzes 6.2.2 die Summation und Integration vertauschen und es folgt:

$$f(z) = \frac{1}{2\pi i} \int_{\gamma_r} \frac{f(w)}{w - z} dw = \sum_{n=0}^{\infty} \left(\frac{1}{2\pi i} \int_{\gamma_r} \frac{f(w)}{(w - z_0)^{n+1}} dw \right) (z - z_0)^n$$

□

Wir haben somit Folgendes bewiesen:

Bemerkungen.

- (a) Die auf einer offenen Menge $U \subset \mathbb{C}$ holomorphen Abbildungen stimmen mit den auf U analytischen Funktionen überein. Insbesondere ist jede holomorphe Funktion unendlich oft komplex differenzierbar.
- (b) Die Nullstellen holomorpher Funktionen $f \not\equiv 0$ sind isoliert, d.h. zu jeder Nullstelle existiert eine Umgebung, in der sich keine weiteren Nullstellen befinden (siehe Satz 2.7.5).
- (c) (b) impliziert den folgenden Identitätssatz: Ist $U \subset \mathbb{C}$ offen und zusammenhängend und sind $f, g : U \rightarrow \mathbb{C}$ zwei holomorphe Funktionen, so gilt: Hat die Menge

$$\{z \in U \mid f(z) = g(z)\}$$

einen Häufungspunkt, so stimmen f und g auf ganz U überein.

Eine wichtige Anwendung des Satzes 17.4.4 ist der Satz von Liouville. Man nennt auf ganz \mathbb{C} holomorphe Funktionen *ganze Funktionen*.

Satz 17.4.5 (Satz von Liouville). *Jede beschränkte ganze Funktion ist konstant.*

Beweis. Ist $f : \mathbb{C} \rightarrow \mathbb{C}$ eine ganze Funktion, so besitzt sie wegen Satz 17.4.4 um $z_0 = 0$ die Potenzreihendarstellung

$$\sum_{n=0}^{\infty} a_n z^n$$

mit den Koeffizienten

$$a_n := \frac{1}{2\pi i} \int_{\gamma_r} \frac{f(w)}{w^{n+1}} dw,$$

wobei $r > 0$ beliebig ist und $\gamma_r : [0, 2\pi] \rightarrow \mathbb{C}$, $\gamma_r(t) = re^{it}$. Ist $|f(z)| \leq M$ für alle $z \in \mathbb{C}$ so folgt mit Hilfe der Standardabschätzung:

$$|a_n| \leq \frac{M}{r^n}$$

für alle $r > 0$. Damit folgt $a_n = 0$ für alle $n \geq 1$.

□

Die Analytizität holomorpher Funktionen hat die folgende überraschende Konsequenz.

Satz 17.4.6 (*Riemannscher Hebbarkeitssatz*). Sei U eine offene Menge und $z_0 \in U$. Ist $f : U \setminus \{z_0\} \rightarrow \mathbb{C}$ eine holomorphe Funktion, die in einer punktierten Umgebung $K(z_0, r) \setminus \{z_0\}$ beschränkt ist, so lässt sich f holomorph nach z_0 fortsetzen.

Beweis. Betrachte die Funktion $g : U \rightarrow \mathbb{C}$ mit

$$g(z) = \begin{cases} (z - z_0)^2 f(z) & \text{für } z \in U \setminus \{z_0\} \\ 0 & \text{für } z = z_0 \end{cases}$$

Diese ist natürlich ebenfalls auf $U \setminus \{z_0\}$ holomorph und wegen

$$\lim_{\substack{z \rightarrow z_0 \\ z \in U \setminus \{z_0\}}} \frac{g(z) - g(z_0)}{z - z_0} = \lim_{\substack{z \rightarrow z_0 \\ z \in U \setminus \{z_0\}}} (z - z_0) f(z) = 0$$

auch in z_0 komplex differenzierbar. Damit ist g auf U holomorph und besitzt eine Potenzreihenentwicklung um z_0 . Da $g(z_0) = g'(z_0) = 0$, ist sie von der Form

$$g(z) = \sum_{n=2}^{\infty} a_n (z - z_0)^n.$$

Damit definiert

$$\sum_{n=2}^{\infty} a_n (z - z_0)^{n-2}$$

die gesuchte Fortsetzung. □

17.5 Holomorphe Funktionen auf Kreisingen und ihre Laurententwicklung

Definition 17.5.1. Seien $R > r \geq 0$ und $z_0 \in \mathbb{C}$, so nennen wir

$$K(z_0, r, R) := K(z_0, R) \setminus B(z_0, r) = \{z \in \mathbb{C} \mid r \leq |z - z_0| \leq R\}$$

den *abgeschlossenen Kreisring* um z_0 mit innerem Radius r und äußerem Radius R . Entsprechend sei der *offene Kreisring* durch

$$B(z_0, r, R) := B(z_0, R) \setminus K(z_0, r) = \{z \in \mathbb{C} \mid r < |z - z_0| < R\}$$

bezeichnet.

Definition 17.5.2. Eine Reihe der Form

$$\sum_{n=-\infty}^{\infty} a_n (z - z_0)^n$$

mit $a_n \in \mathbb{C}$ heißt *Laurentreihe*. Die Reihen

$$\sum_{n=1}^{\infty} a_{-n} (z - z_0)^{-n} \quad \text{bzw.} \quad \sum_{n=0}^{\infty} a_n (z - z_0)^n$$

nennt man *Hauptteil* (*singulärer Teil*) bzw. *Nebenteil* (*regulärer Teil*) der Laurentreihe. Eine Laurentreihe heißt *konvergent* im Punkte z , falls ihre Hauptteile und Nebenteile im Punkte z konvergieren.

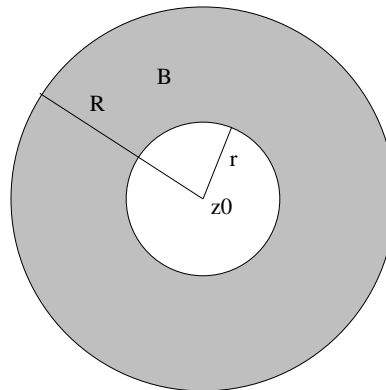


Abbildung 17.15: Kreisring

Lemma 17.5.3. *Sei*

$$\sum_{n=-\infty}^{\infty} a_n (z - z_0)^n$$

eine Laurentreihe. Ist $\frac{1}{r}$ der Konvergenzradius der Potenzreihe

$$\sum_{n=1}^{\infty} a_{-n} z^n,$$

so konvergiert der Hauptteil der Laurentreihe

$$\sum_{n=1}^{\infty} a_{-n} (z - z_0)^{-n}$$

auf

$$\mathbb{C} \setminus K(z_0, r) = \{z \in \mathbb{C} \mid r < |z - z_0|\}$$

und stellt dort eine holomorphe Funktion dar. Ist zusätzlich R der Konvergenzradius des Nebenteils und ist $r < R$, so konvergiert die Laurentreihe auf dem Kreisring

$$B(z_0, r, R) = \{z \in \mathbb{C} \mid r < |z - z_0| < R\}$$

und stellt dort eine holomorphe Funktion dar. Außerhalb des abgeschlossenen Kreisrings

$$K(z_0, r, R) = \{z \in \mathbb{C} \mid r \leq |z - z_0| \leq R\}$$

divergiert die Laurentreihe.

Bemerkung. Ist der Hauptteil einer Laurentreihe von null verschieden und besitzt er nur endlich viele Summanden, so konvergiert er offensichtlich auf $\mathbb{C} \setminus \{0\}$.

Beweis. Ist $\frac{1}{r}$ der Konvergenzradius der Reihe

$$p(z) = \sum_{n=1}^{\infty} a_{-n} z^n,$$

so ist der Hauptteil der Laurentreihe wegen

$$\sum_{n=1}^{\infty} a_{-n}(z - z_0)^{-n} = \sum_{n=1}^{\infty} a_{-n} \left(\frac{1}{z - z_0} \right)^n = p \left(\frac{1}{z - z_0} \right)$$

genau dann konvergent, falls

$$\left| \frac{1}{z - z_0} \right| < \frac{1}{r}$$

und somit

$$|z - z_0| > r$$

gilt. Da p auf $B(0, \frac{1}{r})$ holomorph ist, ist wegen der Kettenregel der Hauptteil $p(\frac{1}{z-z_0})$ auf der Menge

$$\{z \in \mathbb{C} \mid r < |z - z_0|\}$$

holomorph. □

Wir wollen nun die Umkehrung des obigen Lemmas zeigen, nämlich dass auf Kreisringen holomorphe Funktionen eine Darstellung durch eine Laurentreihe besitzen. Dazu werden wir zunächst eine Integralformel für holomorphe Funktionen auf Kreisringen herleiten.

Satz 17.5.4. *Es sei $U \subset \mathbb{C}$ eine offene Menge, die den abgeschlossenen Kreisring $K(z_0, r, R)$ für $0 \leq r < R$ enthält. Sei $f : U \rightarrow \mathbb{C}$ eine holomorphe Funktion und $\gamma_r : [0, 2\pi] \rightarrow \mathbb{C}$, $t \mapsto z_0 + re^{it}$, γ_R entsprechend. Dann gilt für alle $z \in B(z_0, r, R)$*

$$f(z) = \frac{1}{2\pi i} \int_{\gamma_R} \frac{f(w)}{w - z} dw - \frac{1}{2\pi i} \int_{\gamma_r} \frac{f(w)}{w - z} dw.$$

Dabei definiert

$$N_f(z) := \frac{1}{2\pi i} \int_{\gamma_R} \frac{f(w)}{w - z} dw$$

eine auf $B(z_0, R)$ holomorphe Funktion und

$$H_f(z) := -\frac{1}{2\pi i} \int_{\gamma_r} \frac{f(w)}{w - z} dw$$

eine auf $\mathbb{C} \setminus K(z_0, r)$ holomorphe Funktion. Wir nennen H_f den Hauptteil von f und N_f den Nebenteil von f .

Beweis. Wegen der Holomorphie von f auf U ist die Funktion $Q : U \rightarrow \mathbb{C}$ mit

$$Q(w) := \begin{cases} \frac{f(w) - f(z)}{w - z} & \text{für } w \in U \setminus \{z\} \\ f'(z) & \text{für } w = z \end{cases}$$

auf $U \setminus \{z\}$ holomorph und in z stetig. Wegen des Hebbarkeitssatzes ist dann Q aber auch auf ganz U holomorph. Da die Kreise γ_R und γ_r in U homotop sind, gilt gemäß 17.3.9(a):

$$\int_{\gamma_R} \frac{f(w) - f(z)}{w - z} dw = \int_{\gamma_r} \frac{f(w) - f(z)}{w - z} dw.$$

Mit $z \in B(z_0, R)$ folgt

$$\int_{\gamma_R} \frac{f(w) - f(z)}{w - z} dw = \int_{\gamma_R} \frac{f(w)}{w - z} dw - f(z) \int_{\gamma_R} \frac{1}{w - z} dw = \int_{\gamma_R} \frac{f(w)}{w - z} dw - 2\pi i f(z).$$

Wegen $z \in \mathbb{C} \setminus K(z_0, r)$ ist

$$\int_{\gamma_r} \frac{f(w) - f(z)}{w - z} dw = \int_{\gamma_r} \frac{f(w)}{w - z} dw - f(z) \int_{\gamma_r} \frac{1}{w - z} dw = \int_{\gamma_r} \frac{f(w)}{w - z} dw,$$

und die Behauptung ergibt sich durch Subtraktion beider Identitäten und nachfolgende Division durch $2\pi i$. \square

Mit Hilfe der Entwicklung des Cauchykerne in Lemma 17.4.3 erhalten wir nun für den Nebenteil und Hauptteil einer auf einem Kreisring definierten holomorphen Funktion folgende Reihendarstellungen:

Satz 17.5.5 (Laurententwicklung). *Es sei $U \subset \mathbb{C}$ eine offene Menge, die den abgeschlossenen Kreisring $K(z_0, r, R)$ für $0 \leq r < R$ enthält. Sei $f : U \rightarrow \mathbb{C}$ eine holomorphe Funktion. Dann gilt für alle $z \in B(z_0, R)$*

$$N_f(z) := \frac{1}{2\pi i} \int_{\gamma_R} \frac{f(w)}{w - z} dw = \sum_{n=0}^{\infty} \left(\frac{1}{2\pi i} \int_{\gamma_R} \frac{f(w)}{(w - z_0)^{n+1}} dw \right) (z - z_0)^n$$

und

$$H_f(z) := -\frac{1}{2\pi i} \int_{\gamma_r} \frac{f(w)}{w - z} dw = \sum_{n=0}^{\infty} \left(\frac{1}{2\pi i} \int_{\gamma_r} f(w)(w - z_0)^n dw \right) (z - z_0)^{-(n+1)}$$

für alle $z \in \mathbb{C} \setminus K(z_0, r)$. Auf dem Kreisring $B(z_0, r, R)$ konvergieren beide Reihen und wir erhalten

$$f(z) = N_f(z) + H_f(z) = \sum_{n=-\infty}^{\infty} a_n (z - z_0)^n$$

mit

$$a_n = \frac{1}{2\pi i} \int_{\gamma_\rho} \frac{f(w)}{(w - z_0)^{n+1}} dw,$$

wobei $\gamma_\rho(t) = z_0 + \rho e^{it}$ mit $t \in [0, 2\pi]$ einen parametrisierten Kreis um z_0 mit Radius $\rho \in (r, R)$ bezeichnet.

Beispiel. Betrachte die holomorphe Funktion $f : \mathbb{C} \setminus \{1, 2\} \rightarrow \mathbb{C}$ mit

$$f(z) = \frac{1}{(z-1)(z-2)} = \frac{1}{z-2} - \frac{1}{z-1}.$$

Wir wollen die Laurententwicklung auf dem Kreisring $B(0, 1, 2)$ bestimmen. Es gilt für den Nebenteil und Hauptteil

$$N_f(z) = -\frac{1}{z-2} = -\frac{1}{2(1-\frac{z}{2})} = \sum_{n=0}^{\infty} \frac{1}{2^{n+1}} z^n$$

und

$$H_f(z) = \frac{1}{(1-z)} = -\frac{1}{z(1-\frac{1}{z})} = -\sum_{n=0}^{\infty} \frac{1}{z^{n+1}}.$$

Nun wollen wir Singularitäten für holomorphe Funktionen definieren.

Definition 17.5.6. Sei $U \subset \mathbb{C}$ offen und $z_0 \in U$. Ist $f : U \setminus \{z_0\} \rightarrow \mathbb{C}$ eine holomorphe Funktion, so heißt z_0 *isolierte Singularität* von f . Ist

$$f(z) = \sum_{n=-\infty}^{\infty} a_n(z-z_0)^n$$

die Laurententwicklung von f mit Entwicklungspunkt z_0 , so heißt

$$\text{ord}_{z_0}(f) = \inf\{n \in \mathbb{Z} \mid a_n \neq 0\}$$

die *Ordnung* der Singularität. Ist $\text{ord}_{z_0}(f) = -\infty$ so heißt die Singularität *wesentlich*. Ist $\text{ord}_{z_0}(f) < 0$ bzw. $\text{ord}_{z_0}(f) > 0$, so heißt z_0 eine *Polstelle* bzw. *Nullstelle der Ordnung* $|\text{ord}_{z_0}(f)|$.

17.6 Der Residuensatz und seine Anwendungen

Definition 17.6.1. Sei $f : B(z_0, R) \setminus \{z_0\} \rightarrow \mathbb{C}$ eine holomorphe Funktion. Sei $\gamma_r(t) = z_0 + re^{it}$, $t \in [0, 2\pi]$ ein parametrisierter Kreis um z_0 mit Radius $r < R$. Dann heißt

$$\text{Res}_{z_0}(f) = \frac{1}{2\pi i} \int_{\gamma_r} f(z) dz$$

das *Residuum* von f in z_0 .

Bemerkungen.

- (a) Da die parametrisierten Kreise γ_r in $B(z_0, R) \setminus \{z_0\}$ paarweise homotop sind, ist die Definition unabhängig von der Wahl von γ_r .
- (b) Ist $f : B(z_0, R) \setminus \{z_0\} \rightarrow \mathbb{C}$ eine holomorphe Funktion, so besitzt f auf $B(z_0, R) \setminus \{z_0\}$ die Laurententwicklung

$$f(z) = \sum_{n=-\infty}^{\infty} a_n(z-z_0)^n.$$

Dann ist wegen Satz 17.5.5

$$a_{-1} = \frac{1}{2\pi i} \int_{\gamma_p} f(z) dz = \text{Res}_{z_0}(f)$$

Ist allgemeiner $\gamma : [a, b] \rightarrow B(z_0, R) \setminus \{z_0\}$ eine stückweise stetig-differenzierbare Kurve mit Windungszahl $w(\gamma, z_0)$, so gilt:

$$\frac{1}{2\pi i} \int_{\gamma} f(z) dz = \frac{1}{2\pi i} \sum_{n=-\infty}^{\infty} a_n \int_{\gamma} (z-z_0)^n dz = a_{-1} \frac{1}{2\pi i} \int_{\gamma} (z-z_0)^{-1} dz = \text{Res}_{z_0}(f) w(\gamma, z_0),$$

denn für $n \neq -1$ ist $\frac{1}{n+1}(z-z_0)^{n+1}$ Stammfunktion von $(z-z_0)^n$.

Satz 17.6.2 (Residuensatz). *Es sei $U \subset \mathbb{C}$ eine offene Menge und $S \subset U$ eine Menge ohne Häufungspunkte in \mathbb{C} . Sei $f : U \setminus S \rightarrow \mathbb{C}$ eine holomorphe Funktion und $\gamma : [a, b] \rightarrow U \setminus S$ eine geschlossene, in U nullhomotope Kurve. Dann ist die Menge*

$$\{z \in S \mid w(\gamma, z) \neq 0\}$$

eine endliche Menge $\{z_1, \dots, z_k\}$ und es gilt:

$$\int_{\gamma} f(z) dz = 2\pi i \sum_{j=1}^k \text{Res}_{z_j}(f) w(\gamma, z_j).$$

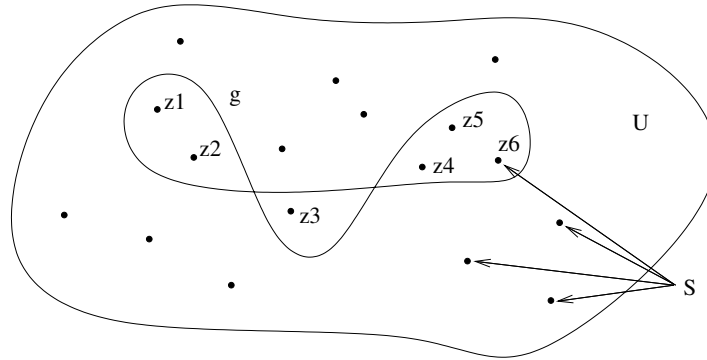


Abbildung 17.16: zum Residuensatz

Beweis. Da $\gamma : [a, b] \rightarrow U$ in U nullhomotop ist, existiert eine Homotopie $H : [0, 1] \times [a, b] \rightarrow U$ mit $H_0(t) = H(0, t) = \gamma(t)$ und $H_1(t) = H(1, t) = p \in U$. Dann ist $K = H([0, 1] \times [a, b])$ kompakt und der Schnitt von S mit K ist endlich, denn unendliche Mengen haben in kompakten Mengen einen Häufungspunkt. Ist $\tilde{S} = S \cap K = \{z_1, \dots, z_k\}$, so ist $\gamma : [a, b] \rightarrow U$ auch in $U \setminus (S \setminus \tilde{S})$ nullhomotop, denn die Homotopie "trifft" die Punkte aus $S \setminus \tilde{S}$ nicht. Seien $h_j : U \setminus \{z_j\} \rightarrow \mathbb{C}$ die Hauptteile von f in den Singularitäten $\{z_1, \dots, z_k\} \subset K$.

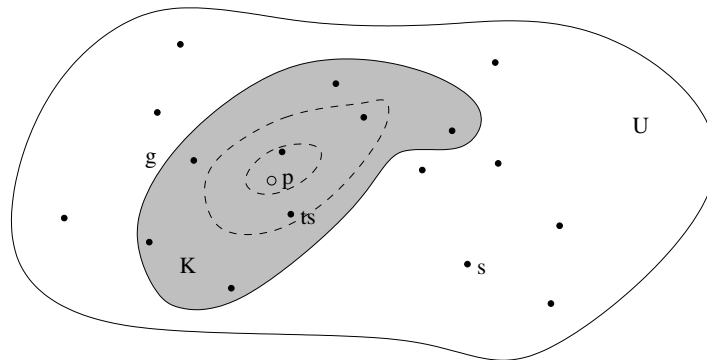


Abbildung 17.17: zum Beweis (der besseren Darstellbarkeit halber ist γ hier einfacher gewählt)

Dann ist

$$F := f - \sum_{j=1}^k h_j$$

eine auf $U \setminus (S \setminus \tilde{S})$ holomorphe Funktion und ihr Kurvenintegral über γ ist wegen des Cauchyschen Integralsatzes gleich null. Also folgt:

$$\int_{\gamma} f(z) dz = \sum_{j=1}^k \int_{\gamma} h_j(z) dz.$$

Da der Hauptteil h_j durch eine auf $\mathbb{C} \setminus \{z_j\}$ konvergierende Reihe der Form

$$\sum_{n=-\infty}^{-1} a_n (z - z_j)^n$$

gegeben ist, folgt aus obiger Bemerkung

$$\int_{\gamma} h_j(z) dz = 2\pi i a_{-1} w(\gamma, z_j)$$

und somit die Behauptung des Satzes. \square

Wir geben noch folgende praktische Methode zur Berechnung von Residuen an.

Satz 17.6.3. *Hat f in z_0 einen Pol k -ter Ordnung, so ist $g(z) = (z - z_0)^k f(z)$ in einer offenen Umgebung von z_0 holomorph und es gilt:*

$$\operatorname{Res}_{z_0}(f) = \frac{1}{(k-1)!} g^{(k-1)}(z_0).$$

Beweis. Da f in z_0 einen Pol k -ter Ordnung hat, besitzt f in z_0 eine Laurententwicklung der Form

$$f(z) = \frac{a_{-k}}{(z - z_0)^k} + \dots + \frac{a_{-1}}{(z - z_0)^{-1}} + \sum_{n=0}^{\infty} a_n (z - z_0)^n.$$

Dann besitzt $g(z) = (z - z_0)^k f(z)$ in z_0 die Potenzreihenentwicklung

$$g(z) = a_{-k} + \dots + a_{-1} (z - z_0)^{k-1} + \sum_{n=0}^{\infty} a_n (z - z_0)^{n+k}.$$

Insbesondere ist g in z_0 holomorph und es gilt

$$\operatorname{Res}_{z_0}(f) = a_{-1} = \frac{1}{(k-1)!} g^{(k-1)}(z_0).$$

\square

Beispiel. Betrachte $f(z) = \frac{e^{iz}}{(z-1)^2}$, so hat f in $z_0 = 1$ eine Polstelle zweiter Ordnung, denn e^{iz} ist holomorph. Da $g(z) = f(z)(z-1)^2 = e^{iz}$ folgt:

$$\operatorname{Res}_{z_0}(f) = g'(1) = ie^i.$$

Nun wollen wir den Residuensatz verwenden, um gewisse reelle Integrale zu berechnen. Zunächst wollen wir folgende uneigentliche Integrale untersuchen: Sei $f : \mathbb{R} \rightarrow \mathbb{C}$ eine Regelfunktion, d.h. Realteil und Imaginärteil von f seien Regelfunktionen. Nach Definition 5.4.1 existiert das uneigentliche Integral

$$\int_{-\infty}^{\infty} f(x) dx$$

genau dann, falls die Grenzwerte

$$\lim_{r \rightarrow \infty} \int_0^r f(x) dx =: \int_0^{\infty} f(x) dx \quad \text{und} \quad \lim_{r \rightarrow \infty} \int_{-r}^0 f(x) dx =: \int_{-\infty}^0 f(x) dx$$

existieren. Wir definieren dann:

$$\int_{-\infty}^{\infty} f(x) dx = \int_0^{\infty} f(x) dx + \int_{-\infty}^0 f(x) dx.$$

Außerdem folgt mit dem Majorantenkriterium (Satz 5.4.2): Ist $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ eine positive Regelfunktion mit $|f(x)| \leq \varphi(x)$ und existiert das Integral $\int_{-\infty}^{\infty} \varphi(x) dx$, so auch das Integral $\int_{-\infty}^{\infty} f(x) dx$. Ist insbesondere f absolut integrierbar, d.h. existiert

$$\int_{-\infty}^{\infty} |f(x)| dx,$$

so auch das Integral von $f : \mathbb{R} \rightarrow \mathbb{C}$.

Bemerkung. Der Grenzwert $\lim_{r \rightarrow \infty} \int_{-r}^r f(x) dx$ kann existieren, ohne dass das Integral $\int_{-\infty}^{\infty} f(x) dx$

existiert. Ein Beispiel hierfür ist die Funktion $f(x) = x$. Den Grenzwert $\lim_{r \rightarrow \infty} \int_{-r}^r f(x) dx$ nennt

man auch *Cauchyschen Hauptwert* des Integrals $\int_{-\infty}^{\infty} f(x) dx$. Existiert das Integral, so stimmt es offensichtlich mit seinem Cauchyschen Hauptwert überein. Wir wollen nun mittels des Residuenkalküls Fourierintegrale berechnen.

Definition 17.6.4 (Fourierintegral). Sei $f : \mathbb{R} \rightarrow \mathbb{C}$ eine Regelfunktion. Existiert für jedes $p \in \mathbb{R}$ das *Fourierintegral*

$$\hat{f}(p) := \int_{-\infty}^{\infty} f(x) e^{-ipx} dx,$$

so heißt $\hat{f} : \mathbb{R} \rightarrow \mathbb{C}$ die *Fouriertransformierte* von f .

Bemerkung. Ist f absolut integrierbar, so ist wegen $|f(x)e^{-ipx}| = |f(x)|$ auch $f(x)e^{-ipx}$ absolut integrierbar und die Fouriertransformierte existiert.

Satz 17.6.5. Sei $S \subset \mathbb{C} \setminus \mathbb{R}$ eine endliche Menge und $f : \mathbb{C} \setminus S \rightarrow \mathbb{C}$ eine holomorphe Funktion mit

$$\lim_{r \rightarrow \infty} \max_{|z|=r} |f(z)| \cdot r = 0.$$

Betrachte für $p \in \mathbb{R}$ die holomorphe Funktion $g_p : \mathbb{C} \setminus S \rightarrow \mathbb{C}$ mit $g_p(z) := f(z)e^{-ipz}$. Dann gilt:

$$\lim_{r \rightarrow \infty} \int_{-r}^r f(x)e^{-ipx} dx = \begin{cases} -2\pi i \sum_{\substack{z \in S \\ \text{im} z < 0}} \text{Res}_z(g_p), & p \geq 0 \\ 2\pi i \sum_{\substack{z \in S \\ \text{im} z > 0}} \text{Res}_z(g_p), & p \leq 0 \end{cases}$$

Beweis. Seien $p \leq 0$ und $r > 0$, so dass $S \subset \{z \in \mathbb{C} \mid |z| < r\}$ sowie

$$S^+ = \{z \in S \mid \text{im} z > 0\}.$$

Sei γ der Integrationsweg, der sich aus dem Intervall $[-r, r]$ und dem Halbkreis $\gamma_r : [0, \pi] \rightarrow \mathbb{C}$ mit $\gamma_r(t) = re^{it}$ zusammensetzt. Dann folgt aus dem Residuensatz:

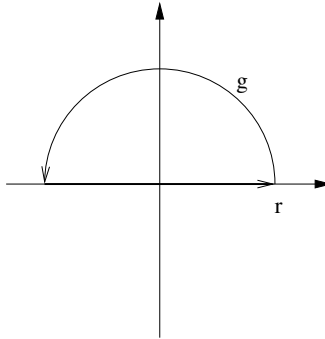


Abbildung 17.18: Integrationsweg für $p \leq 0$

$$\int_{\gamma} f(z) dz = \int_{-r}^r f(x)e^{-ipx} dx + \int_0^{\pi} f(re^{it})e^{-ipre^{it}} ire^{it} dt = 2\pi i \sum_{z \in S^+} \text{Res}_z(g_p).$$

Für $0 \leq t \leq \pi$ gilt wegen $p \leq 0$:

$$\left| e^{-ipre^{it}} \right| = \left| e^{-ipr \cos t} \right| \cdot \left| e^{-ipir \sin t} \right| = e^{pr \sin t} \leq 1.$$

Damit erhalten wir:

$$\left| \int_0^{\pi} f(re^{it})e^{-ipre^{it}} re^{it} dt \right| \leq \pi \max_{|z|=r} |f(z)| \cdot r.$$

Die Annahme an f impliziert somit die Behauptung. Den Fall $p \geq 0$ behandelt man analog mit Integrationsweg $t \mapsto re^{-it}$. □

Beispiele.

(a) Betrachte für $a > 0$ die Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ mit $f(x) = \frac{1}{x^2+a^2}$. Das Integral $\int_{-\infty}^{+\infty} |f(x)| dx$ existiert.

Außerdem ist $f : \mathbb{C} \setminus \{\pm ia\} \rightarrow \mathbb{C}$ mit $f(z) = \frac{1}{z^2+a^2}$ eine holomorphe Funktion. Mit

$$g_p(z) = f(z)e^{-ipz} = \frac{e^{-ipz}}{(z+ia)(z-ia)}$$

folgt

$$\operatorname{Res}_{ia}(g_p) = \frac{e^{pa}}{2ia}$$

und

$$\operatorname{Res}_{-ia}(g_p) = \frac{e^{-pa}}{-2ia}.$$

Also folgt für $p \geq 0$

$$\hat{f}(p) = \int_{-\infty}^{+\infty} \frac{e^{-ipx}}{x^2+a^2} dx = -2\pi i \frac{e^{-pa}}{-2ia} = \frac{\pi e^{-pa}}{a}$$

und für $p \leq 0$

$$\hat{f}(p) = 2\pi i \frac{e^{pa}}{2ia} = \frac{\pi e^{pa}}{a}.$$

(b) Als weitere Anwendung zeigen wir:

$$\int_{-\infty}^{+\infty} \frac{dx}{x^4+1} = \frac{\pi}{\sqrt{2}}.$$

Betrachte

$$(z^4+1) = (z^2-i)(z^2+i) = (z-e^{\frac{\pi i}{4}})(z+e^{\frac{\pi i}{4}})(z-ie^{\frac{\pi i}{4}})(z+ie^{\frac{\pi i}{4}}).$$

Setzen wir $z_0 = e^{\frac{\pi i}{4}}$ und $S = \{z_0, -z_0, iz_0, -iz_0\}$, so ist die Funktion $f : \mathbb{C} \setminus S \rightarrow \mathbb{C}$ mit

$$f(z) = \frac{1}{z^4+1} = \frac{1}{(z-z_0)(z+z_0)(z-iz_0)(z+iz_0)}$$

holomorph. Insbesondere gilt:

$$S^+ = \{z_0, iz_0\} = \{z \in S \mid \operatorname{im}(z) > 0\}.$$

Also folgt aus obigem Satz:

$$\int_{-\infty}^{+\infty} \frac{dx}{x^4+1} = 2\pi i (\operatorname{Res}_{z_0}(f) + \operatorname{Res}_{iz_0}(f)).$$

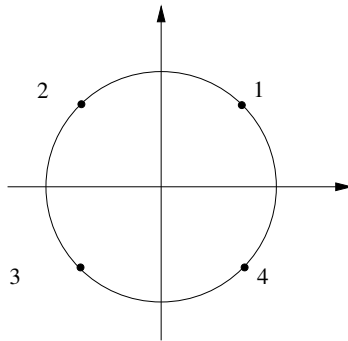


Abbildung 17.19: zu Bsp. (b)

Wir erhalten für die Residuen:

$$\operatorname{Res}_{z_0}(f) = \frac{1}{2z_0(z_0 - iz_0)(z_0 + iz_0)} = \frac{1}{2z_0^3(1 - i)(1 + i)} = \frac{1}{4z_0^3}$$

und

$$\begin{aligned} \operatorname{Res}_{iz_0} &= \frac{1}{(iz_0 - z_0)(iz_0 + z_0)(2iz_0)} = \frac{1}{2iz_0^3(i - 1)(i + 1)} \\ &= \frac{1}{-4iz_0^3} = \frac{i}{4z_0^3}. \end{aligned}$$

Da

$$z_0^3 = e^{\frac{3}{4}\pi i} = i \cdot z_0 = i \frac{(1 + i)}{\sqrt{2}},$$

folgt

$$\int_{-\infty}^{+\infty} \frac{dx}{x^4 + 1} = 2\pi i \frac{1}{4z_0^3}(1 + i) = \frac{2\pi i \sqrt{2}}{4i} = \frac{\pi}{\sqrt{2}}.$$

Kapitel 18

Maß- und Integrationstheorie

Literatur zur Maß- und Integrationstheorie:

- (a) Evans/Gariepy: Measure theory and fine properties of functions. Studies in Advanced Mathematics.
Schneller Einstieg auf den ersten 26 Seiten.
- (b) Amann/Escher: Analysis III. Grundstudium Mathematik, Birkhäuser.
Ausführliche Einführung auf 165 Seiten.
- (c) Heinz Bauer: Maß- und Integrationstheorie, de Gruyter Lehrbuch.
Ausführliche, aber dennoch übersichtliche Darstellung der Maß- und Integrationstheorie (Standardwerk in der deutschen Lehrbuchliteratur).
- (d) Elstrodt: Maß- und Integrationstheorie: Grundwissen Mathematik, Springer.
Ausführliche Darstellung der Maß- und Integrationstheorie, Speziallektüre.

18.1 Maße und messbare Mengen

In der Maßtheorie möchte man Teilmengen eines Raumes X ein Volumen zuordnen. Wir sind hauptsächlich am Fall $X = \mathbb{R}^n$ interessiert, aber für viele Anwendungen ist es notwendig, allgemeinere Räume zu betrachten.

Definition 18.1.1. Sei X eine Menge und $P(X)$ die Potenzmenge von X . Eine Abbildung

$$\mu : P(X) \rightarrow [0, \infty] := \{x \in \mathbb{R} \mid x \geq 0\} \cup \{\infty\}$$

heißt ein *äußeres Maß* auf X , falls Folgendes gilt:

- (i) $\mu(\emptyset) = 0$,
- (ii) $\mu(A) \leq \sum_{k=1}^{\infty} \mu(A_k)$, falls $A \subset \bigcup_{k=1}^{\infty} A_k$ (*Subadditivität*).

Bemerkungen.

- (a) Aus (i) und (ii) folgt die *Monotonie* von μ , d.h. $\mu(A) \leq \mu(B)$, falls $A \subset B$:
Setze $B = \bigcup_{i=1}^{\infty} A_i$ mit $A_1 = B, A_i = \emptyset$ für $i > 1$.

- (b) Die Reihe $\sum_{k=1}^{\infty} \mu(A_k)$ besteht aus Gliedern $\mu(A_k) \geq 0$. Sie konvergiert also genau dann, falls die Partialsummenfolge $\sum_{k=1}^n \mu(A_k) = s_n$ beschränkt ist. Divergiert die Reihe, so setze

$$\sum_{k=1}^{\infty} \mu(A_k) = \infty.$$

Nach dem Umordnungssatz (2.6.2) ist $\sum_{k=1}^{\infty} \mu(A_k)$ unabhängig von der Summationsreihenfolge.

- (c) μ heißt *Wahrscheinlichkeitsmaß*, falls $\mu(X) = 1$.

Beispiele.

- (a) Sei X eine Menge, $x_0 \in X$. Definiere

$$\delta_{x_0} = \begin{cases} 1 & \text{falls } x_0 \in A, \\ 0 & \text{sonst.} \end{cases}$$

δ_{x_0} ist ein äußeres Maß auf X , denn $\delta_{x_0}(\emptyset) = 0$ und sind $A \subset \bigcup_{k=1}^{\infty} A_k$ und $x_0 \in A$, so ist $x_0 \in A_{k_0}$ für ein $k_0 \in \mathbb{N}$. Daraus folgt

$$1 = \delta_{x_0}(A) = \delta_{x_0}(A_{k_0}) \leq \sum_{k=1}^{\infty} \delta_{x_0}(A_k).$$

Dieses äußere Maß heißt *Diracmaß*.

- (b) Sei X eine Menge, so setze

$$\mu(A) = \begin{cases} \text{card}A & \text{falls } A \text{ endlich,} \\ \infty & \text{sonst,} \end{cases}$$

wobei card die Kardinalität bezeichnet. Dieses äußere Maß heißt auch *Zählmaß*, da es die Anzahl der Elemente einer Menge zählt. Ist X endlich, so definiert

$$\mu(A) = \frac{\text{card}A}{\text{card}X}$$

ein Wahrscheinlichkeitsmaß auf X . Dieses Maß findet insbesondere in der elementaren Statistik eine Anwendung.

- (c) Ist μ ein äußeres Maß auf X und $A \subset X$, so heißt $\mu|_A$ mit

$$\mu|_A(B) := \mu(A \cap B)$$

für $B \in P(X)$ das äußere Maß μ *eingeschränkt auf* A .

Das wichtigste Beispiel ist das äußere *Lebesgue-Maß* auf \mathbb{R}^n . Ist I ein beschränktes Intervall, so heißt

$$|I| = b - a$$

die *Länge* von I , falls I durch (a, b) , $(a, b]$, $[a, b)$ oder $[a, b]$ mit $a < b \in \mathbb{R}$ gegeben ist. Eine Menge $Q \subset \mathbb{R}^n$ der Form

$$Q = I_1 \times \dots \times I_n$$

heißt *Quader*. Wir definieren das Volumen des Quaders durch

$$\text{vol}Q := \text{vol}_n Q = |I_1| \cdots |I_n|,$$

d.h. das Volumen eines Quaders ist definiert als das Produkt seiner Kantenlängen.

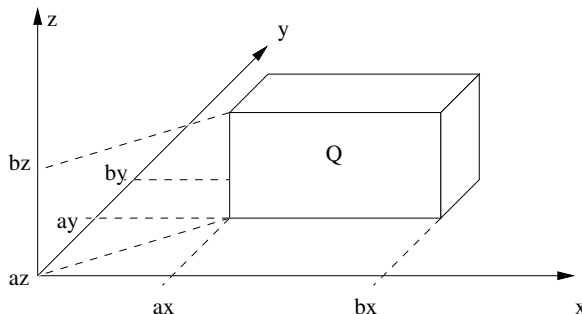


Abbildung 18.1: Quader mit Volumen $\text{vol}_3 Q = (b_x - a_x)(b_y - a_y)(b_z - a_z)$

Definition 18.1.2 (*Äußeres Lebesguemaß*). Die Abbildung $\lambda := \lambda^n : P(\mathbb{R}^n) \rightarrow [0, \infty]$ mit

$$\lambda(A) := \inf \left\{ \sum_{j=1}^{\infty} \text{vol}Q_j \mid A \subset \bigcup_{j=1}^{\infty} Q_j \right\}$$

heißt *äußeres Lebesgue-Maß* auf \mathbb{R}^n .

Satz 18.1.3. λ ist ein äußeres Maß.

Beweis. $\lambda(\emptyset) = 0$, denn die leere Menge ist in jeder Menge enthalten und somit auch in jedem Quader mit beliebig kleinen Kantenlängen.

Sei $A \subset \mathbb{R}^n$ und $A \subset \bigcup_{k=1}^{\infty} A_k$. Wir können annehmen, dass $\sum_{k=1}^{\infty} \lambda(A_k)$ konvergiert, denn sonst ist die Eigenschaft (ii) in der Definition des äußeren Maßes trivialerweise erfüllt. Nach Definition von $\lambda(A_k)$ ist zu gegebenen $\epsilon > 0$ die Zahl $\lambda(A_k) + \frac{\epsilon}{2^k}$ keine untere Schranke von

$$\left\{ \sum_{j=1}^{\infty} \text{vol}Q_j \mid A_k \subset \bigcup_{j=1}^{\infty} Q_j \right\} \subset \mathbb{R}.$$

Daher existieren Quader Q_{k_j} mit $A_k \subset \bigcup_{j=1}^{\infty} Q_{k_j}$ und $\sum_{j=1}^{\infty} \text{vol} Q_{k_j} < \lambda(A_k) + \epsilon/2^k$. Dann gilt

$$A \subset \bigcup_{k=1}^{\infty} A_k \subset \bigcup_{k=1}^{\infty} \bigcup_{j=1}^{\infty} Q_{k_j}$$

und somit mittels geometrischer Reihe

$$\begin{aligned} \lambda(A) &\leq \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} \text{vol} Q_{k_j} \leq \sum_{k=1}^{\infty} \lambda(A_k) + \frac{\epsilon}{2^k} = \sum_{k=1}^{\infty} \lambda(A_k) + \epsilon \left(\frac{1}{1 - \frac{1}{2}} - 1 \right) \\ &= \sum_{k=1}^{\infty} \lambda(A_k) + \epsilon. \end{aligned}$$

Da $\epsilon > 0$ beliebig ist, folgt:

$$\lambda(A) \leq \sum_{k=1}^{\infty} \lambda(A_k).$$

□

Bemerkungen.

(a) Da die Summanden alle positiv sind, folgt:

$$\sum_{k=1}^{\infty} \sum_{j=1}^{\infty} \text{vol} Q_{k_j} = \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} \text{vol} Q_{k_j}.$$

(b) Ist Q ein Quader im \mathbb{R}^n , so ist

$$\lambda(Q) = \text{vol} Q,$$

denn

$$\lambda(Q) = \inf \left\{ \sum_{j=1}^{\infty} \text{vol} Q_j \mid Q \subset \bigcup_{j=1}^{\infty} Q_j \right\} \leq \text{vol} Q.$$

Außerdem kann man zeigen:

$$\text{vol} Q \leq \sum_{j=1}^{\infty} \text{vol} Q_j \tag{18.1}$$

für eine beliebige Überdeckung von Q mit Quadern Q_j , $j \in \mathbb{N}$. Daher ist

$$\text{vol} Q \leq \lambda(Q).$$

Beweis von (18.1):

1. Es gilt

$$\text{vol} Q \leq \sum_{j=1}^N \text{vol} Q_j$$

für jede endliche Überdeckung von Q mit Quadern Q_1, \dots, Q_N .

2. Ist Q kompakt und wird von $\{Q_j\}$ mit $j \in \mathbb{N}$ überdeckt, so ersetze Q_j durch offene Quader Q'_j mit

$$Q_j \subset Q'_j \text{ und } \text{vol}Q'_j \leq \text{vol}Q_j + \frac{\epsilon}{2^j}.$$

Damit ist aber $\bigcup_{j=1}^{\infty} Q'_j$ eine offene Überdeckung von Q und wegen der Kompaktheit wird Q bereits von endlich vielen Q'_1, \dots, Q'_N , überdeckt. Dann folgt aus 1.:

$$\text{vol}Q \leq \sum_{j=1}^N \text{vol}Q'_j \leq \sum_{j=1}^{\infty} \text{vol}Q'_j \leq \sum_{j=1}^{\infty} \text{vol}Q_j + \frac{\epsilon}{2^j} = \sum_{j=1}^{\infty} \text{vol}Q_j + \epsilon.$$

Da $\epsilon > 0$ beliebig ist, folgt (18.1) für kompakte Q . Ist Q nicht kompakt, so existiert für jedes $\epsilon > 0$ ein kompakter Quader $Q' \subset Q$ mit $\text{vol}Q \leq \text{vol}Q' + \epsilon$. Ist $Q \subset \bigcup_{k=1}^{\infty} Q_k$, so ist

$$Q' \subset \bigcup_{k=1}^{\infty} Q_k \text{ und}$$

$$\text{vol}Q \leq \text{vol}Q' + \epsilon \leq \sum_{k=1}^{\infty} \text{vol}Q_k + \epsilon$$

für jedes $\epsilon > 0$. Daraus folgt die Behauptung auch für beliebige Quader.

Äußere Maße haben folgenden Schönheitsfehler: Sie sind nicht unbedingt additiv, d.h. ist μ ein äußeres Maß und sind A_1, \dots, A_n paarweise disjunkte Mengen, so folgt nicht notwendigerweise:

$$\sum_{i=1}^n \mu(A_i) = \mu\left(\bigcup_{i=1}^n A_i\right) \text{ (Additivität).}$$

Als einfaches Beispiel betrachte das äußere Maß $\mu : P(\mathbb{R}^n) \rightarrow [0, \infty]$ mit

$$\mu(A) = \begin{cases} 1 & \text{für } A \neq \emptyset, \\ 0 & \text{sonst.} \end{cases}$$

Sind $A, B \subset \mathbb{R}^n$ nicht leere disjunkte Mengen, so ist wegen

$$1 = \mu(A \cup B) < \mu(A) + \mu(B) = 2$$

die Additivität nicht erfüllt. Jedoch ist, wie wir sehen werden, diese Additivität für die Teilklasse der messbaren Mengen erfüllt.

Definition 18.1.4. Sei μ ein äußeres Maß auf der Menge X . Eine Teilmenge $A \subset X$ heißt *messbar*, falls

$$\mu(E) = \mu(E \cap A) + \mu(E \setminus A)$$

für alle Mengen $E \subset X$ gilt.

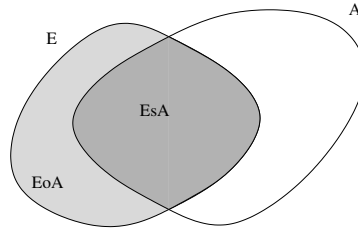


Abbildung 18.2: zu Def. 18.1.4

Bemerkungen.

(a) Da $E = (E \cap A) \cup (E \setminus A)$, folgt aus der Subadditivität von μ

$$\mu(E) \leq \mu(E \cap A) + \mu(E \setminus A)$$

für alle $A, E \subset X$.

(b) Ist $\mu(A) = 0$, so ist

$$\mu(E \cap A) + \mu(E \setminus A) \leq \mu(A) + \mu(E) = \mu(E),$$

d.h. A ist μ -messbar.

(c) Sei $A^c := X \setminus A$ das *Komplement* von A , so gilt:

$$E \cap A^c = E \cap (X \setminus A) = (E \cap X) \setminus A = E \setminus A$$

und damit ist A messbar genau dann, falls

$$\mu(E) = \mu(E \cap A) + \mu(E \cap A^c)$$

für alle $E \subset X$. Wegen $(A^c)^c = A$ folgt somit: A messbar $\Leftrightarrow A^c$ messbar.

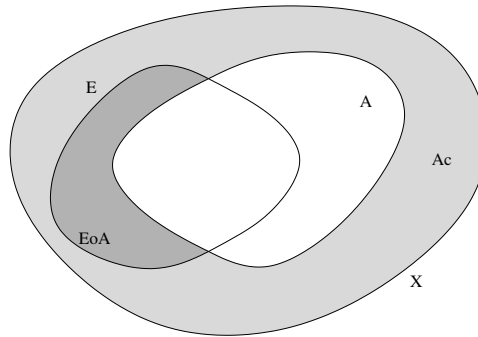


Abbildung 18.3: zu Bem. (c)

Wir werden sehen, dass die μ - messbaren Mengen eine σ -Algebra bilden.

Definition 18.1.5. Sei $\mathcal{A} \subset P(X)$ ein nicht leeres System von Teilmengen von X . Dann heißt \mathcal{A} eine *Algebra*, falls Folgendes gilt:

- (i) $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$,
- (ii) $A, B \in \mathcal{A} \Rightarrow A \cup B \in \mathcal{A}$.

Folgt außerdem aus $A_j \in \mathcal{A}$, $j \in \mathbb{N}$ stets auch $\bigcup_{j=1}^{\infty} A_j \in \mathcal{A}$, so heißt \mathcal{A} eine σ -Algebra. Das Paar (X, \mathcal{A}) heißt dann *Messraum* oder *messbarer Raum* (englische Bezeichnung: measurable space). Die Mengen in \mathcal{A} heißen \mathcal{A} -messbare Mengen.

Bemerkung. Ist \mathcal{A} eine Algebra, so ist \mathcal{A} abgeschlossen unter endlichen Mengenoperationen. Denn sind $A, B \in \mathcal{A}$, so sind wegen $A \cap B = (A^c \cup B^c)^c$ und $A \setminus B = A \cap B^c$ auch $A \cap B \in \mathcal{A}$ und $A \setminus B \in \mathcal{A}$. Ist darüberhinaus \mathcal{A} eine σ -Algebra, so gilt auch $\bigcap_{j=1}^{\infty} A_j \in \mathcal{A}$, falls

$$A_j \in \mathcal{A}, \text{ denn } \bigcap_{j=1}^{\infty} A_j = \left(\bigcup_{j=1}^{\infty} A_j^c \right)^c.$$

Wegen $A \cap A^c = \emptyset$ enthält jede Algebra und somit auch jede σ -Algebra die leere Menge. Wegen $\emptyset^c = X$ enthalten sie auch X . Insbesondere ist das Mengensystem (X, \emptyset) die kleinste σ -Algebra in X .

Lemma 18.1.6. Ist $\mu : P(X) \rightarrow [0, \infty]$ ein äußeres Maß, so bilden die messbaren Mengen eine Algebra \mathcal{A} . Sind $A, B \in \mathcal{A}$, so ist

$$\mu(E \cap (A \cup B)) = \mu(E \cap A) + \mu((E \setminus A) \cap B).$$

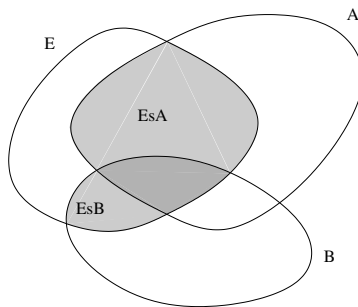


Abbildung 18.4: zu Lemma 18.1.6

Beweis. Wegen Bemerkung (c) nach Def. 18.1.4 genügt es zu zeigen: $A, B \in \mathcal{A} \Rightarrow A \cup B \in \mathcal{A}$. Seien $A, B \in \mathcal{A}$ und $E \subset X$. Dann folgt aus der Messbarkeit von A und B :

$$\mu(E) \stackrel{A \text{ messb.}}{=} \mu(E \cap A) + \mu(E \setminus A) \stackrel{B \text{ messb.}}{=} \mu(E \cap A) + \mu((E \setminus A) \cap B) + \mu((E \setminus A) \setminus B).$$

Da

$$E \cap (A \cup B) = (E \cap A) \cup (E \cap B) = (E \cap A) \cup ((E \setminus A) \cap B),$$

folgt aus der Subadditivität von μ :

$$\mu(E \cap A) + \mu((E \setminus A) \cap B) \geq \mu(E \cap (A \cup B)) \quad (18.2)$$

und somit aus

$$\mu(E) \geq \mu(E \cap (A \cup B)) + \mu((E \setminus A) \setminus B) = \mu(E \cap (A \cup B)) + \mu(E \setminus (A \cup B)) \geq \mu(E)$$

die Messbarkeit von $A \cup B$. Insbesondere muss in (18.2) Gleichheit gelten, d.h.

$$\mu(E \cap (A \cup B)) = \mu(E \cap A) + \mu((E \setminus A) \cap B).$$

□

Nun wollen wir zeigen, dass \mathcal{A} eine σ -Algebra darstellt. Dazu benötigen wir noch folgendes Lemma.

Lemma 18.1.7. *Es seien $\{A_j\}_{j \in \mathbb{N}}$ eine Familie disjunkter, messbarer Mengen und $S = \bigcup_{j=1}^{\infty} A_j$. Dann gilt für alle $E \subset X$:*

$$\mu(E) = \sum_{j=1}^{\infty} \mu(E \cap A_j) + \mu(E \setminus S) = \mu(E \cap S) + \mu(E \setminus S).$$

Insbesondere ist S messbar.

Beweis. Sind $A, B \in \mathcal{A}$ und $A \cap B = \emptyset$, so ist $(E \setminus A) \cap B = E \cap B$ und wegen Lemma 18.1.6 gilt:

$$\mu(E \cap (A \cup B)) = \mu(E \cap A) + \mu(E \cap B).$$

Ist $\{A_j\}_{j \in \mathbb{N}}$ eine disjunkte Familie messbarer Mengen, so folgt durch Induktion

$$\mu(E \cap \bigcup_{j=1}^k A_j) = \sum_{j=1}^k \mu(E \cap A_j).$$

Da $\bigcup_{j=1}^k A_j$ messbar ist, erhalten wir somit für alle $E \subset X$

$$\mu(E) = \mu(E \cap \bigcup_{j=1}^k A_j) + \mu(E \setminus \bigcup_{j=1}^k A_j) \geq \sum_{j=1}^k \mu(E \cap A_j) + \mu(E \setminus S).$$

Dies impliziert:

$$\begin{aligned} \mu(E) &\geq \sum_{j=1}^{\infty} \mu(E \cap A_j) + \mu(E \setminus S) \stackrel{\text{Monotonie}}{\geq} \mu(E \cap \bigcup_{j=1}^{\infty} A_j) + \mu(E \setminus S) \\ &= \mu(E \cap S) + \mu(E \setminus S) \geq \mu(E). \end{aligned}$$

□

Bemerkung. Ist $E = S = \bigcup_{j=1}^{\infty} A_j$ und ist die Familie $\{A_j\}_{j \in \mathbb{N}}$ disjunkt, so folgt

$$\mu\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} \mu(A_j).$$

Diese Eigenschaft wird auch σ -Additivität oder abzählbare Additivität genannt.

Satz 18.1.8. Sei $\mu : P(X) \rightarrow [0, \infty]$ ein äußeres Maß und \mathcal{A} sei das Mengensystem der μ -messbaren Mengen. Dann ist \mathcal{A} eine σ -Algebra und es gilt für jede disjunkte Familie $\{A_j\}_{j \in \mathbb{N}}$ von messbaren Mengen aus \mathcal{A} :

$$\mu\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} \mu(A_j).$$

Beweis. Die σ -Additivität folgt schon aus obiger Bemerkung. Für den Nachweis, dass \mathcal{A} eine σ -Algebra ist, bleibt zu zeigen:

Für eine gegebene Familie $\{A_j\}_{j \in \mathbb{N}}$ von messbaren Mengen ist $\bigcup_{j=1}^{\infty} A_j$ ebenfalls messbar. Ist die Vereinigung disjunkt, so folgt dies aus Lemma 18.1.6. Falls die Vereinigung nicht disjunkt ist, betrachte die messbare Familie $\{C_j\}_{j \in \mathbb{N}}$ mit $C_1 = A_1$ und

$$C_j := A_j \setminus \bigcup_{i=1}^{j-1} A_i$$

für $j > 1$.

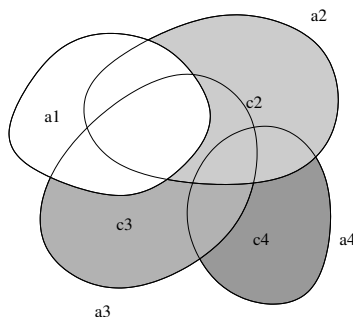


Abbildung 18.5: zum Beweis von Satz 18.1.8

Diese Familie ist auch disjunkt, denn ist $k < j$, so ist $C_k \subset A_k$ und $C_j \cap A_k = \emptyset$.

Da $\bigcup_{j=1}^n C_j = \bigcup_{j=1}^n A_j$ (Induktion über n), ist wegen Lemma 18.1.6

$$\bigcup_{j=1}^{\infty} C_j = \bigcup_{j=1}^{\infty} A_j$$

messbar. □

Definition 18.1.9. Die σ -Algebra der bezüglich des äußeren Lebesguemaßes λ messbaren Mengen des \mathbb{R}^n bezeichnen wir mit $\mathcal{L}(\mathbb{R}^n)$. Ihre Elemente heißen *Lebesgue-messbar*.

Definition 18.1.10. Sei $\mathcal{A} \subset P(X)$ eine σ -Algebra. Eine Abbildung $\mu : \mathcal{A} \rightarrow [0, \infty]$ mit $\mu(\emptyset) = 0$ heißt (σ -additives) *Maß*, falls

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i).$$

für jede disjunkte Familie $A_i \in \mathcal{A}$, $i \in \mathbb{N}$. Ist μ ein Maß auf (X, \mathcal{A}) , so heißt das Tripel (X, \mathcal{A}, μ) *Maßraum*. Die Elemente der σ -Algebra heißen *messbar*.

Bemerkungen.

- (a) σ -additive Maße heißen manchmal auch abzählbar additive Maße. Sprechen wir in Zukunft von Maßen, so meinen wir immer auf einer σ -Algebra definierte σ -additive Maße.
- (b) Insbesondere definiert jedes äußere Maß μ ein Maß auf der σ -Algebra der μ -messbaren Mengen.

Satz 18.1.11 (*Hauptsatz über Maße*). Sei (X, \mathcal{A}, μ) ein Maßraum. Dann gilt:

- (a) Für alle $A, B \in \mathcal{A}$ ist

$$\mu(A \cup B) = \mu(A \setminus B) + \mu(B \setminus A) + \mu(B \cap A).$$

Falls $B \subset A$ und $\mu(B) < \infty$, so folgt insbesondere

$$\mu(A) - \mu(B) = \mu(A \setminus B).$$

und somit die Monotonie $\mu(B) \leq \mu(A)$.

- (b) Ist $\{A_j\}_{j \in \mathbb{N}}$ eine monoton steigende Familie von messbaren Mengen, d.h. $A_1 \subset A_2 \subset \dots \subset A_j \subset A_{j+1} \subset \dots$, so gilt:

$$\lim_{j \rightarrow \infty} \mu(A_j) = \mu\left(\bigcup_{j=1}^{\infty} A_j\right).$$

- (c) Ist $\{A_j\}_{j \in \mathbb{N}}$ eine monoton fallende Familie von messbaren Mengen, d.h. ist $A_1 \supset A_2 \supset \dots \supset A_j \supset A_{j+1} \supset \dots$ und ist $\mu(A_1) < \infty$, so gilt:

$$\lim_{j \rightarrow \infty} \mu(A_j) = \mu\left(\bigcap_{j=1}^{\infty} A_j\right).$$

Bemerkung. Jedes Maß μ besitzt auch die Eigenschaft der Subadditivität, denn ist $\{A_j\}_{j \in \mathbb{N}}$ eine Familie von messbaren Mengen und

$$C_j := A_j \setminus \bigcup_{i=1}^{j-1} A_i,$$

so folgt wie im Beweis des Satzes 18.1.8 zusammen mit der Eigenschaft (a) :

$$\mu\left(\bigcup_{j=1}^{\infty} A_j\right) = \mu\left(\bigcup_{j=1}^{\infty} C_j\right) = \sum_{j=1}^{\infty} \mu(C_j) \leq \sum_{j=1}^{\infty} \mu(A_j).$$

Beweis.

zu (a) Da $A \cup B$ sich als disjunkte Vereinigung dreier Mengen darstellen lässt, nämlich $A \cup B = (A \setminus B) \cup (B \setminus A) \cup (A \cap B)$, folgt aus der Definition des Maßes 18.1.10:

$$\mu(A \cup B) = \mu(A \setminus B) + \mu(B \setminus A) + \mu(A \cap B).$$

Ist $B \subset A$, so ist $B \setminus A = \emptyset, A \cap B = B, A \cup B = A$ und somit

$$\mu(A) = \mu(A \setminus B) + \mu(B).$$

zu (b) Ist

$$A_0 = \emptyset, A_1 \subset A_2 \subset \dots \subset A_j \subset A_{j+1} \subset \dots$$

eine monotone Folge messbarer Mengen, so ist $\{C_j\}_{j \in \mathbb{N}} = \{A_j \setminus A_{j-1}\}_{j \in \mathbb{N}}$ disjunkt.

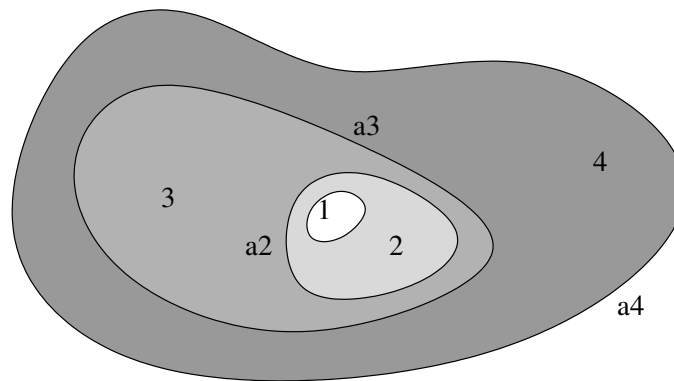


Abbildung 18.6: aufsteigende Mengenfolge

Durch Induktion zeigt man:

$$C_j \cup C_{j-1} \cup \dots \cup C_1 = A_j.$$

Daraus folgt:

$$\bigcup_{j=1}^{\infty} A_j = \bigcup_{j=1}^{\infty} \bigcup_{k=1}^j C_k = \bigcup_{j=1}^{\infty} C_j$$

und somit

$$\mu\left(\bigcup_{j=1}^{\infty} A_j\right) = \mu\left(\bigcup_{j=1}^{\infty} C_j\right) = \sum_{j=1}^{\infty} \mu(C_j) = \lim_{k \rightarrow \infty} \sum_{j=1}^k \mu(C_j) = \lim_{k \rightarrow \infty} \mu\left(\bigcup_{j=1}^k C_j\right) = \lim_{k \rightarrow \infty} \mu(A_k).$$

zu (c) Sei $A_1 \supset A_2 \supset A_3 \supset \dots \supset A_j \supset A_{j+1} \supset \dots$ eine messbare monoton fallende Folge, so betrachte die messbare monoton steigende Folge

$$B_k := A_1 \setminus A_k = A_1 \setminus \bigcap_{j=1}^k A_j = \bigcup_{j=1}^k B_j.$$

Insbesondere ist $\bigcup_{j=1}^{\infty} B_j = A_1 \setminus \bigcap_{j=1}^{\infty} A_j$ messbar und wegen (a) gilt:

$$\mu \left(\bigcup_{k=1}^{\infty} B_k \right) \stackrel{(a)}{=} \mu(A_1) - \mu \left(\bigcap_{k=1}^{\infty} A_k \right).$$

Auf der anderen Seite folgt aus (a) und (b) und der Definition von B_k

$$\mu \left(\bigcup_{k=1}^{\infty} B_k \right) \stackrel{(b)}{=} \lim_{k \rightarrow \infty} \mu(B_k) = \mu(A_1) - \lim_{k \rightarrow \infty} \mu(A_k)$$

und somit die Behauptung. □

Definition 18.1.12. Sei X, \mathcal{A} ein messbarer Raum. Ein Mengensystem $E \subset \mathcal{A}$ heißt *Erzeuger* von \mathcal{A} , falls \mathcal{A} die kleinste σ -Algebra ist, die E enthält. Man schreibt dann $\mathcal{A} = \sigma(E)$.

Sei (X, d) ein metrischer Raum, so nennen wir die von den offenen Mengen erzeugte σ -Algebra die *Borel- σ -Algebra* und bezeichnen sie mit $\mathcal{B}(X)$.

Bemerkungen.

- (a) $\mathcal{B}(X)$ ist somit die kleinste σ -Algebra \mathcal{B} , welche die offenen Mengen in X umfasst. Die Borel- σ -Algebra in einem metrischen Raum wird sowohl durch die offenen Mengen als auch durch die abgeschlossenen Mengen erzeugt.
- (b) Die Borel σ -Algebra $\mathcal{B}(\mathbb{R}^n)$ wird schon von den Quadern erzeugt. Denn jede offene Menge $U \subset \mathbb{R}^n$ lässt sich als abzählbare Vereinigung von Quadern schreiben. Betrachte dazu die Menge aller Quader, deren Eckpunkte rationale Koordinaten besitzen und in U enthalten sind. Diese Menge ist abzählbar und überdeckt U .

Für eine große Klasse von äußeren Maßen μ auf metrischen Räumen ist die Borel- σ -Algebra in der Algebra der μ -messbaren Mengen enthalten. Um ein Kriterium dafür anzugeben, benötigen wir die Definition des Abstandes von Mengen in metrischen Räumen.

Definition 18.1.13. Sei (X, d) ein metrischer Raum und seien A, B nichtleere Teilmengen von X . Dann heißt

$$d(A, B) := \inf\{d(x, y) \mid x \in A, y \in B\}$$

der *Abstand* von A zu B . Ist $A = \{x\}$ einpunktig, so setze

$$d(x, B) = d(\{x\}, B).$$

Ferner heißt $\text{diam}(A) := \sup_{x, y \in A} \{d(x, y)\}$ der *Durchmesser* von A .

Satz 18.1.14 (*Kriterium von Caratheodory*). Sei μ ein äußeres Maß auf X mit

$$\mu(A \cup B) = \mu(A) + \mu(B)$$

für alle Mengen $A, B \subset X$ mit $d(A, B) > 0$. Dann ist die Borel- σ -Algebra in der σ -Algebra der μ -messbaren Mengen enthalten.

Beweis. Es genügt zu zeigen, dass die abgeschlossenen Mengen messbar sind. Sei also $C \subset X$ eine abgeschlossene Menge, so ist zu zeigen

$$\mu(E) \geq \mu(E \cap C) + \mu(E \setminus C)$$

für alle $E \subset X$. Ist $\mu(E) = \infty$, $E \cap C = \emptyset$ oder $E \setminus C = \emptyset$, so ist dies trivialerweise erfüllt. Sei also $\mu(E) < \infty$, $E \cap C \neq \emptyset$ oder $E \setminus C \neq \emptyset$. Definiere

$$C_n := \{x \in X \mid d(x, C) \leq \frac{1}{n}\}.$$

Dann gilt, falls $E \cap C_n \neq \emptyset$:

$$d(E \setminus C_n, E \cap C) \geq d(E \setminus C_n, C) \geq \frac{1}{n}.$$

Also gilt nach Voraussetzung:

$$\mu(E \setminus C_n) + \mu(E \cap C) = \mu((E \setminus C_n) \cup (E \cap C)) \leq \mu(E).$$

Die letzte Beziehung ist für $E \setminus C_n = \emptyset$ trivialerweise erfüllt. Der Satz ist bewiesen, falls Folgendes gilt:

$$\lim_{n \rightarrow \infty} \mu(E \setminus C_n) = \mu(E \setminus C). \tag{18.3}$$

Beweis von (18.3): Betrachte die Menge

$$R_k = \left\{ x \in E \mid \frac{1}{k+1} < d(x, C) \leq \frac{1}{k} \right\} = E \cap (C_k \setminus C_{k+1}).$$

Dann gilt:

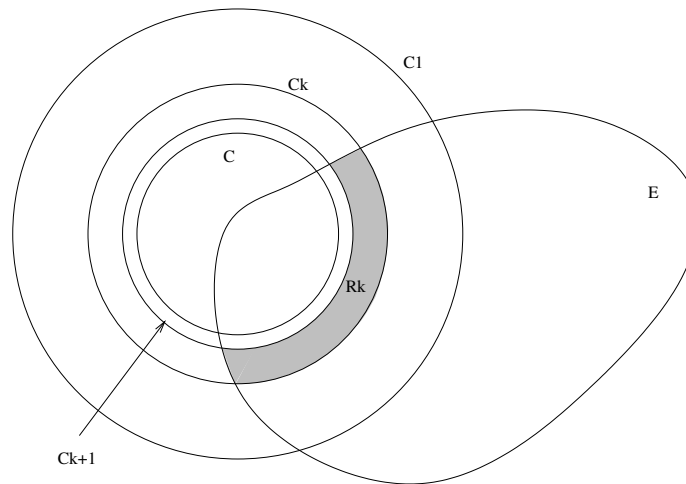


Abbildung 18.7: zu (18.3)

$$E \setminus C = (E \setminus C_n) \cup \bigcup_{k=n}^{\infty} R_k,$$

denn da C abgeschlossen ist, ist $d(x, C) > 0$ für jedes $x \in E \setminus C$. Damit folgt aus der Monotonie und Subadditivität des äußeren Maßes μ :

$$\mu(E \setminus C_n) \leq \mu(E \setminus C) \leq \mu(E \setminus C_n) + \sum_{k=n}^{\infty} \mu(R_k).$$

Um $\lim_{n \rightarrow \infty} \sum_{k=n}^{\infty} \mu(R_k) = 0$ nachzuweisen, reicht es, die Konvergenz von $\sum_{k=1}^{\infty} \mu(R_k)$ zu zeigen. Mit Hilfe der Dreiecksungleichung erhalten wir für $R_k \neq \emptyset$ und $R_{k+2} \neq \emptyset$

$$d(R_k, R_{k+2}) \geq \frac{1}{k+1} - \frac{1}{k+2} > 0.$$

Damit folgt nach Annahme

$$\sum_{k=1}^m \mu(R_{2k}) = \mu\left(\bigcup_{k=1}^m R_{2k}\right) \leq \mu(E).$$

Dies Beziehung gilt auch, falls einige der Mengen R_{2k} leer sind. Genauso folgt:

$$\sum_{k=1}^m \mu(R_{2k+1}) = \mu\left(\bigcup_{k=1}^m R_{2k+1}\right) \leq \mu(E).$$

Addieren wir beide Ungleichungen, so ergibt sich für $m \rightarrow \infty$:

$$\sum_{k=1}^{\infty} \mu(R_k) \leq 2\mu(E) < \infty.$$

□

Wir zeigen nun, dass das äußere Lebesguemaß λ das Kriterium von Caratheodory erfüllt und somit $\mathcal{L}(\mathbb{R}^n)$ die Borel- σ -Algebra $\mathcal{B}(\mathbb{R}^n)$ enthält. Insbesondere sind alle offenen und abgeschlossenen Mengen im \mathbb{R}^n bezüglich des Lebesguemaßes messbar.

Satz 18.1.15. *Das äußere Lebesguemaß λ erfüllt die Bedingung von Caratheodory, d.h. sind $A, B \subset \mathbb{R}^n$ und ist $d(A, B) > 0$, so gilt:*

$$\lambda(A \cup B) = \lambda(A) + \lambda(B).$$

Beweis. Es seien $A, B \subset \mathbb{R}^n$ mit $d(A, B) > 0$. Setze $\delta := \frac{d(A, B)}{2}$. Nach Definition von λ existiert zu jedem $\epsilon > 0$ eine Überdeckung von $A \cup B$ mit Quadern Q_j , so dass

$$\sum_{j=1}^{\infty} \text{vol} Q_j \leq \lambda(A \cup B) + \epsilon.$$

Durch weitere Unterteilung der Quader Q_j können wir annehmen, dass

$$\text{diam}(Q_j) \leq \delta.$$

Seien Q'_k die Teilmenge der Quader aus der Familie der Quader Q_j mit $Q'_k \cap A \neq \emptyset$ und Q''_l die Teilmenge der Quader mit $Q''_l \cap B \neq \emptyset$. Wegen $d(A, B) = 2\delta$ sind die Quader aus der

Familie Q'_k disjunkt zu denen aus der Familie Q''_l . Außerdem überdecken die Quader Q'_k die Menge A und die Quader Q''_l die Menge B . Damit erhalten wir

$$\begin{aligned}\lambda(A) + \lambda(B) &= \sum_k \text{vol} Q'_k + \sum_l \text{vol} Q''_l \leq \sum_j \text{vol} Q_j \\ &\leq \lambda(A \cup B) + \epsilon.\end{aligned}$$

Da dies für jedes $\epsilon > 0$ gilt, folgt

$$\lambda(A) + \lambda(B) \leq \lambda(A \cup B)$$

und somit aus der Subadditivität von λ die Behauptung. \square

Definition 18.1.16. Ist (X, d) metrischer Raum und $\mathcal{A} = \mathcal{B}(X)$ die Borel- σ -Algebra, so heißt μ auch *Borel-Maß*.

Bemerkung. Das äußere Lebesgue-Maß, eingeschränkt auf \mathcal{B} im \mathbb{R}^n , ist ein Borelmaß.

18.2 Messbare Abbildungen

Definition 18.2.1. Es seien (X, \mathcal{A}) und (X', \mathcal{A}') messbare Räume. Eine Abbildung $f : X \rightarrow X'$ heißt *messbar*, falls

$$f^{-1}(A') \in \mathcal{A} \text{ für alle } A' \in \mathcal{A}'.$$

Satz 18.2.2. Sei (X', \mathcal{A}') eine σ -Algebra und $f : X \rightarrow X'$ eine Abbildung. Dann ist

$$f^{-1}\mathcal{A}' = \{f^{-1}(A') \mid A' \in \mathcal{A}'\}$$

eine σ -Algebra über X . Ist E' Erzeuger von \mathcal{A}' , so ist $f^{-1}(E')$ Erzeuger von $f^{-1}\mathcal{A}'$.

Beweis. Sei A'_i eine Familie von Mengen in \mathcal{A}' . Da

$$\bigcup_{i=1}^{\infty} f^{-1}(A'_i) = f^{-1}\left(\bigcup_{i=1}^{\infty} A'_i\right) \in f^{-1}\mathcal{A},$$

ist $A' \in \mathcal{A}'$ so ist auch $X \setminus f^{-1}(A) = f^{-1}(X \setminus A) \in f^{-1}\mathcal{A}$. Der Rest des Beweises sei als Übung überlassen (siehe MfP. IV, Blatt 5). \square

Satz 18.2.3. Seien (X, \mathcal{A}) und (X', \mathcal{A}') messbare Räume und E' ein Erzeuger von \mathcal{A}' . Dann ist $f : X \rightarrow X'$ genau dann messbar, falls $f^{-1}(E') \subset \mathcal{A}$.

Beweis. Ist f messbar, so ist $f^{-1}(E') \subset \mathcal{A}$ für jedes $E' \in \mathcal{A}'$.

Ist umgekehrt $f^{-1}(E') \subset \mathcal{A}$, so ist wegen Satz 18.2.2 die Mengenfamilie $f^{-1}(E')$ ein Erzeuger von $f^{-1}\mathcal{A}'$, d.h. $f^{-1}\mathcal{A}'$ ist die kleinste σ -Algebra, die $f^{-1}(E')$ enthält. Da \mathcal{A} aber nach Annahme auch $f^{-1}(E')$ enthält, folgt: $f^{-1}\mathcal{A}' \subset \mathcal{A}$. Also ist f messbar. \square

Korollar 18.2.4. Seien (X, d) und (X', d') metrische Räume und $f : X \rightarrow X'$ stetig. Sind $\mathcal{B}, \mathcal{B}'$ die Borel- σ -Algebren über X und X' , so ist $f : X \rightarrow X'$ messbar.

Beweis. Da die offenen Mengen Erzeuger der Borel- σ -Algebren sind und für stetige Abbildungen das Urbild offener Mengen offen ist, folgt die Behauptung aus Satz 18.2.2. \square

Mittels messbarer Abbildungen kann man σ -additive Maße auf σ -additive Maße abbilden.

Satz 18.2.5. Seien $(X, \mathcal{A}), (X', \mathcal{A}')$ messbare Räume und μ ein σ -additives Maß auf (X, \mathcal{A}) . Sei $T : X \rightarrow X'$ messbar, so definiert

$$(T\mu)(A') := \mu(T^{-1}A')$$

ein σ -additives Maß auf (X', \mathcal{A}') . $T\mu$ heißt das Bildmaß von μ bzgl. T .

Beweis. Zu zeigen ist nur die σ -Additivität. Sei $\{A'_k\}_{k \in \mathbb{N}}$ eine disjunkte Mengenfamilie mit $A'_k \in \mathcal{A}'$. Dann sind auch die Mengen $T^{-1}A'_k$ disjunkt und es gilt:

$$T\mu\left(\bigcup_{k=1}^{\infty} A'_k\right) := \mu\left(T^{-1}\bigcup_{k=1}^{\infty} A'_k\right) = \mu\left(\bigcup_{k=1}^{\infty} T^{-1}A'_k\right) = \sum_{k=1}^{\infty} \mu(T^{-1}A'_k) = \sum_{k=1}^{\infty} T\mu(A'_k).$$

□

Satz 18.2.6. Seien $T_1 : (X, \mathcal{A}) \rightarrow (X', \mathcal{A}')$ und $T_2 : (X', \mathcal{A}') \rightarrow (X'', \mathcal{A}'')$ messbare Abbildungen, so ist $T_2 \circ T_1 : (X, \mathcal{A}) \rightarrow (X'', \mathcal{A}'')$ messbar und es gilt:

$$(T_2 \circ T_1)\mu = T_2(T_1\mu).$$

Beweis. Sei $A'' \in \mathcal{A}''$, so gilt:

$$(T_2 \circ T_1)\mu(A'') = \mu(T_1^{-1}T_2^{-1}A'') = T_1\mu(T_2^{-1}A'') = T_2(T_1\mu)(A'').$$

□

Definition 18.2.7. Sei $T : (X, \mathcal{A}) \rightarrow (X, \mathcal{A})$ messbar und μ ein Maß auf (X, \mathcal{A}) . Dann heißt μ T -invariant, falls

$$(T\mu)(A) = \mu(T^{-1}A) = \mu(A)$$

für alle $A \in \mathcal{A}$ gilt.

18.3 Translationsinvarianz des Lebesguemaßes

Definition 18.3.1. Sei $A \in \mathbb{R}^n$, so heißt die Abbildung $T_a : \mathbb{R}^n \rightarrow \mathbb{R}^n$ mit $T_a(x) = a + x$ Translation.

Bemerkung. Da T_a stetig ist, ist T_a bezüglich der Borel- σ -Algebren messbar. Da $T_a^{-1} = T_{-a}$, ist die Inverse ebenfalls eine Translation.

Satz 18.3.2. Das Lebesgue-Maß ist translationsinvariant.

Beweis. Sei Q ein Quader, so ist

$$T_a Q = \{x + a \mid x \in Q\} =: Q + a$$

ebenfalls ein Quader mit gleicher Kantenlänge wie Q . Daher gilt: $\text{vol}T_a Q = \text{vol}Q$ Sei $A \subset \mathbb{R}^n$, so gilt somit:

$$\begin{aligned} \lambda(T_a^{-1}A) &= \inf\left\{\sum_{j=1}^{\infty} \text{vol}Q_j \mid \bigcup_{j=1}^{\infty} Q_j \supset T_a^{-1}A\right\} \\ &= \inf\left\{\sum_{j=1}^{\infty} \text{vol}T_a Q_j \mid \bigcup_{j=1}^{\infty} T_a Q_j \supset A\right\} \\ &= \inf\left\{\sum_{j=1}^{\infty} \text{vol}Q'_j \mid \bigcup_{j=1}^{\infty} Q'_j \supset A\right\} \\ &= \lambda(A). \end{aligned}$$

□

Bemerkung. Es gibt jedoch auch andere translationsinvariante Maße auf \mathbb{R}^n . Definiere $\mu : P(\mathbb{R}^n) \rightarrow [0, \infty]$ durch

$$\mu(A) = \begin{cases} 1 & \text{für } A \neq \emptyset, \\ 0 & \text{sonst.} \end{cases}$$

Dann ist $\mu(T_a A) = \mu(A)$. Allerdings besteht die σ -Algebra der μ -messbaren Menge nur aus der leeren Menge und \mathbb{R}^n .

Satz 18.3.3. Sei μ ein auf der Borel- σ -Algebra $\mathcal{B} = \mathcal{B}(\mathbb{R}^n)$ definiertes translationsinvariantes Maß mit

$$\mu([0, 1]^n) = 1.$$

Dann stimmen μ und λ überein.

Beweis.

(1) Das μ -Maß jeder beschränkten Borel-Menge ist endlich, denn für jedes $k \in \mathbb{N}$ ist

$$[-k, k]^n \subset \bigcup_{a=(k_1, \dots, k_n) \in \mathbb{Z}^n, |k_i| \leq k} (T_a([0, 1]^n))$$

und daher ist wegen der Subadditivität von μ

$$\mu([-k, k]^n) \leq \mu([0, 1]^n) \text{card}\{(k_1, \dots, k_n) \in \mathbb{Z}^n \mid |k_i| \leq k\} = 1 \cdot (2k + 1)^n.$$

Da jede beschränkte Menge von einem Würfel der obigen Form überdeckt wird, folgt die Behauptung.

(2) Der Rand ∂Q eines Quaders $Q \subset \mathbb{R}^n$ hat Maß null bzgl. μ . Denn ist $Q = I_1 \times \dots \times I_n$ ein Quader, so besteht der Rand aus Ebenen der Form

$$I_1 \times \dots \times I_{k-1} \times a \times I_{k+1} \times \dots \times I_n$$

mit $a \in \mathbb{R}$. Da μ translationsinvariant, monoton und σ -additiv auf den Borelmengen ist, ist wegen (1)

$$\infty > \mu(I_1 \times \dots \times I_{k-1} \times [a, a+1] \times I_{k+1} \times \dots \times I_n) \quad (18.4)$$

$$> \mu\left(\bigcup_{j=1}^{\infty} I_1 \times \dots \times I_{k-1} \times \left\{a + \frac{1}{j}\right\} \times I_{k+1} \times \dots \times I_n\right) \quad (18.5)$$

$$= \sum_{j=1}^{\infty} \mu(T_1 \times \dots \times I_{k-1} \times a \times I_{k+1} \times \dots \times I_n) \quad (18.6)$$

und daher ist $\mu(I_1 \times \dots \times I_{k-1} \times a \times I_{k+1} \times \dots \times I_n) = 0$.

- (3) Bezeichne für jedes $k \in \mathbb{N}$ und $m = (m_1, \dots, m_n)$, $m_i \in \{1, \dots, k-1\}$ mit $Q(m, k)$ den Würfel (Quader mit gleichen Seitenlängen) in \mathbb{R}^n gegeben durch

$$Q(m, k) = \left[\frac{m_1}{k}, \frac{m_1+1}{k}\right] \times \dots \times \left[\frac{m_n}{k}, \frac{m_n+1}{k}\right].$$

Jeder dieser Würfel ist Translat von $Q(0, k) = [0, \frac{1}{k}] \times \dots \times [0, \frac{1}{k}]$, denn

$$Q(m, k) = T_a\left([0, \frac{1}{k}] \times \dots \times [0, \frac{1}{k}]\right)$$

mit $a = (\frac{m_1}{k}, \dots, \frac{m_n}{k})$. Wegen der Translationsinvarianz von μ besitzen für festes k die Würfel $Q(m, k)$ das gleiche Maß, d.h.

$$\mu(Q(m, k)) = \mu(Q(0, k)).$$

Da

$$\bigcup_{\substack{m=(m_1, \dots, m_n) \\ m_i \in \{0, \dots, k-1\}}} Q(m, k) = [0, 1]^n$$

und das Innere der Quader disjunkt ist, gilt:

$$1 = \mu\left(\bigcup_{\substack{m=(m_1, \dots, m_n) \\ m_i \in \{0, \dots, k-1\}}} Q(m, k)\right) = \sum_{\substack{m=(m_1, \dots, m_n) \\ m_i \in \{0, \dots, k-1\}}} \mu(Q(m, k)) = \mu(Q(0, k)) \cdot k^n.$$

Also folgt für alle Würfel der Form $Q(m, k)$

$$\mu(Q(m, k)) = \frac{1}{k^n} = \lambda(Q(m, k)).$$

Da μ translationsinvariant ist, ist $\mu(Q) = \lambda(Q)$ für alle Q der Form

$$\left[\frac{m_1}{k}, \frac{m_1+1}{k}\right] \times \dots \times \left[\frac{m_k}{k}, \frac{m_k+1}{k}\right] \text{ mit } m_k \in \mathbb{Z}.$$

Ist nun Q ein Quader, dessen Eckpunkte rationale Koordinaten haben, so lässt er sich als endliche Vereinigung von Würfeln der obigen Form schreiben, deren Inneres paarweise disjunkt ist. Aufgrund der σ -Additivität stimmen somit Lebesguemaß und μ auf allen rationalen Quadern überein. Da die rationalen Zahlen dicht in \mathbb{R} liegen, existiert zu

einem beliebigen Quader Q eine aufsteigende Folge $Q_1 \subset Q_2 \subset \dots \subset Q_j \subset \dots$ rationaler Quader mit

$$\bigcup_{j=1}^{\infty} Q_j = Q.$$

Aus dem Hauptsatz über messbare Mengen folgt somit:

$$\mu(Q) = \lim_{j \rightarrow \infty} \mu(Q_j) = \lim_{j \rightarrow \infty} \lambda(Q_j) = \lambda(Q).$$

- (4) Sei nun $A \subset \mathbb{R}^n$ eine Borel-messbare Menge und Q_j eine Familie von Quadern, die A überdeckt. Dann gilt

$$\mu(A) \leq \sum_{j=1}^{\infty} \mu(Q_j) = \sum_{j=1}^{\infty} \lambda(Q_j).$$

Nun folgt aus der Definition des Lebesguemaßes: $\mu(A) \leq \lambda(A)$. Sei nun A eine beschränkte Borelmengung, so ist sie in einem Quader Q enthalten. Dann folgt mit Hilfe des Hauptsatzes

$$\mu(A) = \mu(Q) - \mu(Q \setminus A) = \lambda(Q) - \mu(Q \setminus A) \geq \lambda(Q) - \lambda(Q \setminus A) = \lambda(A)$$

und somit stimmt μ auf allen beschränkten Borelmengen überein.

- (5) Ist A beliebige Borelmengung und $Q_k = [-k, k]^n$, so ist $A_k := A \cap Q_k$ eine monoton steigende Folge von Borelmengen. Da $A = \bigcup_{k=1}^{\infty} A_k$, folgt wiederum aus dem Hauptsatz über Maße:

$$\lambda(A) = \lim_{k \rightarrow \infty} \lambda(A_k) = \lim_{k \rightarrow \infty} \mu(A_k) = \mu(A)$$

und somit ist $\lambda = \mu$.

□

Korollar 18.3.4. Sei μ ein translationsinvariantes Borelmaß auf \mathbb{R}^n mit

$$\mu([0, 1]^n) = \alpha < \infty.$$

Dann ist $\mu = \alpha \cdot \lambda$.

Beweis. Ist $\alpha > 0$, so ist $\frac{1}{\alpha}\mu$ ein translationsinvariantes Maß mit $\frac{1}{\alpha}\mu([0, 1]^n) = 1$. Daher ist $\frac{1}{\alpha}\mu = \lambda$. Ist $\alpha = 0$, so ist μ wegen der Translationsinvarianz trivial. □

Nun zeigen wir, wie sich das Lebesgue-Maß unter bijektiven affinen linearen Abbildungen transformiert.

Satz 18.3.5. Sei $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ eine bijektive affine lineare Abbildung mit $F(x) = a + Lx$ und $L \in GL(n, \mathbb{R})$ sowie $a \in \mathbb{R}^n$. Dann ist F bezüglich der Borel- σ -Algebra messbar und es gilt:

$$F(\lambda)(A) = \lambda(F^{-1}(A)) = \frac{1}{|\det L|} \cdot \lambda(A)$$

für alle Borel-messbaren Mengen A .

Beweis.

- (1) Wegen der Stetigkeit von F ist F auch Borel-messbar. Da $F = T_a \circ L$, ist $F(\lambda) = T_a(L(\lambda))$. Daher genügt es zu zeigen: $L(\lambda) = \frac{1}{|\det L|} \cdot \lambda$, denn dann ist

$$F(\lambda)(A) = T_a(L(\lambda))(A) = T_a\left(\frac{1}{|\det L|} \lambda(A)\right) = \frac{1}{|\det L|} \cdot \lambda(T_{-a}(A)) = \frac{1}{|\det L|} \cdot \lambda(A).$$

- (2) $L(\lambda)$ ist ein translationsinvariantes Borelmaß, denn ist $A \in \mathcal{B}(\mathbb{R}^n)$, so gilt:

$$\begin{aligned} T_a(L(\lambda))(A) &= L(\lambda)(T_{-a}A) = \lambda(L^{-1}(T_{-a}A)) = \lambda(L^{-1}(A - a)) \\ &= \lambda(L^{-1}(A) - L^{-1}(a)) = \lambda(L^{-1}(A)) \\ &= L(\lambda)(A) \end{aligned}$$

für alle Borelmengen $A \in \mathcal{B}(\mathbb{R}^n)$ und $a \in \mathbb{R}^n$.

- (3) Da $L(\lambda)([0, 1]^n) = \lambda(L^{-1}[0, 1]^n) < \infty$ und $L(\lambda)$ translationsinvariant ist, existiert wegen Korollar 18.3.4 ein $\alpha = \alpha(L) > 0$ mit $L(\lambda) = \alpha \cdot \lambda$. Wir werden nun $\alpha(L)$ schrittweise berechnen.

- (i) Sei L orthogonal, so ist $\|Lx\| = \|x\|$ für alle $x \in \mathbb{R}^n$ und somit $L^{-1}(B(0, 1)) = B(0, 1)$. Daraus folgt für α :

$$\alpha \lambda(B(0, 1)) = L(\lambda)(B(0, 1)) = \lambda(L^{-1}(B(0, 1))) = \lambda(B(0, 1))$$

und somit $\alpha = 1$. Da $|\det L| = 1$, ist der Satz für orthogonale lineare Abbildungen bewiesen.

- (ii) Sei nun $e_1, \dots, e_n \in \mathbb{R}^n$ die kanonische Basis des \mathbb{R}^n und $L \in \text{GL}(n, \mathbb{R})$ mit $Le_i = a_i e_i$ und $a_i > 0$. Dann folgt:

$$L^{-1}([0, 1]^n) = \left[0, \frac{1}{a_1}\right] \times \dots \times \left[0, \frac{1}{a_n}\right]$$

und wir erhalten für α :

$$\alpha = \lambda(L^{-1}([0, 1]^n)) = \frac{1}{a_1 \cdot \dots \cdot a_n} = \frac{1}{|\det L|}.$$

- (iii) Sei $L \in \text{GL}(n, \mathbb{R})$ beliebig, so ist LL^t eine symmetrische Abbildung mit positiven Eigenwerten β_i . Daher existiert wegen Korollar 13.6.4 eine orthogonale lineare Abbildung $C \in O(n)$ und eine lineare diagonalisierte Abbildung D mit $De_i = \sqrt{\beta_i} e_i$ und $LL^t = CD^2C^t$. Außerdem ist $(\det L)^2 = \det LL^t = (\det D)^2$ und somit $|\det L| = \det D$. Aus der Orthogonalität von C folgt:

$$L = CDD^{-1}C^tL = CDG,$$

wobei $G = D^{-1}C^tL$. Wegen $LL^t = CD^2C^t$ und

$$GG^t = D^{-1}C^tLL^tC(D^{-1})^t = D^{-1}C^tCD^2C^tCD^{-1} = D^{-1}D^2D^{-1} = \text{id}$$

ist auch die lineare Abbildung G orthogonal. Aus (i) und (ii) folgt mit

$$L(\lambda) = CDG(\lambda) = CD(\lambda) = C\left(\frac{1}{\det D} \lambda\right) = \frac{1}{\det D} C(\lambda) = \frac{1}{|\det L|} \cdot \lambda$$

die Behauptung. □

18.4 Messbare Funktionen

Auf der Maßtheorie aufbauend entwickeln wir nun die Integrationstheorie. Dabei betrachten wir reellwertige messbare Funktionen, die aber auch ∞ und $-\infty$ als Wert annehmen können. Daher müssen wir zunächst den Begriff der offenen Menge von \mathbb{R} auf $\overline{\mathbb{R}}$ fortsetzen. Sei

$$\overline{\mathbb{R}} = \{-\infty\} \cup \mathbb{R} \cup \{+\infty\}.$$

Die Anordnung “ $<$ ” von \mathbb{R} wird auf $\overline{\mathbb{R}}$ durch $-\infty < x < \infty$ für alle $x \in \mathbb{R}$ fortgesetzt. Die Begriffe Supremum und Infimum, die in Definition 1.5.2 nur für nach oben beschränkte bzw. nach unten beschränkte Teilmengen von \mathbb{R} erklärt wurden, lassen sich auf $\overline{\mathbb{R}}$ wie folgt ausdehnen: Ist $A \subset \overline{\mathbb{R}}$, so setze

$$\sup A = \begin{cases} \sup A & \text{falls } A \subset \mathbb{R} \text{ nach oben beschränkt,} \\ \infty & \text{sonst} \end{cases}$$

bzw.

$$\inf A = \begin{cases} \inf A & \text{falls } A \subset \mathbb{R} \text{ nach unten beschränkt,} \\ -\infty & \text{sonst.} \end{cases}$$

Die Intervalle $(a, \infty] = \{x \in \overline{\mathbb{R}} \mid x > a\}$ und $[-\infty, a) = \{x \in \overline{\mathbb{R}} \mid x < a\}$ heißen *offene Intervalle*. Zusammen mit den offenen Intervallen auf \mathbb{R} bilden sie die Menge aller offenen Intervalle auf $\overline{\mathbb{R}}$.

Ein Punkt $x_0 \in A \subset \overline{\mathbb{R}}$ heißt *innerer Punkt* von A , falls ein offenes Intervall I existiert mit $x_0 \in I \subset A$. A heißt *offen*, falls A nur aus inneren Punkten besteht.

Mit dieser Festsetzung lässt sich der Konvergenzbegriff auf $\overline{\mathbb{R}}$ ausdehnen (siehe auch den Abschnitt über uneigentliche Grenzwerte, insbesondere Definition 3.5.2).

Definition 18.4.1. Sei $(x_n)_{n \in \mathbb{N}} \in \overline{\mathbb{R}}$ eine Folge, so schreiben wir $\lim_{n \rightarrow \infty} x_n = c \in \overline{\mathbb{R}}$, falls in jeder offenen Umgebung von c fast alle (d.h. alle bis auf endlich viele) Folgenglieder liegen.

Bemerkung. Da in jeder offenen Umgebung ein offenes Intervall liegt, reicht es, offene Intervalle zu betrachten. Ist $c = \infty$, so bedeutet dies:

$$\forall a > 0 \quad \exists n_0 \quad \forall n \geq n_0 : x_n > a.$$

Definition 18.4.2. Sei $(x_n)_{n \in \mathbb{N}} \in \overline{\mathbb{R}}$ eine Folge, so heißt $c \in \overline{\mathbb{R}}$ *Häufungspunkt (HP)* von $(x_n)_{n \in \mathbb{N}}$, falls in jeder offenen Umgebung von c unendlich viele Folgenglieder x_n liegen. Desweiteren definiert man:

$$\overline{\lim}_{n \rightarrow \infty} x_n = \sup\{a \mid a \text{ HP von } (x_n)_{n \in \mathbb{N}}\}$$

und

$$\underline{\lim}_{n \rightarrow \infty} x_n = \inf\{a \mid a \text{ HP von } (x_n)_{n \in \mathbb{N}}\}.$$

Bemerkungen.

- (a) Jede Folge $(x_n)_{n \in \mathbb{N}}$ in $\overline{\mathbb{R}}$ hat einen Häufungspunkt, denn ist die Folge beschränkt, so folgt dies aus dem Satz 2.3.5 von Bolzano-Weierstraß (siehe auch Satz 2.3.11).
Ist $(x_n)_{n \in \mathbb{N}}$ nicht nach oben (unten) beschränkt, so ist $+\infty$ ($-\infty$) ein Häufungspunkt.

Insbesondere existiert in diesem verallgemeinerten Sinne der Limes superior und Limes inferior für eine beliebige Folge. Ist die Folge $(x_n)_{n \in \mathbb{N}}$ nicht nach oben beschränkt bzw. nicht nach unten beschränkt, so gilt

$$\overline{\lim}_{n \rightarrow \infty} x_n = \infty \text{ bzw. } \underline{\lim}_{n \rightarrow \infty} x_n = -\infty.$$

(b) Sei $(a_n)_{n \in \mathbb{N}}$ die monoton fallende Folge mit $a_n = \sup\{x_k \mid k \geq n\}$, so gilt:

$$\overline{\lim}_{n \rightarrow \infty} x_n = \lim_{n \rightarrow \infty} a_n = \inf\{a_n \mid n \in \mathbb{N}\}.$$

Genauso gilt für die monoton steigende Folge $b_n = \inf\{x_k \mid k \geq n\}$:

$$\underline{\lim}_{n \rightarrow \infty} x_n = \lim_{n \rightarrow \infty} b_n = \inf\{b_n \mid n \in \mathbb{N}\}.$$

Aus Gründen, die im Folgenden ersichtlich werden, ist es in der Maßtheorie zweckmäßig, die Rechenregeln in \mathbb{R} auf $\overline{\mathbb{R}}$ teilweise auszudehnen. Allerdings wird $\overline{\mathbb{R}}$ mit diesen erweiterten Rechenregeln nicht zu einem Körper, ja nicht einmal zu einer abelschen Gruppe bezüglich der Addition. Für die Addition setzen wir

$$a + (\pm\infty) = (\pm\infty) + a = \pm\infty$$

für alle $a \in \mathbb{R}$ und

$$+\infty + \infty = +\infty \text{ sowie } (-\infty) + (-\infty) = -\infty.$$

Für die Multiplikation setzt man

$$a \cdot (\pm\infty) = \begin{cases} \pm\infty & \text{für } a \in (0, \infty], \\ \mp\infty & \text{für } a \in [-\infty, 0). \end{cases}$$

Weitere zweckmäßige Festlegungen sind (diese werden nicht in allen Büchern über Maßtheorie zugelassen)

$$0 \cdot (\pm\infty) = (\pm\infty) \cdot 0 = 0 \text{ und } +\infty + (-\infty) = (-\infty) + (+\infty) = 0.$$

Diese Regeln sind mit Vorsicht zu genießen! Zum Beispiel gelten weder das Assoziativgesetz bezüglich der Addition, noch das Distributivgesetz, denn

$$(-\infty) = (+\infty + (-\infty)) + (-\infty) \text{ und } (+\infty) + ((-\infty)) + (-\infty) = (+\infty) + ((-\infty)) = 0$$

sowie

$$+\infty = (5 - 3)(+\infty) \text{ und } 5(+\infty) - 3(+\infty) = (+\infty) + (-\infty) = 0.$$

Definition 18.4.3. Wir bezeichnen mit $\mathcal{B}(\overline{\mathbb{R}})$ die Borel- σ -Algebra von $\overline{\mathbb{R}}$, d.h. die σ -Algebra, die von den offenen Mengen in $\overline{\mathbb{R}}$ erzeugt wird.

Bemerkung. Jedes einzelne der vier Mengensysteme

$$\{(a, \infty] \mid a \in \mathbb{R}\}, \{[a, \infty] \mid a \in \mathbb{R}\}$$

$$\{[-\infty, a) \mid a \in \mathbb{R}\}, \{[-\infty, a] \mid a \in \mathbb{R}\}$$

ist ein Erzeuger von $\mathcal{B}(\overline{\mathbb{R}})$. Der einfache Beweis sei als Übung überlassen.

Satz 18.4.4. Sei \mathcal{A} eine σ -Algebra über X und $f : X \rightarrow \overline{\mathbb{R}}$ eine Funktion. Dann sind äquivalent:

- (a) f ist $(\mathcal{A}, \mathcal{B}(\overline{\mathbb{R}}))$ -messbar,
- (b) $f^{-1}(a, \infty]$ ist messbar für alle $a \in \mathbb{R}$,
- (c) $f^{-1}[a, \infty]$ ist messbar für alle $a \in \mathbb{R}$,
- (d) $f^{-1}[-\infty, a)$ ist messbar für alle $a \in \mathbb{R}$,
- (e) $f^{-1}[-\infty, a]$ ist messbar für alle $a \in \mathbb{R}$.

Bemerkung. Funktionen $f : X \rightarrow \overline{\mathbb{R}}$ werden in der Maßtheorie im Unterschied zu den reellwertigen Funktionen $f : X \rightarrow \mathbb{R}$ oft *numerische Funktionen* genannt. Da wir die Wertebereiche dieser Funktionen immer angeben, können wir auf diese Unterscheidung verzichten. Wollen wir numerische Funktionen addieren oder multiplizieren, so werden wir die obigen Rechenregeln mit den entsprechenden Vorsichtsmaßnahmen verwenden.

Beweis. Da jedes einzelne der Mengensysteme

$$\{(a, \infty] \mid a \in \mathbb{R}\}, \{[a, \infty] \mid a \in \mathbb{R}\}, \{[-\infty, a) \mid a \in \mathbb{R}\} \text{ und } \{[-\infty, a] \mid a \in \mathbb{R}\}$$

$\mathcal{B}(\overline{\mathbb{R}})$ erzeugt, folgt die Behauptung des Satzes aus Satz 18.2.2 (Urbild des Erzeugendensystems einer σ -Algebra messbar $\Leftrightarrow f$ messbar). \square

Ist $f_n : X \rightarrow \overline{\mathbb{R}}$ eine Folge von Funktionen, so definiere

$$(\sup_{n \in \mathbb{N}} f_n)(x) := \sup\{f_n(x) \mid n \in \mathbb{N}\} \text{ und } (\inf_{n \in \mathbb{N}} f_n)(x) := \inf\{f_n(x) \mid n \in \mathbb{N}\},$$

sowie

$$(\overline{\lim}_{n \rightarrow \infty} f_n)(x) = \overline{\lim}_{n \rightarrow \infty} f_n(x) \text{ und } (\underline{\lim}_{n \rightarrow \infty} f_n)(x) = \underline{\lim}_{n \rightarrow \infty} f_n(x).$$

Die Funktion $\sup f_n$ heißt *obere Einhüllende*, die Funktion $\inf f_n$ heißt *untere Einhüllende* der Folge $(f_n)_{n \in \mathbb{N}}$.

Satz 18.4.5. Für jede Folge $(f_n)_{n \in \mathbb{N}}$ von messbaren Funktionen $f_n : (X, \mathcal{A}) \rightarrow (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ sind die Funktionen

$$\sup_{n \in \mathbb{N}} f_n, \inf_{n \in \mathbb{N}} f_n, \overline{\lim}_{n \rightarrow \infty} f_n \text{ und } \underline{\lim}_{n \rightarrow \infty} f_n$$

messbar. Insbesondere ist $\lim_{n \rightarrow \infty} f_n$ messbar, falls der punktweise Grenzwert existiert.

Beweis. Da für alle $a \in \mathbb{R}$ die Mengen $(\sup_{n \in \mathbb{N}} f_n)^{-1}[-\infty, a]$ und $(\inf_{n \in \mathbb{N}} f_n)^{-1}[a, \infty]$ wegen

$$(\sup_{n \in \mathbb{N}} f_n)^{-1}[-\infty, a] = \{x \in X \mid f_n(x) \leq a \text{ für alle } n \in \mathbb{N}\} = \bigcap_{n=1}^{\infty} f_n^{-1}[-\infty, a]$$

und

$$(\inf_{n \in \mathbb{N}} f_n)^{-1}[a, \infty] = \{x \in X \mid f_n(x) \geq a \text{ für alle } n \in \mathbb{N}\} = \bigcap_{n=1}^{\infty} f_n^{-1}[a, \infty]$$

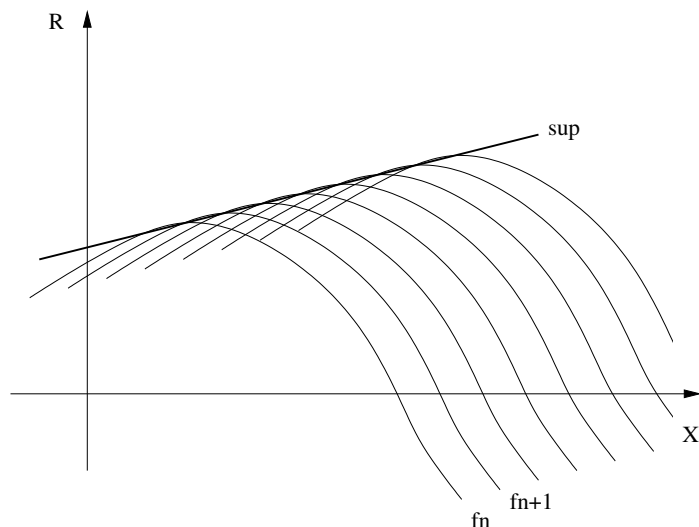


Abbildung 18.8: Einhüllende einer Funktionenfolge

messbar sind, folgt die Messbarkeit von $\sup_{n \in \mathbb{N}} f_n$ und $\inf_{n \in \mathbb{N}} f_n$ aus Satz 18.4.4. Aus der alternativen Darstellung des Limes superior (siehe obige Bemerkung) folgt:

$$\overline{\lim}_{n \rightarrow \infty} f_n(x) = \lim_{n \rightarrow \infty} \sup\{f_k(x) \mid k \geq n\} = \inf_{n \in \mathbb{N}} \sup\{f_k(x) \mid k \geq n\}$$

und somit ergibt sich aus der Messbarkeit von

$$F_n(x) := \sup\{f_k(x) \mid k \geq n\}$$

auch die Messbarkeit von $\overline{\lim}_{n \rightarrow \infty} f_n(x) = \inf_{n \in \mathbb{N}} F_n$. Da

$$\underline{\lim}_{n \rightarrow \infty} f_n(x) = \lim_{n \rightarrow \infty} \inf\{f_k(x) \mid k \geq n\} = \sup_{n \in \mathbb{N}} \inf\{f_k(x) \mid k \geq n\},$$

folgt auch die Messbarkeit des Limes inferior. \square

Korollar 18.4.6. Sind $f_1, \dots, f_n : (X, \mathcal{A}) \rightarrow (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ messbar, so auch $\max(f_1, \dots, f_n)$ und $\min(f_1, \dots, f_n)$.

Beweis. Betrachte die Folge $f_1, \dots, f_{n-1}, f_n, f_n, f_n, \dots$ und wende Satz 18.4.5 an. \square

Wir wollen zeigen, dass die Summe und das Produkt messbarer Funktionen messbar sind. Dazu beweisen wir zunächst folgendes Lemma.

Lemma 18.4.7. Sei $f : (X, \mathcal{A}) \rightarrow (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ eine Abbildung mit $f(x) = (f_1(x), \dots, f_n(x))$, $f_i : X \rightarrow \mathbb{R}$. Dann ist f genau dann messbar, falls ihre Komponenten f_j messbar sind.

Beweis. Die Projektionen $p_j : \mathbb{R}^n \rightarrow \mathbb{R}$ mit $p_j(x) = x_j$ sind Borel-messbar, da sie stetig sind. Ist f messbar, so auch $f_j = p_j \circ f$.

Seien f_1, \dots, f_n messbar und $Q = I_1 \times \dots \times I_n \subset \mathbb{R}^n$ ein Quader. Dann ist

$$f^{-1}(Q) = \bigcap_{j=1}^n f_j^{-1} I_j \in \mathcal{A},$$

denn $f_j^{-1}I_j \in \mathcal{A}$. Da die Quader die Borel- σ -Algebra von \mathbb{R}^n erzeugen, ist f wegen Satz 18.2.2 messbar. \square

Satz 18.4.8. Sind $f, g : (X, \mathcal{A}) \rightarrow (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ messbar und $a, \beta \in \overline{\mathbb{R}}$. Dann sind auch $\alpha f + \beta g$, $f \cdot g$ und $|f|$ messbar.

Beweis. Seien $f, g : X \rightarrow \mathbb{R}$ reellwertig und messbar. Dann ist $h : X \rightarrow \mathbb{R}^2$ mit $h(x) = (f(x), g(x))$ messbar. Da die Funktionen $s, p : \mathbb{R}^2 \rightarrow \mathbb{R}$ mit $s(x_1, x_2) = x_1 + x_2$ und $p(x_1, x_2) = x_1 \cdot x_2$ Borel-messbar sind, sind

$$f(x) + g(x) = s \circ h(x), \text{ sowie } f(x) \cdot g(x) = p \circ h(x)$$

messbar. Es seien nun $f, g : X \rightarrow \overline{\mathbb{R}}$ messbar. Dann sind die Funktionen $f_n, g_n : X \rightarrow \mathbb{R}$ mit

$$f_n(x) = \max(-n, \min(f(x), n)), \text{ sowie } g_n(x) = \max(-n, \min(g(x), n))$$

messbar. Wegen

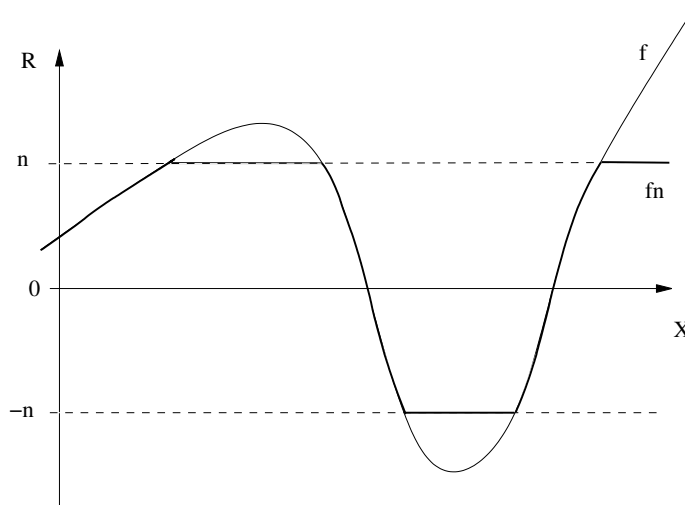


Abbildung 18.9: zum Beweis von Satz 18.4.8

$$(f + g)(x) = \lim_{n \rightarrow \infty} (f_n + g_n)(x) \text{ und } (f \cdot g)(x) = \lim_{n \rightarrow \infty} f_n(x) \cdot g_n(x)$$

folgt die Messbarkeit von $f \cdot g$ und $f + g$. Da die konstanten Funktionen messbar sind, sind auch die Funktionen $\alpha f, \beta g$ und $\alpha f + \beta g$ messbar. Insbesondere ist $-f$ messbar und damit auch $|f| = \max(f, -f)$. \square

Für jede Funktion $f : X \rightarrow \overline{\mathbb{R}}$ definiere ihren *positiven Anteil* $f^+ : X \rightarrow [0, \infty]$ und *negativen Anteil* $f^- : X \rightarrow [0, \infty]$ durch

$$f^+(x) = \max\{f(x), 0\} \text{ und } f^-(x) = \max\{-f(x), 0\} = (-f)^+(x).$$

Dann gilt: $f = f^+ - f^-$ und $|f| = f^+ + f^-$.

Korollar 18.4.9. Eine Funktion $f : (X, \mathcal{A}) \rightarrow (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ ist genau dann messbar, falls f^+ und f^- messbar sind.

Beweis. Ist f messbar, so auch $-f$ und wegen Korollar 18.4.6 auch f^+ und f^- . Sind f^+ und f^- messbar, so auch $f = f^+ - f^-$. \square

Von großer Bedeutung für die Integrationstheorie sind die messbaren Funktionen, die nur endlich viele Werte annehmen. Sie sind Verallgemeinerungen der in der elementaren Integrationstheorie eingeführten Treppenfunktionen (siehe Kapitel 5, Definition 5.1.1).

Definition 18.4.10. Eine messbare Funktion $f : (X, \mathcal{A}) \rightarrow (\mathbb{R}, \mathcal{B})$ heißt \mathcal{A} -Treppenfunktion, falls f nur endlich viele Werte annimmt. Es sei $T(X, \mathcal{A})$ die Menge aller \mathcal{A} -Treppenfunktionen und $T^+(X, \mathcal{A}) = \{f \in T \mid f \geq 0\}$ die der nicht negativen Treppenfunktionen. Die nicht negativen Treppenfunktionen werden manchmal auch *elementare Funktionen* genannt.

Bemerkungen.

(a) Ist $f \in T(X, \mathcal{A})$, so ist $f(X) = \{\alpha_1, \dots, \alpha_m\} \in \mathbb{R}$. Insbesondere sind die Mengen

$$A_j = f^{-1}(\alpha_j) \in \mathcal{A}$$

paarweise disjunkt und messbar und ihre Vereinigung ist X . Dann gilt:

$$f = \sum_{j=1}^m \alpha_j \chi_{A_j} = \sum_{y \in f(X)} y \chi_{f^{-1}(y)}, \quad (18.7)$$

wobei für eine Teilmenge $A \subset X$

$$\chi_A = \begin{cases} 1 & x \in A \\ 0 & \text{sonst} \end{cases}$$

die *charakteristische Funktion* von A bezeichnet. Die Darstellung (18.7) heißt auch kanonische Darstellung der Treppenfunktion f .

(b) $T(X, \mathcal{A})$ ist ein Vektorraum über \mathbb{R} , nämlich ein Untervektorraum aller Funktionen $f : X \rightarrow \mathbb{R}$. Denn wegen 18.4.8 ist die Summe und das Vielfache messbarer Funktionen wieder messbar. Außerdem nimmt die Summe zweier Funktionen mit endlich vielen Werten ebenfalls nur endlich viele Werte an.

(c) Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ die charakteristische Funktion $f = \chi_{\mathbb{Q}}$, so ist $f \in T(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Im Unterschied zu den Treppenfunktionen im Sinne von Definition 5.1.1 müssen daher die Mengen, auf denen die Funktion konstant ist, keine Intervalle oder Vereinigungen von Intervallen sein.

(d) Mit

$$M(X, \mathcal{A}) = \{f : (X, \mathcal{A}) \rightarrow (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}})) \mid f \text{ messbar}\}$$

bezeichnen wir die Menge der *messbaren Funktionen* und mit

$$M^+(X, \mathcal{A}) = \{f \in M(X, \mathcal{A}) \mid f \geq 0\}$$

die Menge der *nicht negativen messbaren Funktionen*. Beide Mengen sind keine Vektorräume, wie die Bemerkung nach 18.4.2 zeigt. Im ersten Fall versagt das Assoziativgesetz der Addition, im zweiten Fall ist die Existenz eines inversen Elementes nicht gegeben. Hingegen bilden die messbaren reellwertigen Funktionen einen Vektorraum.

Wir zeigen nun, dass die nicht negativen messbaren Funktionen genau solche sind, die sich als Limes einer monoton wachsenden Folge von nicht negativen Treppenfunktionen darstellen lassen.

Satz 18.4.11. *Sei $f : X \rightarrow [0, \infty]$ eine Funktion. Dann sind äquivalent:*

- (i) $f \in M^+(X, \mathcal{A})$
- (ii) *es existiert eine Folge $(u_n)_{n \in \mathbb{N}} \in T^+(X, \mathcal{A})$ mit $u_1 \leq u_2 \leq u_3 \leq \dots \leq u_n \leq u_{n+1} \dots$ und $\lim_{n \rightarrow \infty} u_n(x) = f$.*

Falls f beschränkt ist, kann $(u_n)_{n \in \mathbb{N}}$ gleichmäßig konvergent gewählt werden.

Beweis. Aus (ii) folgt (i), denn jeder Limes von messbaren Funktionen ist messbar. Sei also $f \in M^+(X, \mathcal{A})$ und $n \in \mathbb{N}$. Dann sind die Mengen

$$A_{j,n} = \left\{ x \in X \mid \frac{j}{2^n} \leq f(x) < \frac{j+1}{2^n} \right\}$$

für $j \in \{0, \dots, n2^n - 1\}$ und

$$A_{n2^n, n} = \{x \in X \mid f(x) \geq n\}$$

paarweise disjunkt und liegen wegen der Messbarkeit von f in \mathcal{A} . Außerdem folgt

$$\bigcup_{j=0}^{n2^n} A_{j,n} = X.$$

Damit ist

$$u_n(x) := \sum_{j=0}^{n2^n} \frac{j}{2^n} \chi_{A_{j,n}} \in T^+(X, \mathcal{A})$$

eine Treppenfunktion. Für jedes $x \in X$ ist $(u_n(x))_{n \in \mathbb{N}}$ eine monoton wachsende Zahlenfolge, denn

$$A_{j,n} = A_{2j, n+1} \cup A_{2j+1, n+1}$$

für $j \in \{0, \dots, n2^n - 1\}$ und

$$A_{n2^n, n} = \bigcup_{j=n2^{n+1}}^{(n+1)2^{n+1}} A_{j, n+1}$$

sind disjunkte Vereinigungen. Falls $x \in A_{j,n}$ und $j \in \{0, \dots, n2^n - 1\}$, gilt daher:

$$u_n(x) = \frac{j}{2^n} = \frac{2j}{2^{n+1}} \leq u_{n+1}(x).$$

Ist $x \in A_{n2^n, n}$, so folgt

$$u_n(x) = n = \frac{n2^{n+1}}{2^{n+1}} \leq u_{n+1}(x).$$

Für jedes $x \in X$ ist $\lim_{n \rightarrow \infty} u_n(x) = f(x)$. Denn ist $f(x) = \infty$, so ist $x \in A_{n2^n, n}$ für alle $n \in \mathbb{N}$ und somit $u_n(x) = n$.

Ist $f(x) < \infty$, so existiert für alle $n > f(x)$ ein

$$j \in \{0, \dots, n2^n - 1\} \text{ mit } x \in A_{j,n}.$$

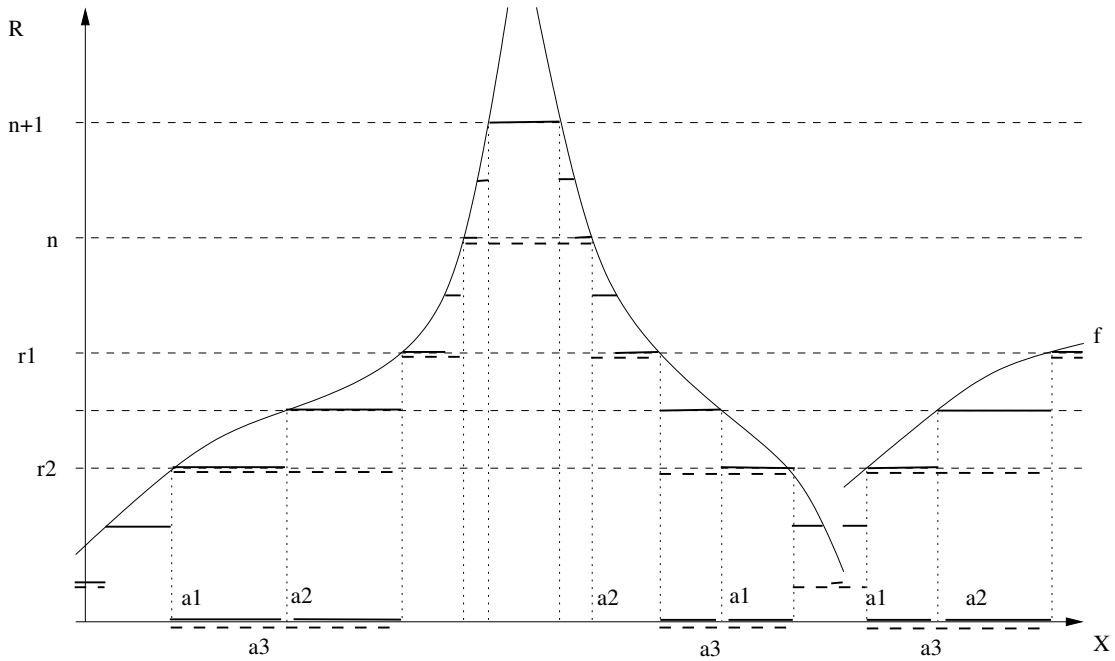


Abbildung 18.10: zum Beweis von Satz 18.4.11; die breiten gestrichelten bzw. durchgezogenen Linien deuten u_n bzw. u_{n+1} an, entsprechendes gilt für $A_{*,n}$ bzw. $A_{*,n+1}$

Damit erhalten wir:

$$0 \leq f(x) - u_n(x) < \frac{j+1}{2^n} - \frac{j}{2^n} = \frac{1}{2^n}$$

und somit gilt:

$$\lim_{n \rightarrow \infty} u_n(x) = f(x)$$

für alle $x \in X$. Ist f beschränkt, so existiert ein $n_0 \in \mathbb{N}$ mit $f(x) \leq n_0$ für alle $x \in X$. Damit gilt für alle $n \geq n_0$ und $x \in X$:

$$0 \leq f(x) - u_n(x) < \frac{j+1}{2^n} - \frac{j}{2^n} = \frac{1}{2^n}.$$

□

Korollar 18.4.12. (a) Zu jeder beschränkten \mathcal{A} -messbaren Funktion $f : X \rightarrow \mathbb{R}$ existiert eine wachsende Folge $(v_n)_{n \in \mathbb{N}} \in T(X, \mathcal{A})$, die gleichmäßig gegen f konvergiert.

(b) Zu jeder nach unten beschränkten messbaren Funktion $f : X \rightarrow \overline{\mathbb{R}}$ existiert eine monoton wachsende Folge $(v_n)_{n \in \mathbb{N}} \in T(X, \mathcal{A})$, die punktweise gegen f konvergiert.

(c) Ist $f : X \rightarrow \overline{\mathbb{R}}$ messbar, so existiert eine Folge $(v_n)_{n \in \mathbb{N}} \in T(X, \mathcal{A})$, die punktweise gegen f konvergiert.

Beweis.

zu (a) Ist f beschränkt, so existieren $\alpha, \beta > 0$ mit

$$0 \leq f(x) + \alpha < \beta.$$

Wie in Satz 18.4.11 konstruieren wir eine wachsende Folge $(u_n)_{n \in \mathbb{N}} \in T^+(X, \mathcal{A})$ mit

$$\lim_{n \rightarrow \infty} u_n(x) = f(x) + \alpha.$$

Da $f(x) + \alpha$ beschränkt ist, konvergiert u_n sogar gleichmäßig, denn für $n \geq \beta$ ist

$$0 \leq f(x) - u_n(x) \leq \frac{1}{2^n}.$$

Setze nun $v_n = u_n - \alpha$.

zu (b) Wähle $\alpha \geq 0$ mit $0 \leq f(x) + \alpha$ und argumentiere wie in (a).

zu (c) Da f messbar ist, sind auch f^+ und f^- messbar und es existieren $u_n \in M^+(X, \mathcal{A})$ und $u'_n \in M^+(X, \mathcal{A})$ mit

$$f^+(x) = \lim_{n \rightarrow \infty} u_n(x) \quad \text{und} \quad f^-(x) = \lim_{n \rightarrow \infty} u'_n(x).$$

Dann folgt:

$$f(x) = f^+(x) - f^-(x) = \lim_{n \rightarrow \infty} (u_n(x) - u'_n(x)).$$

□

18.5 Das Lebesgue-Integral

Definition 18.5.1. Sei μ ein Maß auf (X, \mathcal{A}) , $f \in T^+(X, \mathcal{A})$ eine nicht negative Treppenfunktion mit kanonischer Darstellung

$$f = \sum_{j=1}^n \alpha_j \chi_{A_j} = \sum_{y \in f(X)} y \chi_{f^{-1}(\{y\})}.$$

Dann heißt

$$\int f d\mu = \sum_{j=1}^n \alpha_j \mu(A_j) = \sum_{y \in f(X)} y \mu(f^{-1}(\{y\}))$$

das *Integral* von f bzgl. μ .

Bemerkungen.

- (a) Der Grund, weshalb man das Integral nicht für beliebige Treppenfunktionen definiert ist der Folgende: Die Maße der Mengen A_j müssen nicht endlich sein. Hätten die Koeffizienten α_j unterschiedliche Vorzeichen und hätten wenigstens drei Mengen A_j unendliches Maß, so wäre die Summe $\sum_{j=1}^n \alpha_j \mu(A_j)$ nach den Abschnitt 18.4 vereinbarten Rechenregeln von der Summationsreihenfolge abhängig.

- (b) Sei $f = \sum_{j=1}^m \beta_j \chi_{B_j}$ eine andere Darstellung von f mit disjunkten $(B_j)_{j \leq m} \in \mathcal{A}$, deren Vereinigung die Menge X ergibt. Im Unterschied zur kanonischen Darstellung müssen die Koeffizienten β_j nicht verschieden sein. Sei $\sum_{j=1}^n \alpha_j \chi_{A_j}$ die kanonische Darstellung von f und $x \in B_j \cap A_i \neq \emptyset$, so folgt: $\beta_j = f(x) = \alpha_i$ und somit

$$\sum_{j=1}^m \beta_j \mu(B_j) = \sum_{j=1}^m \sum_{i=1}^n \beta_j \mu(A_i \cap B_j) = \sum_{i=1}^n \sum_{j=1}^m \alpha_i \mu(A_i \cap B_j) = \sum_{i=1}^n \alpha_i \mu(A_i) = \int f d\mu.$$

Diese einfache Beobachtung wird für den Beweis des folgenden Satzes nützlich sein.

Satz 18.5.2. Sei (X, \mathcal{A}, μ) ein Maßraum, so gilt:

(a) $\int_X \chi_A d\mu = \mu(A)$, falls $A \in \mathcal{A}$.

(b) Falls $f, g \in T^+(X, \mathcal{A})$ und $0 \leq \alpha, \beta < \infty$, so gilt

$$\int_X (\alpha f + \beta g) d\mu = \alpha \int_X f d\mu + \beta \int_X g d\mu.$$

(c) Seien $f, g \in T^+(X, \mathcal{A})$ mit $f \leq g$, so gilt:

$$\int_X f d\mu \leq \int_X g d\mu.$$

Beweis.

- zu (a) Die charakteristische Funktion der messbaren Menge A ist eine Treppenfunktion und ihre kanonische Darstellung ist im Falle $X \setminus A \neq \emptyset$ von der Form $\chi_A = 0 \cdot \chi_{X \setminus A} + 1 \cdot \chi_A$ und somit ist

$$\int_X \chi_A d\mu = 0 \cdot \mu(X \setminus A) + 1 \cdot \mu(A) = \mu(A).$$

Ist $A = X$, so ist χ_X die kanonische Darstellung und die Behauptung folgt sofort aus der Definition.

- zu (b) Die Beziehung $\int_X \alpha f d\mu = \alpha \int_X f d\mu$ folgt unmittelbar aus der Definition. Seien nun f, g zwei nicht negative Treppenfunktionen mit den kanonischen Darstellungen

$$f = \sum_{i=1}^n \alpha_i \chi_{A_i} \quad \text{und} \quad g = \sum_{j=1}^m \beta_j \chi_{B_j}.$$

Dann sind die Mengen $A_i \cap B_j$ paarweise disjunkt und ihre Vereinigung ist X . Des Weiteren ist

$$f + g = \sum_{i=1}^n \sum_{j=1}^m (\alpha_i + \beta_j) \chi_{A_i \cap B_j}.$$

Wegen Bemerkung (b) folgt somit

$$\begin{aligned}
 \int_X f + g d\mu &= \sum_{i=1}^n \sum_{j=1}^m (\alpha_i + \beta_j) \mu(A_i \cap B_j) = \sum_{i=1}^n \sum_{j=1}^m \alpha_i \mu(A_i \cap B_j) + \sum_{i=1}^n \sum_{j=1}^m \beta_j \mu(A_i \cap B_j) \\
 &= \sum_{i=1}^n \alpha_i \sum_{j=1}^m \mu(A_i \cap B_j) + \sum_{j=1}^m \beta_j \sum_{i=1}^n \mu(A_i \cap B_j) \\
 &= \sum_{i=1}^n \alpha_i \mu(A_i) + \sum_{j=1}^m \beta_j \mu(B_j) = \int_X f d\mu + \int_X g d\mu.
 \end{aligned}$$

zu (c) Schreibe $g = g - f + f$. Dann ist $g - f \in T^+(X, \mathcal{A})$ und aus (a) folgt

$$\int_X g d\mu = \int_X (g - f) d\mu + \int_X f d\mu \geq \int_X f d\mu.$$

□

Bemerkung. Aus (b) folgt insbesondere: Sind $A_1, \dots, A_n \in \mathcal{A}$ beliebig und $\alpha_1, \dots, \alpha_n \geq 0$, so ist $f = \sum_{j=1}^n \alpha_j \chi_{A_j}$ eine Treppenfunktion und es gilt:

$$\int f d\mu = \sum_{j=1}^n \alpha_j \mu(A_j).$$

Definition 18.5.3 (Integral für $f \in M^+(X, \mathcal{A})$). Sei (X, \mathcal{A}, μ) ein Maßraum und $f \in M^+(X, \mathcal{A})$. Ist $(u_n)_{n \in \mathbb{N}} \in T^+(X, \mathcal{A})$ eine monoton wachsende Folge mit $\lim_{n \rightarrow \infty} u_n = f$, so definiere

$$\int f d\mu = \lim_{n \rightarrow \infty} \int_X u_n d\mu.$$

Bemerkungen.

(a) Da $u_n \leq u_{n+1}$ mit $\int_X u_n d\mu \leq \int_X u_{n+1} d\mu$, existiert der Grenzwert in $\overline{\mathbb{R}}$.

(b) Wir haben in Satz 18.4.11 gezeigt, dass eine solche Folge $(u_n)_{n \in \mathbb{N}}$ existiert. Wir werden nun zeigen, dass die obige Definition nicht von der Wahl der Folge abhängt.

Lemma 18.5.4. Sei $(u_n)_{n \in \mathbb{N}}$ eine wachsende Folge von Funktionen aus $T^+(X, \mathcal{A})$. Ist $v \in T^+(X, \mathcal{A})$ und $v \leq \lim_{n \rightarrow \infty} u_n$, so gilt:

$$\int_X v d\mu \leq \lim_{n \rightarrow \infty} \int_X u_n d\mu.$$

Beweis. Es sei $v = \sum_{j=1}^m \alpha_j \chi_{A_j}$ mit $\alpha_j \geq 0$ und $A_j \in \mathcal{A}$ für $j \in \{1, \dots, m\}$. Für $\gamma > 1$ und $n \in \mathbb{N}$ betrachte die Menge

$$B_n := \{x \in X \mid \gamma u_n(x) \geq v(x)\} \in \mathcal{A}.$$

Da $(u_n)_{n \in \mathbb{N}}$ eine wachsende Folge von Funktionen ist, ist $B_n \subset B_{n+1}$. Falls $v(x) = 0$, so folgt $x \in B_n$ für alle $n \in \mathbb{N}$. Ist $v(x) > 0$, so folgt aus $v(x) \leq \lim_{n \rightarrow \infty} u_n(x)$

$$\lim_{n \rightarrow \infty} \gamma u_n(x) > v(x).$$

Also ist $x \in B_n$ für n genügend groß und somit

$$\bigcup_{n=1}^{\infty} B_n = X.$$

Daher ist $(A_j \cap B_n)_{n \in \mathbb{N}}$ für jedes $j \in \{1, \dots, m\}$ eine monoton steigende Folge von messbaren Mengen und es folgt aus dem Hauptsatz über Maße:

$$\lim_{n \rightarrow \infty} \mu(A_j \cap B_n) = \mu(A_j \cap \bigcup_{n=1}^{\infty} B_n) = \mu(A_j).$$

Außerdem folgt aus der Definition von B_n , dass $\gamma \cdot u_n(x) \geq v(x) \chi_{B_n}(x)$. Wir erhalten

$$\begin{aligned} \int_X v \, d\mu &= \sum_{j=1}^m \alpha_j \mu(A_j) = \lim_{n \rightarrow \infty} \sum_{j=1}^m \alpha_j \mu(A_j \cap B_n) = \lim_{n \rightarrow \infty} \sum_{j=1}^m \alpha_j \int_X \chi_{A_j \cap B_n} \, d\mu \\ &= \lim_{n \rightarrow \infty} \int_X \sum_{j=1}^m \alpha_j \chi_{A_j} \cdot \chi_{B_n} \, d\mu = \lim_{n \rightarrow \infty} \int_X v \cdot \chi_{B_n} \, d\mu \leq \lim_{n \rightarrow \infty} \gamma \int_X u_n \, d\mu. \end{aligned}$$

Da $\gamma > 1$ beliebig ist, folgt die Behauptung. \square

Korollar 18.5.5. Sind $(\mu_n)_{n \in \mathbb{N}}$ und $(v_n)_{n \in \mathbb{N}}$ zwei wachsende Folgen von Funktionen aus $T^+(X, \mathcal{A})$ mit $\lim_{n \rightarrow \infty} \mu_n = \lim_{n \rightarrow \infty} v_n$, so gilt

$$\lim_{n \rightarrow \infty} \int_X u_n \, d\mu = \lim_{n \rightarrow \infty} \int_X v_n \, d\mu.$$

Insbesondere ist das Integral in 18.5.3 nicht von der Wahl der Folge abhängig, und somit ist das Integral wohldefiniert.

Beweis. Da $(v_n)_{n \in \mathbb{N}}$ eine monoton wachsende Folge ist, gilt:

$$v_k(x) \leq \lim_{n \rightarrow \infty} v_n(x) = \lim_{n \rightarrow \infty} u_n(x)$$

für alle $x \in X$. Deshalb folgt aus Lemma 6.4

$$\int_X v_k(x) \, d\mu \leq \lim_{n \rightarrow \infty} \int_X u_n(x) \, d\mu.$$

und somit

$$\lim_{k \rightarrow \infty} \int_X v_k(x) \, d\mu \leq \lim_{n \rightarrow \infty} \int_X u_n(x) \, d\mu.$$

Vertauschen wir die Rollen von $(v_n)_{n \in \mathbb{N}}$ und $(u_n)_{n \in \mathbb{N}}$, so erhalten wir die umgekehrte Ungleichung. \square

Satz 18.5.6. (a) Es seien $f, g \in M^+(X, \mathcal{A})$ und $\alpha, \beta \in [0, \infty]$. Dann gilt:

$$\int_X (\alpha f + \beta g) d\mu = \alpha \int_X f d\mu + \beta \int_X g d\mu.$$

(b) Ist $f \leq g$, so folgt:

$$\int_X f d\mu \leq \int_X g d\mu.$$

Beweis.

zu (a) Wir beweisen zunächst:

$$\int_X (f + g) d\mu = \int_X f d\mu + \int_X g d\mu$$

für alle $f, g \in M^+(X, \mathcal{A})$. Wähle monoton wachsende Folgen $(u_n)_{n \in \mathbb{N}}$ und $(v_n)_{n \in \mathbb{N}}$ nicht negativer Treppenfunktionen mit $\lim_{n \rightarrow \infty} u_n(x) = f(x)$ sowie $\lim_{n \rightarrow \infty} v_n(x) = g(x)$ punktweise. Dann ist ihre Summe $(u_n + v_n)_{n \in \mathbb{N}}$ ebenfalls monoton wachsend und konvergiert punktweise gegen $f + g$. Nach Definition des Integrals 18.5.3 folgt mit Hilfe des Satzes 18.5.2

$$\int_X (f + g) d\mu = \lim_{n \rightarrow \infty} \int_X (u_n + v_n) d\mu = \lim_{n \rightarrow \infty} \int_X u_n d\mu + \lim_{n \rightarrow \infty} \int_X v_n d\mu = \int_X f d\mu + \int_X g d\mu.$$

Zu zeigen bleibt:

$$\int_X \alpha f d\mu = \alpha \int_X f d\mu \tag{18.8}$$

für alle $f \in M^+(X, \mathcal{A})$ und $\alpha \in [0, \infty]$. Ist $\alpha < \infty$, so folgt der Beweis wie oben. Ist $\alpha = \infty$, so betrachte die Menge $A := \{x \in X \mid f(x) > 0\}$. Diese Menge ist messbar und die Treppenfunktionen $u_n = n\chi_A$ konvergieren punktweise gegen $\infty \cdot f$. Also gilt

$$\int_X \infty \cdot f d\mu = \lim_{n \rightarrow \infty} \int_X u_n d\mu = \lim_{n \rightarrow \infty} n\mu(A) = \begin{cases} \infty & \text{falls } \mu(A) > 0, \\ 0 & \text{falls } \mu(A) = 0. \end{cases}$$

Betrachte für $n \in \mathbb{N}$ die monoton wachsende Folge

$$A_n := \left\{x \in X \mid f(x) \geq \frac{1}{n}\right\}$$

messbarer Mengen. Wegen des Hauptsatzes über Maße folgt:

$$\lim_{n \rightarrow \infty} \mu(A_n) = \mu(A).$$

Aus der Definition der Mengen A_n ergibt sich: $\frac{1}{n}\chi_{A_n} \leq f$. Sei nun $(u_k)_{k \in \mathbb{N}}$ eine monoton wachsende punktweise konvergente Folge von nicht negativen Treppenfunktionen mit $\lim_{k \rightarrow \infty} u_k = f$. Wegen Lemma 18.5.4 gilt für jedes $n \in \mathbb{N}$

$$\frac{1}{n}\mu(A_n) = \int_X \frac{1}{n}\chi_{A_n} d\mu \leq \lim_{k \rightarrow \infty} \int_X u_k d\mu = \int_X f d\mu.$$

Ist nun $\mu(A) > 0$, so existiert ein $n \in \mathbb{N}$ mit $\mu(A_n) > 0$ und somit ist $\int_X f d\mu > 0$. Wegen $\infty \int_X f d\mu = \infty$ ist somit (18.8) im Falle $\mu(A) > 0$ bewiesen. Sei nun $\mu(A) = 0$ und u eine nicht negative Treppenfunktion mit $u \leq f$, so folgt:

$$u \leq \max\{u(x) \mid x \in X\} \chi_A \quad \text{und somit} \quad \int_X u d\mu \leq \max\{u(x) \mid x \in X\} \mu(A) = 0.$$

Damit ergibt sich aus der Definition des Integrals: $\int_X f d\mu = 0$. Wegen der Regel $\infty \cdot 0 = 0$ erhält man somit (18.8) auch im Falle $\mu(A) = 0$.

zu (b) Der Beweis von (b) folgt durch Anwendung von (a) auf die Summen $g = (g - f) + f$. □

Satz 18.5.7. Für alle $f \in M^+(X, \mathcal{A})$ gilt:

$$\int_X f d\mu = \sup \left\{ \int_X u d\mu \mid u \in T^+(X, \mathcal{A}), u \leq f \right\} =: s.$$

Beweis. Für alle $u \in T^+(X, \mathcal{A})$ mit $u \leq f$ gilt: $\int_X u d\mu \leq \int_X f d\mu$, d.h. $\int_X f d\mu \geq s$. Die umgekehrte Ungleichung folgt aus der Definition des Integrals, denn

$$\int_X f d\mu = \lim_{n \rightarrow \infty} \int_X u_n d\mu \leq s$$

mit $u_n \leq f$. □

Satz 18.5.8. Sei $f \in M^+(X, \mathcal{A})$. Dann ist $\int_X f d\mu = 0$ genau dann, wenn

$$\{x \in X \mid f(x) > 0\} =: A$$

eine μ -Nullmenge ist, d.h. $\mu(A) = 0$.

Beweis. Sei $A_n := \{x \in X \mid f(x) \geq \frac{1}{n}\}$, so ist $(A_n)_{n \in \mathbb{N}}$ monoton wachsend und $\bigcup_{n=1}^{\infty} A_n = A$. Somit ist wegen des Hauptsatzes 18.1.11 über Maße

$$\mu(A) = \lim_{n \rightarrow \infty} \mu(A_n).$$

Es sei $\int_X f d\mu = 0$. Aus $0 \leq \frac{1}{n} \chi_{A_n} \leq f$ folgt:

$$0 \leq \frac{1}{n} \mu(A_n) = \int_X \frac{1}{n} \chi_{A_n} d\mu \leq \int_X f d\mu = 0.$$

Insbesondere ist $\mu(A) = 0$. Sei nun $\mu(A) = 0$, so folgt aus $f \leq \infty \cdot \chi_A$ nach Satz 18.5.6

$$\int_X f d\mu \leq \infty \cdot \int_X \chi_A d\mu = \infty \cdot 0 = 0,$$

also $\int_X f d\mu = 0$. □

Bemerkung. Sei (X, \mathcal{A}, μ) ein Maßraum und $B(x)$ eine Aussage, die von $x \in X$ abhängt (Aussageform siehe Definition 1.1.2). Man sagt, die Aussage gilt μ -fast überall (*fast sicher*), falls eine μ -Nullmenge $N \subset X$ existiert, so dass die Aussage $B(x)$ wahr ist für alle $x \in X \setminus N = N^c$. Ist also $f \in M^+(X, \mathcal{A})$ und $\int f d\mu = 0$, so ist wegen Satz 18.5.8 $f(x) = 0$ μ -fast überall richtig.

Satz 18.5.9 (Satz von der monotonen Konvergenz / Satz von Beppo Levi). Sei $(f_n)_{n \in \mathbb{N}}$ eine monoton wachsende Folge von Funktionen aus $M^+(X, \mathcal{A})$. Dann gilt:

$$\int_X \left(\lim_{n \rightarrow \infty} f_n \right) d\mu = \lim_{n \rightarrow \infty} \int_X f_n d\mu.$$

Beweis. Es sei $f := \lim_{n \rightarrow \infty} f_n$. Da $(f_n)_{n \in \mathbb{N}}$ eine monoton wachsende Folge ist, gilt für jedes $k \in \mathbb{N} : f_k \leq f_{k+1} \leq f$ und somit $\int_X f_k d\mu \leq \int_X f_{k+1} d\mu \leq \int_X f d\mu$. Insbesondere ist aufgrund von Satz 18.5.6(b)

$$\lim_{k \rightarrow \infty} \int_X f_k d\mu \leq \int_X f d\mu.$$

Für den Beweis der umgekehrten Ungleichung betrachte für $u \in T^+(X, \mathcal{A})$ mit $u \leq f$ und $\gamma > 1$ die Mengen $B_n = \{x \in X \mid \gamma f_n(x) \geq u(x)\}$. Dann ist $B_n \subset B_{n+1}$ und $\bigcup_{n=1}^{\infty} B_n = X$. Insbesondere ist $\gamma f_n \geq u \cdot \chi_{B_n}$. Da $(u \cdot \chi_{B_n})_{n \in \mathbb{N}} \in T^+(X, \mathcal{A})$ eine monoton wachsende Folge mit $\lim_{n \rightarrow \infty} u \cdot \chi_{B_n}(x) = u(x)$ ist, folgt aus Lemma 18.5.4

$$\int_X u d\mu \leq \lim_{n \rightarrow \infty} \int_X u \chi_{B_n} d\mu \leq \gamma \lim_{n \rightarrow \infty} \int_X f_n d\mu.$$

Da $\gamma > 1$ beliebig ist, erhalten wir $\int_X u d\mu \leq \lim_{n \rightarrow \infty} \int_X f_n d\mu$ und aus Satz 18.5.7 die umgekehrte Ungleichung. □

Bemerkung. Ohne Monotonieannahme ist der Satz falsch. Betrachte $(\mathbb{R}, \mathcal{B}, \lambda)$. Dann ist $f_n = \frac{1}{n} \chi_{[0,n]} \in M^+(\mathbb{R}, \mathcal{B})$ und $\int_{\mathbb{R}} f_n d\mu = 1$. Jedoch konvergiert $(f_n)_{n \in \mathbb{N}}$ auf \mathbb{R} gleichmäßig gegen 0 und somit ist

$$\int_{\mathbb{R}} \lim_{n \rightarrow \infty} f_n d\mu = 0.$$

Nun definieren wir das Lebesgue-Integral einer Funktion f , indem wir sie in ihren Positivteil f^+ und Negativteil f^- zerlegen.

Definition 18.5.10. Sei (X, \mathcal{A}, μ) ein Maßraum. Eine Funktion $f : X \rightarrow \overline{\mathbb{R}}$ heißt μ -integrierbar, falls sie messbar ist und die Integrale von $f^\pm \in M^+(X, \mathcal{A})$ endlich sind.

Die reelle Zahl

$$\int_X f d\mu := \int_X f^+ d\mu - \int_X f^- d\mu$$

heißt das *Lebesgue-Integral* von f (über X bezüglich μ).

Bemerkung. Ist wenigstens eines der Integrale $\int_X f^+ d\mu, \int_X f^- d\mu$ endlich, so nennen wir f *quasiintegrierbar*.

Satz 18.5.11. Für jede Funktion $f : X \rightarrow \overline{\mathbb{R}}$ sind folgende Aussagen äquivalent:

- (a) f ist integrierbar.
- (b) Es existieren integrierbare Funktionen $u, v \in M^+(X, \mathcal{A})$ mit $f = u - v$.
- (c) f ist messbar und es existiert eine integrierbare Funktion g mit $|f| \leq g$.
- (d) f ist messbar und $|f|$ integrierbar.

Beweis.

- (a) $a \Rightarrow b$

Ist f integrierbar, so auch f^+, f^- .

- (b) \Rightarrow (c)

f ist messbar, denn u und v sind messbar. Außerdem ist

$$g(x) := u(x) + v(x) \geq |u(x) - v(x)| = |f(x)|$$

integrierbar. Da $u, v \in M^+(X, \mathcal{A})$ und integrierbar sind, gilt

$$\int_X (u + v) d\mu = \int_X u d\mu + \int_X v d\mu < \infty.$$

- (c) \Rightarrow (d)

Da f messbar ist, sind f^+, f^- messbar und somit auch $|f| = f^+ + f^- \in M^+(X, \mathcal{A})$.

Da $\int_X |f| d\mu \leq \int_X g d\mu < \infty$ ist $|f|$ integrierbar.

- (d) \Rightarrow (a)

Da f messbar ist, sind auch f^+, f^- messbar und da $\int_X f^\pm d\mu \leq \int_X |f| d\mu < \infty$, ist f integrierbar.

□

Nun wollen wir einige Eigenschaften integrierbarer Funktionen diskutieren.

Korollar 18.5.12. Ist $f : X \rightarrow \overline{\mathbb{R}}$ integrierbar, so ist

$$A = \{x \in X \mid |f(x)| = \infty\}$$

eine μ -Nullmenge, d.h. $f(x) \in \mathbb{R}$ fast überall.

Beweis. Da f integrierbar ist, ist wegen Satz 18.5.11 auch $|f|$ integrierbar und insbesondere messbar. Daher ist die Menge A ebenfalls messbar. Da $\infty \cdot \chi_A \leq |f|$, folgt aus Satz 18.5.6

$$\infty \cdot \mu(A) = \infty \cdot \int_X \chi_A d\mu = \int_X \infty \chi_A d\mu \leq \int_X |f| d\mu < \infty,$$

und somit ist $\mu(A) = 0$.

□

Korollar 18.5.13. Sei $f : X \rightarrow \overline{\mathbb{R}}$ messbar und μ -fast überall null, d.h.

$$\mu\{x \in X \mid f(x) \neq 0\} = 0.$$

Dann ist f integrierbar und $\int_X f d\mu = 0$.

Beweis. Sei $A = \{x \in X \mid f(x) \neq 0\}$, so ist $|f| \leq \infty \cdot \chi_A$. Da wegen Satz 18.5.11 $\int_X \infty \chi_A d\mu = \infty \mu(A) = 0$ gilt, ist $\infty \cdot \chi_A$ integrierbar. Die Integrierbarkeit von f folgt somit aus Satz 18.5.11(c). \square

Korollar 18.5.14. Sind $f, g : X \rightarrow \overline{\mathbb{R}}$ integrierbar, so auch $\max(f, g)$ und $\min(f, g)$.

Beweis. $\max(f, g)$ und $\min(f, g)$ sind messbar und $|\max(f, g)| \leq |f| + |g|$, sowie $|\min(f, g)| \leq |f| + |g|$. Wegen Satz 18.5.11(c) sind $|f|, |g|$ und daher auch $\max(f, g)$ sowie $\min(f, g)$ integrierbar. \square

Satz 18.5.15 (Linearität des Integrals). Seien $f, g : X \rightarrow \overline{\mathbb{R}}$ integrierbar und $\alpha, \beta \in \mathbb{R}$. Dann ist auch $\alpha f + \beta g$ integrierbar und es gilt:

$$\int_X (\alpha f + \beta g) d\mu = \alpha \int_X f d\mu + \beta \int_X g d\mu.$$

Beweis. Da $|\alpha f + \beta g| \leq |\alpha| |f| + |\beta| |g|$, folgt die Integrierbarkeit von $\alpha f + \beta g$ mit Hilfe von Satz 18.5.11 aus der Integrierbarkeit von f und g . Der Rest folgt durch Zerlegung der Funktionen f und g in ihren positiven und negativen Anteil und aus der Definition des Integrals. Die Details seien als Übung überlassen. \square

Korollar 18.5.16. Seien $f, g : X \rightarrow \overline{\mathbb{R}}$ messbar und $f(x) = g(x)$ fast überall. Ist f integrierbar, so auch g und $\int_X g d\mu = \int_X f d\mu$.

Beweis. Da $(g - f)(x)$ fast überall gleich null ist, ist wegen Korollar 18.5.13 die Funktion $g - f$ integrierbar mit $\int_X (g - f) d\mu = 0$. Aus $g = g - f + f$ folgt mit Satz 18.5.15 die Integrierbarkeit von g und wir erhalten: $\int_X g d\mu = \int_X f d\mu$. \square

Bemerkung. Ändern wir also eine Funktion auf einer Menge von Maß null durch eine messbare Funktion ab, so ändert sich nichts an ihrer Integrierbarkeit. Genauer: Sei f eine integrierbare Funktion und A eine Menge von Maß null. Dann ist für jede messbare Funktion g auch die Funktion $\tilde{f} = f \chi_{X \setminus A} + g \chi_A$ integrierbar und die Integrale von f und \tilde{f} stimmen überein.

Satz 18.5.17 (Monotonie des Integrals). Sind $f, g : X \rightarrow \overline{\mathbb{R}}$ quasiintegrierbar und gilt $f \leq g$ μ -fast überall, so gilt:

$$\int_X f d\mu \leq \int_X g d\mu.$$

Beweis. Sei zunächst $f \leq g$ überall. Aus der Voraussetzung $f \leq g$ ergibt sich $f^+ \leq g^+$ sowie $f^- \geq g^-$. Dann folgt aus 18.5.6:

$$\int_X f d\mu = \int_X f^+ d\mu - \int_X f^- d\mu \leq \int_X g^+ d\mu - \int_X g^- d\mu = \int_X g d\mu.$$

Anderenfalls betrachte

$$\tilde{g}(x) := \begin{cases} g(x) & f(x) \leq g(x), \\ \infty & f(x) > g(x). \end{cases}$$

mit gleichem Integral wie g . □

Korollar 18.5.18. Es sei $f : X \rightarrow \overline{\mathbb{R}}$ messbar und $g : X \rightarrow \overline{\mathbb{R}}$ integrierbar mit

$$|f(x)| \leq g(x)$$

fast überall. Dann ist auch f integrierbar.

Beweis. Da $|f| \in M^+(X, \mathcal{A})$, ist $|f|$ quasiintegrierbar und wegen Satz 18.5.17 gilt $\int_X |f| d\mu \leq \int_X g d\mu < \infty$. Daher ist $|f|$ integrierbar und wegen Satz 18.5.11 auch f . □

18.6 Konvergenzsätze

Bisher haben wir als Konvergenzsatz den Satz von B. Levi kennengelernt, der besagt: Ist $(f_n)_{n \in \mathbb{N}} \in M^+(X, \mathcal{A})$ monoton wachsend mit $f = \lim_{n \rightarrow \infty} f_n$, so gilt:

$$\lim_{n \rightarrow \infty} \int_X f_n d\mu = \int_X f d\mu,$$

falls μ ein Maß auf \mathcal{A} definiert. Aus diesem Satz lässt sich leicht das Lemma von Fatou beweisen.

Lemma 18.6.1 (Lemma von Fatou). Sei $f_n \in M^+(X, \mathcal{A})$ und μ ein Maß auf \mathcal{A} . Dann gilt:

$$\int_X \liminf_{n \rightarrow \infty} f_n d\mu \leq \liminf_{n \rightarrow \infty} \int_X f_n d\mu.$$

Bemerkung. Der wesentliche Unterschied zum Satz von Beppo Levi ist, dass $(f_n)_{n \in \mathbb{N}}$ nicht monoton sein muss. Daher kann man Gleichheit nicht erwarten, wie das schon aufgeführte Beispiel $f_n = \frac{1}{n} \chi_{[0, n]}$ zeigt.

Beweis. Sei $f(x) = \liminf_{n \rightarrow \infty} f_n(x) = \lim_{n \rightarrow \infty} g_n(x)$ mit $g_n(x) = \inf\{f_k(x) \mid k \geq n\}$. Dann ist die Folge $(g_n)_{n \in \mathbb{N}}$ monoton wachsend und wegen des Satzes von B. Levi folgt: $\int_X f d\mu = \lim_{n \rightarrow \infty} \int_X g_n d\mu$. Da $\int_X g_n d\mu \leq \int_X f_k d\mu$ für alle $k \geq n$, gilt $\int_X g_n d\mu \leq \inf\{\int_X f_k d\mu \mid k \geq n\}$ und somit

$$\lim_{n \rightarrow \infty} \int_X g_n d\mu \leq \lim_{n \rightarrow \infty} \inf\{\int_X f_k d\mu \mid k \geq n\} = \lim_{k \rightarrow \infty} \int_X f_k d\mu.$$

□

Satz 18.6.2 (Satz von der majorisierten Konvergenz / Satz von Lebesgue). Sei (X, \mathcal{A}, μ) ein Maßraum und für die messbaren Funktionen $f, f_n : X \rightarrow \overline{\mathbb{R}}$ gelte:

1. $\lim_{n \rightarrow \infty} f_n(x) = f(x)$ für μ -fast alle $x \in X$,
2. es existiert eine integrierbare Majorante g der Folge $(f_n)_{n \in \mathbb{N}}$, d.h. $|f_n(x)| \leq g(x)$ für fast alle $x \in X$.

Dann sind alle f_n sowie f integrierbar und es gilt:

$$\lim_{n \rightarrow \infty} \int_X |f_n - f| d\mu = 0.$$

Bemerkung.

- (a) Da $|\int_X f_n d\mu - \int_X f d\mu| = |\int_X (f_n - f) d\mu| \leq \int_X |f_n - f| d\mu$, folgt auch $\int_X f_n \rightarrow \int_X f d\mu$.
- (b) Die Existenz einer integrierbaren Majorante ist notwendig. Wie in der Bemerkung nach dem Satz von Beppo Levi betrachte wiederum das Beispiel $f_n = \frac{1}{n} \chi_{[0,n]}$. Wie schon dort bemerkt wurde, lassen sich Limesbildung und Integration nicht vertauschen. Dabei sind alle f_n integrierbar und sogar beschränkt, aber es existiert keine integrierbare Majorante. Insbesondere ist $\int \sup_{n \in \mathbb{N}} f_n = \infty$.

Beweis. Da f_n messbar ist und $|f_n| \leq g$, sind die $|f_n|$ und f_n integrierbar. Da für μ -fast alle $x \in X$ $\lim_{n \rightarrow \infty} f_n(x) = f(x)$, ist auch $|f| \leq g$ μ -fast überall. Daher sind auch $|f|$ und f integrierbar. Durch Abändern von f und f_n auf einer Menge von Maß null können wir annehmen, dass $\lim_{n \rightarrow \infty} f_n(x) = f(x)$ für alle $x \in X$ gilt (z.B. setze f, f_n gleich null auf der Menge der $x \in X$, für die keine Konvergenz der Folge $(f_n(x))_{n \in \mathbb{N}}$ gegen $f(x)$ vorliegt). Dann ist

$$g_n := |f| + g - |f_n - f| \geq |f| + g - |f_n| - |f| = g - |f_n| \geq 0,$$

und es folgt $g_n \in M^+(X, \mathcal{A})$ und $\lim_{n \rightarrow \infty} g_n = |f| + g$. Mit dem Lemma von Fatou erhalten wir:

$$\begin{aligned} \int_X (|f| + g) d\mu &= \int_X \liminf_{n \rightarrow \infty} g_n d\mu \leq \liminf_{n \rightarrow \infty} \int_X g_n d\mu = \liminf_{n \rightarrow \infty} \left(\int_X (|f| + g) d\mu - \int_X |f_n - f| d\mu \right) \\ &= \int_X (|f| + g) d\mu - \overline{\lim}_{n \rightarrow \infty} \int_X |f_n - f| d\mu. \end{aligned}$$

und es folgt

$$\lim_{n \rightarrow \infty} \int_X |f_n - f| d\mu = 0.$$

□

Es sei $X = [a, b]$ und λ das Lebesgue-Maß auf der Borel σ -Algebra von $[a, b] \subset \mathbb{R}$. Wir hatten in Definition 5.1.9 das Integral für Regelfunktionen $f : [a, b] \rightarrow \mathbb{R}$ eingeführt. Für solche Funktionen stimmt dieses Integral mit dem Integral bezüglich des Lebesgue-Maßes λ überein, d.h. es ist $\int_a^b f(x) dx = \int_{[a,b]} f(x) d\lambda(x)$ (siehe Übungsaufgabe 2, Bl. 7, MfP IV).

Als erste Anwendung des Satzes von der majorisierten Konvergenz beweisen wir folgende Aussage.

Satz 18.6.3. *Es sei $f : [a, b] \rightarrow \mathbb{R}$ differenzierbar und die Ableitung f' sei beschränkt. Dann ist f' integrierbar und es gilt:*

$$\int_a^b f'(x) d\lambda(x) = f(b) - f(a).$$

Bemerkung. Diesen Satz haben wir unter der Voraussetzung, dass f' stetig ist, schon einmal bewiesen. Allerdings muss die Ableitung einer differenzierbaren Funktion nicht stetig sein. Man kann Beispiele von differenzierbaren Funktionen angeben, deren Ableitungen nicht einmal Regelfunktionen sind.

Beweis. Wir können f auf $[a - 1, b + 1]$ differenzierbar fortsetzen, so dass f' auch auf $[a - 1, b + 1]$ beschränkt ist, d.h. es existiert ein $M > 0$ mit $|f'(x)| \leq M$ für alle $x \in [a - 1, b + 1]$. Betrachte nun die Folge $(g_n)_{n \in \mathbb{N}} : [a, b] \rightarrow \mathbb{R}$ mit

$$g_n(x) = n \left(f\left(x + \frac{1}{n}\right) - f(x) \right).$$

Da f stetig ist, ist g_n stetig und somit Borel-messbar. Dann ist auch

$$f'(x) = \lim_{n \rightarrow \infty} g_n(x)$$

Borel-messbar. Aus dem Mittelwertsatz der Differentialrechnung folgt $|g_n(x)| \leq M$, denn $g_n(x) = f'(x_n)$ für ein $x_n \in (x, x + \frac{1}{n})$. Damit haben die $\{g_n\}_{n \in \mathbb{N}}$ eine integrierbare Majorante. Mit dem Satz von der majorisierten Konvergenz erhalten wir

$$\int_a^b f'(x) d\lambda(x) = \lim_{n \rightarrow \infty} \int_a^b g_n(x) d\lambda(x).$$

Aus dem Hauptsatz der Differential- und Integralrechnung folgt: Ist $F(x) = \int_a^x f(t) d\lambda(t)$ und f stetig, so ist $F'(x) = f(x)$. Dies impliziert

$$\int_a^b g_n(x) d\lambda(x) = n \int_a^b \left(f\left(x + \frac{1}{n}\right) - f(x) \right) d\lambda(x) = n \left(F\left(b + \frac{1}{n}\right) - F(b) - F\left(a + \frac{1}{n}\right) + F(a) \right)$$

und somit erhalten wir

$$\lim_{n \rightarrow \infty} \int_a^b g_n(x) d\lambda(x) = f(b) - f(a).$$

□

Es lässt sich aus dem Satz von der majorisierten Konvergenz eine Version für Reihen ableiten:

Satz 18.6.4. *Sei (X, \mathcal{A}, μ) ein Maßraum und für die messbaren Funktionen $f, f_n : X \rightarrow \overline{\mathbb{R}}$ gelte:*

- $f(x) = \sum_{k=1}^{\infty} f_k(x)$ für μ -fast alle $x \in X$,

2. Es existiere eine integrierbare Funktion $g \in M^+(X, \mathcal{A})$ mit $|\sum_{k=1}^n f_k| \leq g$ μ -fast überall für alle $n \in \mathbb{N}$.

Dann sind f und alle f_k mit $k \in \mathbb{N}$ integrierbar und es gilt:

$$\sum_{k=1}^{\infty} \int_X f_k d\mu = \int_X f d\mu.$$

Beweis. Betrachte die Partialsummenfolge $g_n = \sum_{k=1}^n f_k$. Nach Voraussetzung besitzt die Folge $(g_n)_{n \in \mathbb{N}}$ eine integrierbare Majorante. Ferner gilt $\lim_{n \rightarrow \infty} g_n(x) = \sum_{k=1}^{\infty} f_k(x) = f(x)$ für μ -fast alle $x \in X$. Also erfüllen f und g_n alle Voraussetzungen des Lebesgueschen Satzes von der majorisierten Konvergenz und wir erhalten:

$$\int_X f d\mu \stackrel{\text{Lebesgue}}{=} \lim_{n \rightarrow \infty} \int_X g_n d\mu = \lim_{n \rightarrow \infty} \int_X \sum_{k=1}^n f_k d\mu = \sum_{k=1}^{\infty} \int_X f_k d\mu.$$

□

Weitere wichtige Anwendungen des Satzes von der majorisierten Konvergenz sind:

Satz 18.6.5 (Stetige Abhängigkeit des Integrals von einem Parameter). *Es sei M ein metrischer Raum, (X, \mathcal{A}, μ) ein Maßraum und $f : M \times X \rightarrow \mathbb{R}$ habe folgende Eigenschaften:*

- (a) Für alle $p \in M$ sei $f(p, \cdot) : X \rightarrow \mathbb{R}$ integrierbar.
- (b) Für μ -fast alle $x \in X$ sei $f(\cdot, x) : M \rightarrow \mathbb{R}$ stetig in $p_0 \in M$.
- (c) Es existiert eine Umgebung $U \subset M$ von $p_0 \in M$ und eine integrierbare Funktion $g \in M^+(X, \mathcal{A})$, so dass für alle $p \in U$ gilt $|f(p, \cdot)| \leq g$ μ -fast überall.

Dann ist die Funktion $F : M \rightarrow \mathbb{R}$ mit

$$F(p) = \int_X f(p, x) d\mu(x)$$

stetig in p_0 .

Beweis. Betrachte eine Folge $(p_n)_{n \in \mathbb{N}}$ in U mit $\lim_{n \rightarrow \infty} p_n = p_0$. Setze $f_n(x) := f(p_n, x)$ für $n \in \mathbb{N}$. Nach Voraussetzung existiert eine integrierbare Funktion $g \in M^+(X, \mathcal{A})$ mit $|f_n(x)| \leq g(x)$ für μ -fast alle $x \in X$. Ferner ist $\lim_{n \rightarrow \infty} f_n(x) = f(p_0, x)$ für μ -fast alle $x \in X$. Somit lässt sich der Lebesguesche Satz 18.6.2 von der majorisierten Konvergenz auf $(f_n)_{n \in \mathbb{N}}$ anwenden. Wir erhalten:

$$\lim_{n \rightarrow \infty} F(p_n) = \lim_{n \rightarrow \infty} \int_X f_n(x) d\mu = \int_X \lim_{n \rightarrow \infty} f_n(x) d\mu = \int_X f(p_0, x) d\mu = F(p_0).$$

und somit die Stetigkeit von F in p_0 .

□

Eine weitere Anwendung ist die sogenannte

Satz 18.6.6 (*Differentiation unter dem Integralzeichen*). *Es sei $I \subset \mathbb{R}$ ein Intervall und $f : I \times X \rightarrow \mathbb{R}$ eine Funktion mit folgenden Eigenschaften.*

- (a) *Für alle $t \in I$ sei $f(t, \cdot) : X \rightarrow \mathbb{R}$ integrierbar.*
- (b) *Für jedes $x \in X$ sei die Funktion $t \mapsto f(t, x)$ auf I differenzierbar.*
- (c) *Es existiere eine integrierbare Funktion $g \in M^+(X, \mathcal{A})$, so dass für jedes $t \in I$ und für μ -fast alle $x \in X$ gilt:*

$$\left| \frac{\partial f}{\partial t}(t, x) \right| \leq g(x).$$

Dann ist die Funktion $F : I \rightarrow \mathbb{R}$ mit $F(t) := \int_X f(t, x) d\mu$ auf I differenzierbar und es gilt für jedes $t_0 \in I$:

$$F'(t_0) = \int_X \frac{\partial f}{\partial t}(t_0, x) d\mu.$$

Beweis. Zu gegebenem $t_0 \in I$ sei $(t_n)_{n \in \mathbb{N}}$ eine Folge in I mit $t_n \neq t_0$ für alle $n \in \mathbb{N}$ und $\lim_{n \rightarrow \infty} t_n = t_0$. Dann gilt:

$$\frac{F(t_n) - F(t_0)}{t_n - t_0} = \int_X \frac{f(t_n, x) - f(t_0, x)}{t_n - t_0} d\mu(x).$$

Wegen des Mittelwertsatzes der Differentialrechnung existiert ein $s_n \in (t_0, t_n)$ (bzw. $s_n \in (t_n, t_0)$, falls $t_0 > t_n$) mit

$$f(t_n, x) - f(t_0, x) = (t_n - t_0) \cdot \frac{\partial f}{\partial t}(s_n, x).$$

Damit ist

$$f_n(x) := \frac{f(t_n, x) - f(t_0, x)}{t_n - t_0} = \frac{\partial f}{\partial t}(s_n, x).$$

Aus (c) folgt daher die Existenz einer integrierbaren Funktion $g \in M^+(X, \mathcal{A})$ mit $|f_n(x)| \leq g(x)$ für alle $n \in \mathbb{N}$ und μ -fast alle $x \in X$. Aus der Differenzierbarkeit der Funktion $t \mapsto f(t, x)$ ergibt sich $\lim_{n \rightarrow \infty} f_n(x) = \frac{\partial f}{\partial t}(t_0, x)$ für alle $x \in X$. Somit erhalten wir aus dem Lebesgueschen Satz 18.6.2 von der majorisierten Konvergenz:

$$\begin{aligned} \int_X \frac{\partial f}{\partial t}(t_0, x) d\mu(x) &= \int_X \lim_{n \rightarrow \infty} f_n(x) d\mu(x) \stackrel{\text{Lebesgue}}{=} \lim_{n \rightarrow \infty} \int_X f_n(x) d\mu(x) \\ &= \lim_{n \rightarrow \infty} \frac{F(t_n) - F(t_0)}{t_n - t_0} = F'(t_0), \end{aligned}$$

woraus sich zugleich die Differenzierbarkeit von F auf I ergibt. □

18.7 Produkt-Maß und der Satz von Fubini

Definition 18.7.1. Es seien (X, \mathcal{A}, μ) und (Y, \mathcal{B}, ν) Maßräume. So heißt

$$\mu \otimes \nu : P(X \times Y) \rightarrow [0, \infty]$$

mit

$$\mu \otimes \nu(S) = \inf \left\{ \sum_{i=1}^{\infty} \mu(A_i) \nu(B_i) \mid S \subset \bigcup_{i=1}^{\infty} A_i \times B_i, A_i \in \mathcal{A}, B_i \in \mathcal{B} \right\}$$

das *äußere Produktmaß*. Mit \mathcal{C} bezeichnen wir die σ -Algebra der $\mu \otimes \nu$ -messbaren Mengen.

Definition 18.7.2. Sind (X, \mathcal{A}) und (Y, \mathcal{B}) messbare Räume, so nennen wir

$$\mathcal{A} * \mathcal{B} = \{A \times B \mid A \in \mathcal{A}, B \in \mathcal{B}\}$$

das Mengensystem der *Rechteckmengen (Produktmengen)* von \mathcal{A} und \mathcal{B} . Die von $\mathcal{A} * \mathcal{B}$

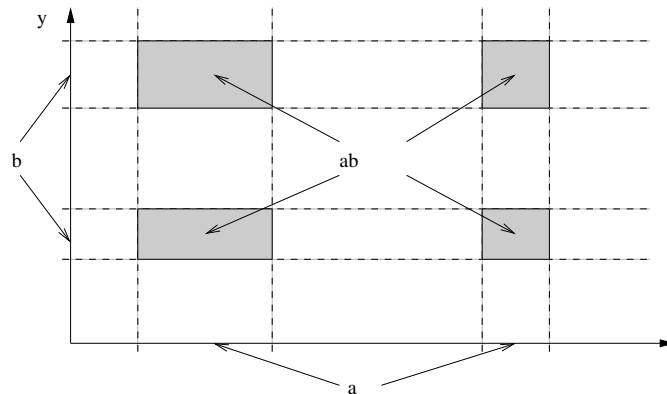


Abbildung 18.11: Beispiel einer Rechteckmenge

erzeugte σ -Algebra bezeichnen wir mit $\mathcal{A} \otimes \mathcal{B}$. Sie heißt die *Produkt- σ -Algebra* über $X \times Y$.

Wir werden zeigen, dass $\mathcal{A} \otimes \mathcal{B}$ in \mathcal{C} enthalten ist.

Lemma 18.7.3. Es seien $A_1 \times B_1 \in \mathcal{A} * \mathcal{B}$ und $A_2 \times B_2 \in \mathcal{A} * \mathcal{B}$, so ist

$$(A_1 \times B_1) \cap (A_2 \times B_2) = (A_1 \cap A_2) \times (B_1 \cap B_2) \in \mathcal{A} * \mathcal{B}$$

und

$$(A_1 \times B_1) \setminus (A_2 \times B_2) = ((A_1 \setminus A_2) \times B_1) \cup ((A_1 \cap A_2) \times (B_1 \setminus B_2))$$

die disjunkte Vereinigung zweier Elemente in $\mathcal{A} * \mathcal{B}$. Insbesondere ist

$$(X \times Y) \setminus (A_1 \times B_1) = (A_1 \times B_1)^c$$

die disjunkte Vereinigung von Elementen in $\mathcal{A} * \mathcal{B}$.

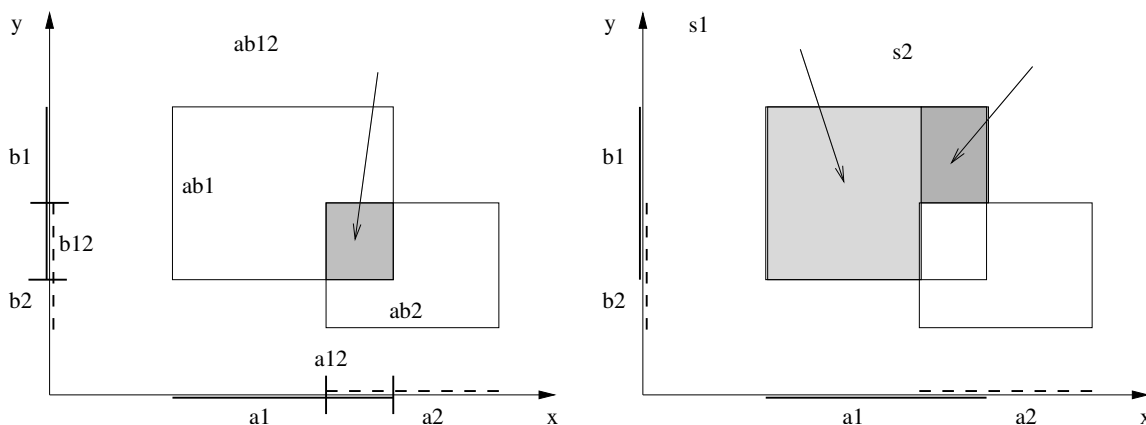


Abbildung 18.12: zu Lemma 18.7.3

Lemma 18.7.4. *Betrachte das Mengensystem*

$$\Omega := \left\{ \bigcup_{j=1}^{\infty} R_j \mid R_j \in \mathcal{A} * \mathcal{B} \right\}.$$

Dann gilt:

$$\Omega = \left\{ \bigcup_{j=1}^{\infty} R'_j \mid R'_j \in \mathcal{A} * \mathcal{B} \text{ und } R'_j \cap R'_k = \emptyset, \text{ falls } j \neq k \right\}.$$

Außerdem ist $S \setminus R \in \Omega$ und $S \cap R \in \Omega$ für $S \in \Omega$ und $R \in \mathcal{A} * \mathcal{B}$.

Beweis. Ist $(R_j)_{j \in \mathbb{N}} \in \mathcal{A} * \mathcal{B}$, so sind die Mengen $B_n := R_n \setminus \bigcup_{j=1}^{n-1} R_j$ disjunkt und $\bigcup_{n=1}^{\infty} B_n =$

$\bigcup_{n=1}^{\infty} R_n$. Da

$$B_n = R_n \setminus \bigcup_{j=1}^{n-1} R_j = R_n \cap (R_1 \cup \dots \cup R_{n-1})^c = R_n \cap R_1^c \cap \dots \cap R_{n-1}^c,$$

lässt sich B_n wie folgt als disjunkte Vereinigung von Elementen in $\mathcal{A} * \mathcal{B}$ schreiben: Wegen Lemma 18.7.3 sind die Mengen R_j^c von der Form $R_j^c = R_j^1 \cup R_j^2$ mit $R_j^1, R_j^2 \in \mathcal{A} * \mathcal{B}$ und $R_j^1 \cap R_j^2 = \emptyset$. Aus dem Distributivgesetz der Mengenlehre (siehe Satz 1.1.7) folgt

$$B_n = \bigcup_{k_j \in \{1,2\}} (R_n \cap R_1^{k_1} \cap \dots \cap R_{n-1}^{k_{n-1}}).$$

Wegen Lemma 18.7.3 sind die Mengen $R_n \cap R_1^{k_1} \cap \dots \cap R_{n-1}^{k_{n-1}}$ alle in $\mathcal{A} * \mathcal{B}$ enthalten. Sie sind paarweise disjunkt, denn

$$(R_n \cap R_1^{k_1} \cap \dots \cap R_{n-1}^{k_{n-1}}) \cap (R_n^{l_1} \cap \dots \cap R_{n-1}^{l_{n-1}}) = \emptyset,$$

falls $k_j \neq l_j$ für ein $j \in \{1, \dots, n-1\}$. □

Sei $S \subset X \times Y$ und $x \in X, y \in Y$. Dann heißen die Teilmengen

$$S^y = \{x \in X \mid (x, y) \in S\} \text{ bzw. } S_x = \{y \in Y \mid (x, y) \in S\}$$

Horizontalschnitt bzw. Vertikalschnitt von S bezüglich x bzw. y . Die Horizontalschnitte be-

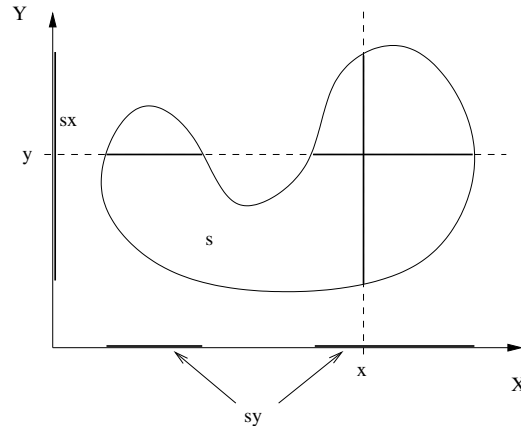


Abbildung 18.13: Horizontal- und Vertikalschnitte

sitzen folgende Eigenschaften: Für $(S_i)_{i \in \mathbb{N}} \subset X \times Y$ gilt

$$\left(\bigcup_{i=1}^{\infty} S_i\right)^y = \bigcup_{i=1}^{\infty} (S_i)^y \quad \text{und} \quad \left(\bigcap_{i=1}^{\infty} S_i\right)^y = \bigcap_{i=1}^{\infty} (S_i)^y$$

sowie

$$(S_1 \setminus S_2)^y = (S_1)^y \setminus (S_2)^y \quad \text{und} \quad (S^c)^y = (S^y)^c$$

für $S_1, S_2 \subset X \times Y$. Entsprechendes gilt für die Vertikalschnitte S_x .

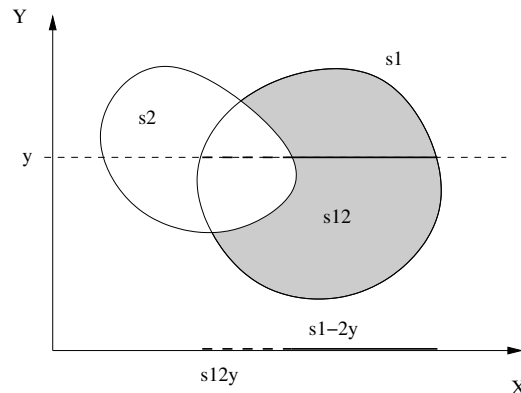


Abbildung 18.14: zu den Eigenschaften von Schnitten

Lemma 18.7.5. Die beiden Mengensysteme

$$H := \{S \subset X \times Y \mid S^y \in \mathcal{A} \forall y \in Y\} \quad \text{bzw.} \quad V := \{S \subset X \times Y \mid S_x \in \mathcal{B} \forall x \in X\}$$

der Teilmengen von $X \times Y$ mit messbaren Horizontalschnitten bzw. Vertikalschnitten sind σ -Algebren, die jeweils die σ -Algebra $\mathcal{A} \otimes \mathcal{B}$ enthalten.

Beweis. Wir zeigen dies für H . Zunächst ist $\mathcal{A} * \mathcal{B}$ in diesem Mengensystem enthalten. Denn ist

$$R = A \times B \in \mathcal{A} * \mathcal{B},$$

so ist für jedes $y \in Y$

$$R^y = \begin{cases} A, & \text{falls } y \in B \\ \emptyset, & \text{falls } y \notin B \end{cases} \in \mathcal{A}$$

und somit sind die Horizontalschnitte von R messbar (d.h. in \mathcal{A}). Haben alle S_i für $i \in \mathbb{N}$ messbare Horizontalschnitte, so auch $\bigcup_{i=1}^{\infty} S_i$ wegen

$$\left(\bigcup_{i=1}^{\infty} S_i \right)^y = \bigcup_{i=1}^{\infty} (S_i)^y.$$

Besitzen $S_1, S_2 \subset X \times Y$ messbare Horizontalschnitte, so auch $S_1 \setminus S_2$ wegen

$$(S_1 \setminus S_2)^y = (S_1)^y \setminus (S_2)^y.$$

Damit bildet H eine σ -Algebra, die $\mathcal{A} * \mathcal{B}$ enthält. Da $\mathcal{A} \otimes \mathcal{B}$ die kleinste σ -Algebra ist, die $\mathcal{A} * \mathcal{B}$ umfasst, ist $\mathcal{A} \otimes \mathcal{B}$ in der σ -Algebra H enthalten. Genauso bildet V eine σ -Algebra, die $\mathcal{A} \otimes \mathcal{B}$ umfasst. \square

Bemerkung. Ist also $S \in \mathcal{A} \otimes \mathcal{B}$, so sind sowohl ihre Horizontalschnitte S^y als auch ihre Vertikalschnitte S_x messbar, d.h. $S^y \in \mathcal{A}$ und $S_x \in \mathcal{B}$. Damit sind auch $\mu(S^y)$ und $\nu(S_x)$ erklärt. Betrachte die folgende Teilmenge

$$\mathcal{F} = \{S \in \mathcal{A} \otimes \mathcal{B} \mid y \mapsto \mu(S^y) \in M^+(Y, \mathcal{B})\}$$

von $\mathcal{A} \otimes \mathcal{B}$. Für $S \in \mathcal{F}$ ist dann das Integral

$$\rho(S) := \int_Y \mu(S^y) d\nu(y) = \int_Y \int_X \chi_{S^y}(x) d\mu(x) d\nu(y) = \int_Y \int_X \chi_S(x, y) d\mu(x) d\nu(y)$$

erklärt. Wir wollen zeigen, dass \mathcal{F} unter gewissen Voraussetzungen mit der σ -Algebra $\mathcal{A} \otimes \mathcal{B}$ übereinstimmt. Als ersten Schritt beweisen wir einige wichtige allgemeine Eigenschaften des Mengensystemes \mathcal{F} .

Lemma 18.7.6. Das Mengensystem $\mathcal{A} * \mathcal{B}$ ist in \mathcal{F} enthalten. Sind $A \in \mathcal{A}$ und $B \in \mathcal{B}$, so gilt für $R = A \times B$:

$$\rho(R) = \mu(A)\nu(B).$$

Ist darüber hinaus $(M_n)_{n \in \mathbb{N}}$ eine disjunkte Folge in \mathcal{F} , so ist auch ihre Vereinigung $\bigcup_{n=1}^{\infty} M_n$ in \mathcal{F} enthalten und es gilt

$$\rho\left(\bigcup_{n=1}^{\infty} M_n\right) = \sum_{n=1}^{\infty} \rho(M_n).$$

Da jedes $S \in \Omega$ sich als disjunkte Vereinigung von Elementen aus $\mathcal{A} * \mathcal{B}$ darstellen lässt, ist auch Ω in \mathcal{F} enthalten.

Beweis. Sei $R = A \times B \in \mathcal{A} * \mathcal{B}$, so folgt:

$$\mu(R^y) = \begin{cases} \mu(A) & y \in B \\ 0 & y \notin B. \end{cases}$$

Also erhalten wir: $y \mapsto \mu(R^y) = \mu(A)\chi_B(y) \in M^+(Y, \mathcal{B})$ sowie

$$\rho(R) := \int_Y \mu(R^y) d\nu(y) = \int_Y \mu(A)\chi_B(y) d\nu(y) = \mu(A)\nu(B).$$

Betrachte nun eine disjunkte Folge (M_n) von Mengen in \mathcal{F} . Dann sind auch die Horizontalschnitte $(M_n)^y$ für jedes $y \in Y$ disjunkt und es gilt für $M = \bigcup_{n=1}^{\infty} M_n$:

$$\mu(M^y) = \mu\left(\bigcup_{n=1}^{\infty} M_n^y\right) = \sum_{n=1}^{\infty} \mu(M_n^y).$$

Insbesondere ist $y \mapsto \mu(M^y)$ messbar als Limes der messbaren Funktionen $y \mapsto \sum_{n=1}^k \mu(M_n^y)$ für $k \rightarrow \infty$ gemäß Satz 18.4.5.

Außerdem folgt mit Hilfe des Satzes von Beppo Levi:

$$\begin{aligned} \rho\left(\bigcup_{n=1}^{\infty} M_n\right) &= \int_Y \mu\left(\bigcup_{n=1}^{\infty} M_n^y\right) d\nu(y) = \int_Y \sum_{n=1}^{\infty} \mu(M_n^y) d\nu(y) \\ &\stackrel{\text{B. Levi}}{=} \sum_{n=1}^{\infty} \int_Y \mu(M_n^y) d\nu(y) = \sum_{n=1}^{\infty} \rho(M_n). \end{aligned}$$

□

Es bestehen folgende Beziehungen zwischen dem äußeren Maß $\mu \otimes \nu$ und ρ :

Lemma 18.7.7. Sei $S \in \Omega$, so gilt

$$\rho(S) = \mu \otimes \nu(S).$$

Sei $E \subset X \times Y$ eine beliebige Menge, so folgt

$$\mu \otimes \nu(E) = \inf\{\rho(S) \mid E \subset S \in \Omega\}.$$

Beweis. Sei $S \in \Omega$ und $S \subset \bigcup_{i=1}^{\infty} A_i \times B_i$ mit $A_i \in \mathcal{A}$ und $B_i \in \mathcal{B}$. Dann folgt

$$\rho(S) = \int_Y \mu(S^y) d\nu(y) \leq \int_Y \mu\left(\bigcup_{i=1}^{\infty} (A_i \times B_i)^y\right) d\nu(y) \leq \sum_{i=1}^{\infty} \int_Y (A_i \times B_i)^y d\nu(y) = \sum_{i=1}^{\infty} \mu(A_i)\nu(B_i). \tag{18.9}$$

Nach Definition von $\mu \otimes \nu$ gilt somit: $\rho(S) \leq \mu \otimes \nu(S)$. Auf der anderen Seite lässt sich $S \in \Omega$ als abzählbare disjunkte Vereinigung $S = \bigcup_{i=1}^{\infty} A'_i \times B'_i$ mit $A'_i \times B'_i \in \mathcal{A} \times \mathcal{B}$ schreiben. Dann folgt:

$$\rho(S) = \sum_{i=1}^{\infty} \rho(A'_i \times B'_i) = \sum_{i=1}^{\infty} \mu(A'_i)\nu(B'_i).$$

Aus der Definition von $\mu \otimes \nu$ folgt daher auch: $\rho(S) \geq \mu \otimes \nu(S)$. Sei nun $E \subset X \times Y$ eine beliebige Menge. Dann gilt mit (18.9):

$$\begin{aligned} \mu \otimes \nu(E) &= \inf \left\{ \sum_{i=1}^{\infty} \mu(A_i) \nu(B_i) \mid S = \bigcup_{i=1}^{\infty} A_i \times B_i \supset E \right\} \\ &\geq \inf \{ \rho(S) \mid E \subset S \in \Omega \} = \inf \{ \mu \otimes \nu(S) \mid E \subset S \in \Omega \} \\ &\geq \mu \otimes \nu(E). \end{aligned}$$

□

Satz 18.7.8. *Seien (X, \mathcal{A}, μ) und (Y, \mathcal{B}, ν) zwei Maßräume. Sei $M \in \mathcal{A} \otimes \mathcal{B}$, so ist M messbar bezüglich des äußeren Maßes $\mu \otimes \nu$. Insbesondere definiert $\mu \otimes \nu$ ein Maß auf $\mathcal{A} \otimes \mathcal{B}$.*

Beweis. Da die $\mu \otimes \nu$ -messbaren Mengen eine σ -Algebra bilden und $\mathcal{A} \otimes \mathcal{B}$ die kleinste σ -Algebra darstellt, die $\mathcal{A} * \mathcal{B}$ enthält, genügt es, die Messbarkeit der Mengen $A \times B \in \mathcal{A} * \mathcal{B}$ zu zeigen. Seien also $A \times B \in \mathcal{A} * \mathcal{B}$ und $E \subset X \times Y$ gegeben. Sei $S \in \Omega$ und $E \subset S \in \Omega$. Dann sind auch $S \setminus (A \times B)$ und $S \cap (A \times B)$ Elemente von Ω und aus Lemma 18.7.7 folgt:

$$\mu \otimes \nu(E \setminus (A \times B)) + \mu \otimes \nu(E \cap (A \times B)) \leq \rho(S \setminus (A \times B)) + \rho(S \cap (A \times B)) = \rho(S)$$

für alle $S \supset E$ mit $S \in \Omega$. Wieder mit Lemma 18.7.7 erhalten wir durch Bildung des Infimum:

$$\mu \otimes \nu(E \setminus (A \times B)) + \mu \otimes \nu(E \cap (A \times B)) \leq \inf \{ \rho(S) \mid S \supset E \} = \mu \otimes \nu(E),$$

d.h. $A \times B$ ist messbar. □

Nun wollen wir unter gewissen Voraussetzungen zeigen, dass auch ρ ein Maß auf $\mathcal{A} \otimes \mathcal{B}$ definiert, welches mit $\mu \otimes \nu$ übereinstimmt. Dazu müssen wir nur noch zeigen, dass \mathcal{F} mit der von $\mathcal{A} * \mathcal{B}$ erzeugten σ -Algebra $\mathcal{A} \otimes \mathcal{B}$ übereinstimmt. Die σ -Additivität ist schon in Lemma 18.7.6 bewiesen worden.

Wie wir jetzt sehen werden, stimmt \mathcal{F} mit $\mathcal{A} \otimes \mathcal{B}$ überein, falls das Maß μ σ -endlich ist.

Definition 18.7.9. Sei (X, \mathcal{A}, μ) ein Maßraum. Dann heißt μ σ -endlich, falls eine Folge $(A_n)_{n \in \mathbb{N}}$ von \mathcal{A} -messbaren Mengen existiert mit $\mu(A_n) < \infty$ für alle n und $\bigcup_{n=1}^{\infty} A_n = X$.

Bemerkungen.

- Ist μ σ -endlich, so existiert eine monoton wachsende Folge $(B_n)_{n \in \mathbb{N}} \in \mathcal{A}$ mit den Eigenschaften $\mu(B_n) < \infty$ und $\bigcup_{n=1}^{\infty} B_n = X$. Betrachte dazu die Folge $B_n := A_1 \cup \dots \cup A_n$.
- Während das Lebesgue-Maß σ -endlich ist (denn der \mathbb{R}^n lässt sich als abzählbare Vereinigung von Quadern darstellen), ist das Zählmaß (siehe Beispiel (b) nach Definition 18.1.1) einer überabzählbaren Menge X nicht σ -endlich. Denn nach Definition ist das Zählmaß einer Menge $A \subset X$ nur dann endlich, falls A endlich. Da eine abzählbare Vereinigung endlicher Mengen abzählbar ist, kann somit das Zählmaß einer überabzählbaren Mengen nicht σ -endlich sein.

Lemma 18.7.10. *Sei X eine Menge und \mathcal{A} eine σ -Algebra, die von einem durchschnittstabilen Mengensystem E erzeugt wird, d.h. sind $A_1, A_2 \in E$, so ist auch $A_1 \cap A_2 \in E$. Sei nun \mathcal{F} ein Mengensystem mit den Eigenschaften:*

1. $E \subset \mathcal{F} \subset \mathcal{A}$ und $X \in \mathcal{F}$.
2. Sind $A, B \in \mathcal{F}$ und $A \subset B$, so folgt $B \setminus A \in \mathcal{F}$.
3. Ist $(A_n)_{n \in \mathbb{N}} \in \mathcal{F}$ eine disjunkte Folge, so folgt: $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$.

Dann stimmen \mathcal{F} und \mathcal{A} überein.

Beweis. Den Beweis dieser Aussage kann man zum Beispiel in dem Buch von Heinz Bauer, "Maß- und Integrationstheorie", finden (siehe Satz 2.4). Man hat zu zeigen, dass die Vereinigung einer beliebigen Folge von Mengen aus \mathcal{F} wieder in \mathcal{F} liegt. Denn dann ist \mathcal{F} eine σ -Algebra. Sie stimmt mit \mathcal{A} überein, denn \mathcal{A} ist die kleinste σ -Algebra, die E enthält.

Um zu zeigen, dass die Vereinigung einer beliebigen Folge von Mengen aus \mathcal{F} wieder in \mathcal{F} liegt, betrachtet man die zugeordnete disjunkte Folge. Ihre Elemente sind dann in \mathcal{F} , falls der Durchschnitt zweier Elemente von \mathcal{F} wieder in \mathcal{F} liegt. Für den Erzeuger E ist dies nach Annahme richtig. Mit Hilfe eines kleinen aber genialen Tricks beweist man es dann für beliebige Elemente. \square

Satz 18.7.11. Seien (X, \mathcal{A}, μ) und (Y, \mathcal{B}, ν) zwei Maßräume und das Maß μ sei σ -endlich. Dann gilt

$$\mathcal{F} := \{S \in \mathcal{A} \otimes \mathcal{B} \mid y \mapsto \mu(S^y) \in M^+(Y, \mathcal{B})\} = \mathcal{A} \otimes \mathcal{B}.$$

Beweis. Wir haben schon in Lemma 18.7.6 gezeigt, dass \mathcal{F} den durchschnittstabilen Erzeuger $\mathcal{A} * \mathcal{B}$ von $\mathcal{A} \otimes \mathcal{B}$ enthält. Wir haben dort auch gezeigt, dass die abzählbare Vereinigung disjunkter Mengen aus \mathcal{F} wieder in \mathcal{F} liegt. Es bleibt daher zu zeigen: Sind $S_1 \subset S_2 \in \mathcal{F}$, so auch $S_2 \setminus S_1$. Wir nehmen zunächst an, dass das Maß μ endlich ist, d.h. $\mu(X) < \infty$. Dann ist

$$\mu((S_2 \setminus S_1)^y) = \mu((S_2)^y \setminus (S_1)^y) = \mu((S_2)^y) - \mu((S_1)^y),$$

denn $\mu((S_2)^y) < \infty$. Da nach Annahme die Abbildungen $y \mapsto \mu((S_i)^y)$ für $i \in \{1, 2\}$ messbar sind, ist somit auch $y \mapsto \mu((S_2 \setminus S_1)^y)$ messbar. Also ist $S_2 \setminus S_1 \in \mathcal{F}$.

Sei μ nun σ -endlich. Dann existiert eine monoton wachsende Folge $(A_n)_{n \in \mathbb{N}}$ in \mathcal{A} mit $\mu(A_n) < \infty$ und $\bigcup_{n=1}^{\infty} A_n = X$. Betrachte das endliche Maß $\mu_n : \mathcal{A} \rightarrow [0, \infty)$ mit $\mu_n(A) := \mu(A \cap A_n)$.

Sei $S \in \mathcal{A} \otimes \mathcal{B}$, so ist wegen obiger Ausführungen die Funktion $y \mapsto \mu_n(S^y)$ eine Element von $M^+(Y, \mathcal{B})$ und somit insbesondere messbar. Da wegen Satz 18.1.11

$$\mu(S^y) = \lim_{n \rightarrow \infty} \mu(S^y \cap A_n) = \lim_{n \rightarrow \infty} \mu_n(S^y),$$

ist auch $y \mapsto \mu(S^y)$ als Limes messbarer Funktionen wieder messbar. \square

Bemerkung. Sei $M \in \mathcal{A} \otimes \mathcal{B}$, so folgt aus Lemma 18.7.7

$$\rho(M) \leq \inf\{\rho(S) \mid M \subset S \in \Omega\} = \mu \otimes \nu(M).$$

Wir werden nun zeigen: Sind sowohl μ als auch ν σ -endlich, so stimmen die Maße $\mu \otimes \nu$ und ρ überein.

Satz 18.7.12 (*Prinzip von Cavalieri*). Seien (X, \mathcal{A}, μ) und (Y, \mathcal{B}, ν) zwei Maßräume mit σ -endlichen Maßen μ und ν . Sei $M \in \mathcal{A} \otimes \mathcal{B}$, so gilt $y \mapsto \mu(M^y) \in M^+(Y, \mathcal{B})$ und $x \mapsto \nu(M_x) \in M^+(X, \mathcal{A})$ und es folgt:

$$\mu \otimes \nu(M) = \int_Y \mu(M^y) d\nu(y) = \int_X \nu(M_x) d\mu(x).$$

Beweis. Die Messbarkeit der Funktionen $y \mapsto \mu(M^y)$ und $x \mapsto \nu(M_x)$ ist eine Konsequenz von Satz 18.7.11.

Für die Integralgleichheit genügt es nun zu zeigen:

$$\rho(M) := \int_Y \mu(M^y) d\nu(y) = \mu \otimes \nu(M),$$

denn durch Vertauschen der Rollen von X und Y erhalten wir auch

$$\int_X \nu(M_x) d\mu(x) = \mu \otimes \nu(M).$$

Wie oben bemerkt, gilt $\rho(M) \leq \mu \otimes \nu(M)$ für alle $M \in \mathcal{A} \otimes \mathcal{B}$. Sei nun zunächst $M \subset A \times B \in \mathcal{A} * \mathcal{B}$ und $\mu(A), \nu(B) < \infty$. Dann folgt mit 18.7.6:

$$\begin{aligned} \mu \otimes \nu((A \times B) \setminus M) &= \mu \otimes \nu(A \times B) - \mu \otimes \nu(M) \leq \mu \otimes \nu(A \times B) - \rho(M) \\ &= \rho((A \times B) \setminus M) \leq \mu \otimes \nu((A \times B) \setminus M). \end{aligned}$$

Wir erhalten

$$\mu \otimes \nu(A \times B) - \mu \otimes \nu(M) = \mu \otimes \nu(A \times B) - \rho(M)$$

und somit $\rho(M) = \mu \otimes \nu(M)$. Ist nun $M \subset \mathcal{A} \otimes \mathcal{B}$ beliebig, so existieren wegen der σ -Endlichkeit von μ und ν monoton wachsende Folgen $(A_n)_{n \in \mathbb{N}}$, $A_n \in \mathcal{A}$ und $(B_n)_{n \in \mathbb{N}}$, $B_n \in \mathcal{B}$ mit jeweils endlichen Maßen und $\bigcup_{n=1}^{\infty} A_n = X$ sowie $\bigcup_{n=1}^{\infty} B_n = Y$. Dann ist auch $C_n = A_n \cup B_n$

eine monoton wachsende messbare Folge mit $\bigcup_{n=1}^{\infty} C_n = X \times Y$ und es gilt

$$\mu \otimes \nu(M) = \lim_{n \rightarrow \infty} \mu \otimes \nu(M \cap C_n) = \lim_{n \rightarrow \infty} \rho(M \cap C_n) = \rho(M).$$

□

Bemerkung. Das Prinzip von Cavalieri ist das wichtigste Werkzeug zur Berechnung von Produktmaßen. Die Verallgemeinerung auf Integrale liefert der Satz von Fubini.

Satz 18.7.13 (*Fubini*). Es seien (X, \mathcal{A}, μ) und (Y, \mathcal{B}, ν) zwei σ -endliche Maßräume. Dann gelten die folgenden Aussagen:

(a) Für jedes $f \in M^+(X \times Y, \mathcal{A} \otimes \mathcal{B})$ folgt

$$y \mapsto \int_X f(x, y) d\mu(x) \in M^+(Y, \mathcal{B}) \quad \text{und} \quad x \mapsto \int_Y f(x, y) d\nu(y) \in M^+(X, \mathcal{A})$$

und

$$\int_{X \times Y} f(x, y) d(\mu \otimes \nu)(x, y) = \int_Y \int_X f(x, y) d\mu(x) d\nu(y) = \int_X \int_Y f(x, y) d\nu(y) d\mu(x).$$

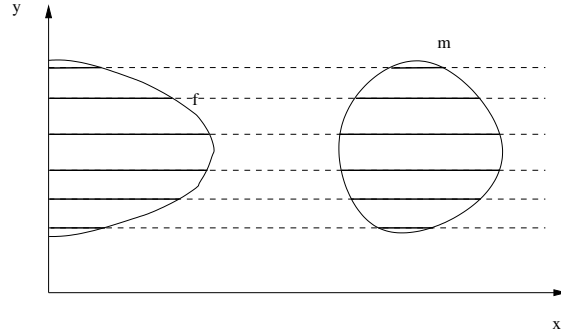


Abbildung 18.15: Zum Prinzip von Cavalieri. Speziell für $(X, \mu) = (\mathbb{R}, \lambda)$ kann man $f(y) := \lambda(M^y)$ als Graph über Y veranschaulichen.

(b) Sei $f : X \times Y \rightarrow \overline{\mathbb{R}}$ integrierbar bezüglich $\mu \otimes \nu$. Dann ist $f(x, \cdot)$ ν -integrierbar für μ -fast alle $x \in X$ und $f(\cdot, y)$ ist μ -integrierbar für ν -fast alle $y \in Y$. Außerdem sind die fast überall definierten Funktionen

$$y \mapsto \int_X f(x, y) d\mu(x) \quad \text{und} \quad x \mapsto \int_Y f(x, y) d\nu(y)$$

ν - bzw. μ -integrierbar und es folgt:

$$\int_{X \times Y} f(x, y) d(\mu \otimes \nu)(x, y) = \int_Y \int_X f(x, y) d\mu(x) d\nu(y) = \int_X \int_Y f(x, y) d\nu(y) d\mu(x).$$

Beweis. Wir beweisen Teil (a) in 3 Schritten:

1. Im ersten Schritt beweisen wir (a) für eine charakteristische Funktion χ_M mit $M \in \mathcal{A} \otimes \mathcal{B}$. In diesem Fall ist der Beweis eine direkte Konsequenz des Prinzips von Cavalieri. Wegen $\chi_M(x, y) = \chi_{M^y}(x) = \chi_{M_x}(y)$ folgt

$$y \mapsto \mu(M^y) = \int_X \chi_{M^y}(x) d\mu(x) = \int_X \chi_M(x, y) d\mu(x) \in M^+(Y, \mathcal{B})$$

und

$$x \mapsto \nu(M_x) = \int_Y \chi_{M_x}(y) d\nu(y) = \int_Y \chi_M(x, y) d\nu(y) \in M^+(X, \mathcal{A}).$$

Außerdem gilt:

$$\begin{aligned} \int_{X \times Y} \chi_M(x, y) d(\mu \otimes \nu)(x, y) &= \mu \otimes \nu(M) \stackrel{\text{Prinzip von Cavalieri}}{=} \int_Y \mu(M^y) d\nu(y) \\ &= \int_Y \int_X \chi_{M^y}(x) d\mu(x) d\nu(y) = \int_Y \int_X \chi_M(x, y) d\mu(x) d\nu(y). \end{aligned}$$

Genauso folgt

$$\mu \otimes \nu(M) = \int_X \nu(M_x) d\mu(x) = \int_X \int_Y \chi_M(x, y) d\nu(y) d\mu(x).$$

2. Die Aussage in (a) gilt auch für alle Treppenfunktionen $\varphi \in T^+(X \times Y, \mathcal{A} \otimes \mathcal{B})$. Denn ist $\varphi = \sum_{j=1}^m c_j \chi_{M_j}$ mit $c_j \geq 0$ und $M_j \in \mathcal{A} \otimes \mathcal{B}$, so folgt aus Schritt 1.

$$y \mapsto \int_X \varphi(x, y) d\mu(x) = \sum_{j=1}^m c_j \int_X \chi_{M_j}(x, y) d\mu(x) \in M^+(Y, \mathcal{B})$$

und

$$x \mapsto \int_Y \varphi(x, y) d\nu(y) = \sum_{j=1}^m c_j \int_Y \chi_{M_j}(x, y) d\nu(y) \in M^+(X, \mathcal{A})$$

und somit mit Hilfe der Linearität des Integrals

$$\begin{aligned} \int_{X \times Y} \varphi d(\mu \otimes \nu) &\stackrel{\text{Def}}{=} \sum_{j=1}^m c_j \int_{X \times Y} \chi_{M_j} d(\mu \otimes \nu) \stackrel{1.}{=} \sum_{j=1}^m c_j \int_Y \int_X \chi_{M_j}(x, y) d\mu(x) d\nu(y) \\ &= \int_Y \int_X \varphi(x, y) d\mu(x) d\nu(y). \end{aligned}$$

Genauso folgt:

$$\int_{X \times Y} \varphi d(\mu \otimes \nu) = \int_X \int_Y \varphi(x, y) d\mu(x) d\nu(y).$$

3. Sei nun $f \in M^+(X \times Y, \mathcal{A} \otimes \mathcal{B})$. Dann existiert eine monoton wachsende Folge von Treppenfunktionen $\varphi_n \in T^+(X \times Y, \mathcal{A} \otimes \mathcal{B})$, $n \in \mathbb{N}$ mit $\lim_{n \rightarrow \infty} \varphi_n(x, y) = f(x, y)$. Somit erhalten wir für jedes $y \in Y$ eine monoton wachsende Folge $x \mapsto \varphi_n(x, y)$ von Treppenfunktionen in $T^+(X, \mathcal{A})$, die punktweise gegen $x \mapsto f(x, y)$ konvergiert, und nach Definition des Integrals folgt

$$\int_X f(x, y) d\mu(x) = \lim_{n \rightarrow \infty} \int_X \varphi_n(x, y) d\mu(x). \quad (18.10)$$

Außerdem ist die Funktion $y \mapsto \int_X f(x, y) d\mu(x)$ als Limes messbarer Funktionen selbst messbar. Des Weiteren folgt mit Hilfe von Schritt 2 und Anwendung des Satzes 18.5.9 von Beppo Levi

$$\begin{aligned} \int_{X \times Y} f d(\mu \otimes \nu) &\stackrel{\text{Def}}{=} \lim_{n \rightarrow \infty} \int_{X \times Y} \varphi_n d(\mu \otimes \nu) \stackrel{2.}{=} \lim_{n \rightarrow \infty} \int_Y \int_X \varphi_n(x, y) d\mu(x) d\nu(y) \\ &\stackrel{B. Levi}{=} \int_Y \lim_{n \rightarrow \infty} \left(\int_X \varphi_n(x, y) d\mu(x) \right) d\nu(y) \\ &\stackrel{(18.10)}{=} \int_Y \left(\int_X f(x, y) d\mu(x) \right) d\nu(y). \end{aligned}$$

Analog folgt die Messbarkeit von $x \mapsto \int_Y f(x, y) d\nu(y)$ und die Identität

$$\int_{X \times Y} f(x, y) d(\mu \otimes \nu)(x, y) = \int_X \int_Y f(x, y) d\mu(x) d\nu(y).$$

Beweis von (b):

Sei nun f eine $\mu \otimes \nu$ -integrierbare Funktion und

$$f = f^+ - f^- \quad \text{mit} \quad f^+, f^- \in M^+(X \times Y, \mathcal{A} \otimes \mathcal{B})$$

die Zerlegung von f in seinen positiven und negativen Anteil. Wegen Teil (a) sind die Funktionen $y \mapsto \int_X f^+(x, y) d\mu(x)$ und $y \mapsto \int_X f^-(x, y) d\mu(x)$ beide ν -messbar. Aus der Integrierbarkeit von f folgt wegen Satz 18.5.11 auch die $\mu \otimes \nu$ -Integrierbarkeit von $|f|$ und wegen Teil (a) des Satzes ist die Abbildung $y \mapsto \int_X |f(x, y)| d\mu(x)$ ein Element in $M^+(Y, \mathcal{B})$. Außerdem ist deshalb

$$\int_Y \left(\int_X |f(x, y)| d\mu(x) \right) d\nu(y) = \int_{X \times Y} |f| d(\mu \otimes \nu) < \infty$$

und wegen Korollar 18.5.12 gilt $\int_X |f(x, y)| d\mu(x) < \infty$ für ν -fast alle $y \in Y$. Da nun

$$|f| = f^+ + f^- \quad \text{und} \quad f^+, f^- \in M^+(X \times Y, \mathcal{A} \otimes \mathcal{B}),$$

sind auch

$$\int_X f^+(x, y) d\mu(x) < \infty \quad \text{und} \quad \int_X f^-(x, y) d\mu(x) < \infty$$

für ν -fast alle $y \in Y$, nämlich für alle diejenigen $y \in Y$, für die das Integral $\int_X |f(x, y)| d\mu(x)$ endlich ist. Für solche y ist das Integral

$$\int_X f(x, y) d\mu(x) = \int_X f^+(x, y) d\mu(x) - \int_X f^-(x, y) d\mu(x) \quad (18.11)$$

definiert. Damit erhalten wir:

$$\begin{aligned} \int_{X \times Y} f d(\mu \otimes \nu) &\stackrel{\text{Def}}{=} \int_{X \times Y} f^+ d(\mu \otimes \nu) - \int_{X \times Y} f^- d(\mu \otimes \nu) \\ &\stackrel{(a)}{=} \int_Y \int_X f^+(x, y) d\mu(x) d\nu(y) - \int_Y \int_X f^-(x, y) d\mu(x) d\nu(y) \\ &= \int_Y \left(\int_X f^+(x, y) d\mu(x) - \int_X f^-(x, y) d\mu(x) \right) d\nu(y) \\ &\stackrel{(18.11)}{=} \int_Y \int_X f(x, y) d\mu(x) d\nu(y). \end{aligned}$$

□

Wie lässt sich die Produkt σ -Algebra aus Erzeugendensystemen berechnen?

Satz 18.7.14. *Es seien $(X, \mathcal{A}), (Y, \mathcal{B})$ messbare Räume und \mathcal{A} bzw. \mathcal{B} habe die Erzeugendensysteme $E_{\mathcal{A}}$ bzw. $E_{\mathcal{B}}$. Weiter nehme man an, es gebe Folgen $(A_n)_{n \in \mathbb{N}} \in E_{\mathcal{A}}$ und $(B_n)_{n \in \mathbb{N}} \in E_{\mathcal{B}}$ mit $\bigcup_{n=1}^{\infty} A_n = X$ und $\bigcup_{n=1}^{\infty} B_n = Y$. Wie in Definition 18.1.12 bezeichne $\sigma(E)$ die von E erzeugte σ -Algebra. Dann gilt:*

$$\mathcal{A} \otimes \mathcal{B} = \sigma(\{A \times B \mid A \in E_{\mathcal{A}}, B \in E_{\mathcal{B}}\})$$

Beweis. Da $\bigcup_{n=1}^{\infty} A_n = X$ mit $A_n \in E_{\mathcal{A}}$ und $\bigcup_{n=1}^{\infty} B_n = Y$ mit $B_n \in E_{\mathcal{B}}$ und weil ferner $E_{\mathcal{A}}$ und $E_{\mathcal{B}}$ Erzeugendensysteme sind, gilt:

$$\begin{aligned} \sigma(\{A \times B \mid A \in E_{\mathcal{A}}, B \in E_{\mathcal{B}}\}) &= \sigma(\{A \times Y \mid A \in E_{\mathcal{A}}\} \cup \{X \times B \mid B \in E_{\mathcal{B}}\}) \\ &= \sigma(\{A \times Y \mid A \in \mathcal{A}\} \cup \{X \times B \mid B \in \mathcal{B}\}) = \sigma(\{A \times B \mid A \in \mathcal{A}, B \in \mathcal{B}\}) = \mathcal{A} \otimes \mathcal{B}. \end{aligned}$$

□

Korollar 18.7.15. *Ist $\mathcal{B}(\mathbb{R}^k)$ die Borel σ -Algebra von \mathbb{R}^k , so gilt für $m, n \in \mathbb{N}$:*

$$\mathcal{B}(\mathbb{R}^m) \otimes \mathcal{B}(\mathbb{R}^n) = \mathcal{B}(\mathbb{R}^{m+n})$$

und

$$\lambda^m \otimes \lambda^n = \lambda^{m+n}.$$

Beweis. Ein Erzeugendensystem von $\mathcal{B}(\mathbb{R}^k)$ ist durch die Menge aller Quader $Q^k \subset \mathbb{R}^k$ gegeben. Wegen Satz 18.7.14 wird $\mathcal{B}(\mathbb{R}^m) \otimes \mathcal{B}(\mathbb{R}^n)$ durch

$$\{Q^m \times Q^n \mid Q^m \subset \mathbb{R}^m, Q^n \subset \mathbb{R}^n, \text{ mit } Q^m, Q^n \text{ Quader}\}$$

erzeugt. Da jeder Quader $Q \subset \mathbb{R}^{m+n}$ von der Form $Q^m \times Q^n$ ist, folgt die erste Behauptung. Des Weiteren gilt:

$$\lambda^{m+n}(Q^m \times Q^n) = \lambda^m(Q^m)\lambda^n(Q^n) \stackrel{\text{Lemma 18.7.7}}{=} \lambda^m \otimes \lambda^n(Q^m \times Q^n).$$

Ist $A \in \mathcal{B}(\mathbb{R}^{m+n})$, so folgt

$$\lambda^{m+n}(A) = \inf \left\{ \sum_{j=1}^{\infty} \lambda^m(Q_j^m)\lambda^n(Q_j^n) \mid A \subset \bigcup_{j=1}^{\infty} Q_j^m \times Q_j^n \right\} \geq \lambda^m \otimes \lambda^n(A).$$

Ist $A \subset Q \subset \mathbb{R}^{m+n}$, so folgt:

$$\lambda^{m+n}(Q) = \lambda^m \otimes \lambda^n(Q) = \lambda^m \otimes \lambda^n(Q \setminus A) + \lambda^m \otimes \lambda^n(A) \leq \lambda^{m+n}(Q \setminus A) + \lambda^m \otimes \lambda^n(A)$$

und somit auch die umgekehrte Ungleichung

$$\lambda^{m+n}(A) \leq \lambda^m \otimes \lambda^n(A).$$

Ist $A \subset \mathcal{B}(\mathbb{R}^{m+n})$ beliebig, so existiert eine Folge $(Q_j)_{j \in \mathbb{N}} \subset \mathbb{R}^{m+n}$ von Quadern mit $Q_j \subset Q_{j+1}$ und $\bigcup_{j=1}^{\infty} Q_j = \mathbb{R}^{m+n}$. Dann ist $\bigcup_{j=1}^{\infty} (A \cap Q_j) = A$ und

$$\lambda^{m+n}(A) = \lim_{j \rightarrow \infty} \lambda^{m+n}(A \cap Q_j) = \lim_{j \rightarrow \infty} \lambda^m \otimes \lambda^n(A \cap Q_j) = \lambda^m \otimes \lambda^n(A).$$

□

Bemerkung. Statt $\int_{\mathbb{R}^m} f d\lambda^m(x)$ schreiben wir auch $\int_{\mathbb{R}^m} f(x) d^m x$ oder $\int f(x) dx$. Ist $f : \mathbb{R}^{m+n} \rightarrow \overline{\mathbb{R}}$ integrierbar, so folgt aus dem Satz von Fubini:

$$\int_{\mathbb{R}^{m+n}} f(w) d^{m+n}(w) = \int_{\mathbb{R}^m} \left(\int_{\mathbb{R}^n} f(x, y) d^n x \right) d^m y.$$

Beispiel. (Volumenberechnung eines Drehkörpers) Sei $c : [a, b] \rightarrow \mathbb{R}^3$ eine Kurve mit $c(t) = (g(t), 0, t)$, wobei $g : [a, b] \rightarrow \mathbb{R}$ eine stetige Funktion ist mit $g(t) > 0$. Betrachte die zugehörige Drehfläche mit der Parametrisierung $f : [a, b] \times [0, 2\pi] \rightarrow \mathbb{R}^3$, wobei

$$f(t, \varphi) = (\cos(\varphi)g(t), \sin(\varphi)g(t), t).$$

Bezeichne mit

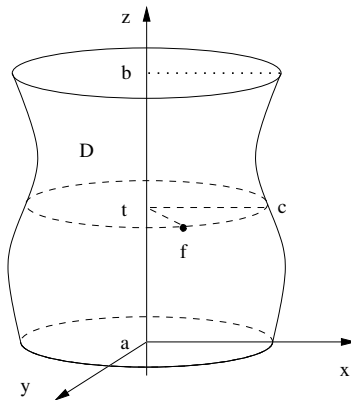


Abbildung 18.16: Drehkörper

$$D = \{(x, y, t) \mid t \in [a, b], \sqrt{x^2 + y^2} \leq g(t)\}$$

den zugehörigen "Drehkörper". Da D abgeschlossen ist, ist $D \in \mathcal{B}(\mathbb{R}^3)$. Schreibe $\lambda^3 = \lambda^2 \otimes \lambda^1$, so gilt nach dem Prinzip von Cavalieri (und der Kreisflächenformel):

$$\lambda^3(D) = \int_a^b \lambda^2(D^t) dt = \int_a^b g^2(t) \pi dt = \pi \int_a^b g^2(t) dt.$$

Ist $g(t) = r$, so ist D ein Zylinder und es gilt die Formel:

$$\lambda^3(D) = \pi r^2(b - a) \quad (\text{"Grundfläche mal Höhe"}).$$

18.8 Transformationsformel

Wir haben gesehen, wie sich das Lebesgue-Maß $\lambda : \mathcal{B}(\mathbb{R}^n) \rightarrow \mathbb{R}$ unter affin linearen Abbildungen $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ transformiert (siehe Satz 18.3.5): Ist T invertierbar mit $T(x) = a + L(x)$ und $L \in GL(n, \mathbb{R})$, so folgt:

$$\lambda(T^{-1}(B)) = \frac{\lambda(B)}{|\det L|}$$

für alle $B \in \mathcal{B}(\mathbb{R}^n)$. Da T ein Homöomorphismus ist, ist mit $A \in \mathcal{B}(\mathbb{R}^n)$ auch $T(A) \in \mathcal{B}(\mathbb{R}^n)$ und es folgt, indem man $T(A)$ für B einsetzt und umstellt:

$$\lambda(T(A)) = \lambda(A) |\det L| = \int_{\mathbb{R}^n} |\det L| \chi_A d\lambda = \int_A |\det L| d\lambda = \int_A |\det DT(x)| d\lambda,$$

denn $DT(x) = L$. Sind allgemeiner $X, Y \subset \mathbb{R}^n$ offen und $T : X \rightarrow Y$ ein Diffeomorphismus, so liegt daher die Vermutung nahe, dass die Identität

$$\lambda(T(A)) = \int_A |\det DT(x)| d\lambda(x)$$

für alle Borel-messbaren Mengen $A \subset X$ gilt. Für die Formulierung der Transformationsformel benötigen wir noch folgende Bezeichnung: Für $X \in \mathcal{B}(\mathbb{R}^n)$ sei

$$\mathcal{B}(X) := \{A \in \mathcal{B}(\mathbb{R}^n) \mid A \subset X\}.$$

Dann ist $\mathcal{B}(X)$ eine σ -Algebra.

Satz 18.8.1 (Transformationsformel). *Es seien $X, Y \subset \mathbb{R}^n$ offen, $T : X \rightarrow Y$ ein C^1 -Diffeomorphismus. Dann folgt:*

(a) *Für alle $A \in \mathcal{B}(X)$ gilt*

$$\lambda(T(A)) = \int_A |\det DT(x)| d\lambda(x).$$

(b) *Für alle $f \in M^+(Y, \mathcal{B}(Y))$ ist*

$$\int_Y f(y) d\lambda(y) = \int_X f \circ T(x) |\det DT(x)| d\lambda(x).$$

(c) *Eine Funktion $f : Y \rightarrow \overline{\mathbb{R}}$ ist genau dann λ -integrierbar über Y , falls $f \circ T |\det DT|$ über X λ -integrierbar ist. Dann gilt:*

$$\int_Y f(y) d\lambda(y) = \int_X f \circ T(x) |\det DT(x)| d\lambda(x).$$

Beweis. Wir werden diesen Satz in mehreren Schritten beweisen.

1. Schritt:

Sei $Q \subset X$ ein abgeschlossener Würfel, d.h. ein Quader mit fester Kantenlänge $d > 0$. Dann gilt:

$$\lambda(T(Q)) \leq \int_Q |\det DT(x)| d\lambda(x).$$

Beweis: Da Q kompakt und die Funktion $x \mapsto \|(DT(x))^{-1}\|$ stetig ist, existiert

$$\alpha := \max_{x \in Q} \|(DT(x))^{-1}\|.$$

(Für $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ ist $\|A\| := \sup\{\|A(v)\| \mid \|v\| \leq 1\}$, wobei wir auf \mathbb{R}^m bzw. \mathbb{R}^n die euklidischen Normen benutzen.)

Zu jedem $\epsilon > 0$ existiert wegen der gleichmäßigen Stetigkeit der Abbildung $x \mapsto DT(x) \in L(\mathbb{R}^n, \mathbb{R}^n)$ auf Q ein $\delta' > 0$ mit

$$\|DT(x) - DT(y)\| \leq \frac{\epsilon}{\alpha}, \quad (18.12)$$

sobald $\|x - y\| \leq \delta'$. Wähle $m \in \mathbb{N}$ so, dass $m > \frac{\sqrt{nd}}{\delta'}$ und setze $k := m^n$. Seien $\{Q_j\}_{1 \leq j \leq k}$ die abgeschlossenen Würfel der Kantenlänge $\frac{d}{m} =: \delta_0$, welche durch Unterteilung von Q in k gleich große Teilwürfel entstehen. Dann gilt

$$\text{diam}Q_j = \max\{\|x - y\| \mid x, y \in Q_j\} = \sqrt{n}\delta_0 =: \delta < \delta'$$

für $j \in \{1, \dots, k\}$. Zu jedem Q_j wähle $q_j \in Q_j$ mit

$$|\det DT(q_j)| = \min\{|\det DT(x)| \mid x \in Q_j\}.$$

Ist $h : Q_j \rightarrow \mathbb{R}^n$ eine stetig differenzierbare Abbildung, so folgt aus dem Mittelwertsatz (siehe Korollar 15.3.16) für $x, y \in Q_j$:

$$\|h(x) - h(y)\| \leq \|x - y\| \cdot \max\{\|Dh(z)\| \mid z \in Q_j\}.$$

Wende diese Ungleichung auf $h : Q_j \rightarrow \mathbb{R}^n$ mit $h(x) = T(x) - DT(q_j)(x - q_j)$ an. Dann ist ihre Ableitung gegeben durch $Dh(x) = DT(x) - DT(q_j)$ und es folgt mit Hilfe von (18.12):

$$\begin{aligned} & \|T(x) - T(q_j) - DT(q_j)(x - q_j)\| \\ &= \|h(x) - h(q_j)\| \leq \|x - q_j\| \cdot \max\{\|DT(z) - DT(q_j)\| \mid z \in Q_j\} \leq \|x - q_j\| \frac{\epsilon}{\alpha}, \end{aligned}$$

denn $\|x - q_j\| \leq \delta \leq \delta'$. Aus der Dreiecksungleichung ergibt sich somit:

$$\|T(x)\| \leq \|T(q_j)\| + \|DT(q_j)(x - q_j)\| + \delta \frac{\epsilon}{\alpha}.$$

Verwenden wir für Teilmengen A, B des \mathbb{R}^n die Bezeichnung

$$A + B := \{x + y \mid x \in A, y \in B\},$$

so impliziert diese Abschätzung :

$$T(Q_j) \subset T(q_j) + DT(q_j)(Q_j - q_j) + B(0, \delta \frac{\epsilon}{\alpha}).$$

Nun folgt aus der Definition von α

$$B(0, \delta \frac{\epsilon}{\alpha}) = DT(q_j)DT(q_j)^{-1}(B(0, \delta \frac{\epsilon}{\alpha})) \subset DT(q_j)(B(0, \delta\epsilon))$$

und somit

$$T(Q_j) \subset T(q_j) + DT(q_j)(Q_j + B(0, \delta\epsilon) - q_j).$$

Aus der linearen Transformationsformel 18.3.5 erhalten wir mittels der Translationsinvarianz von λ :

$$\lambda(T(Q_j)) \leq |\det DT(q_j)| \cdot \lambda(Q_j + B(0, \delta\epsilon) - q_j) = |\det DT(q_j)| \cdot \lambda(Q_j + B(0, \delta\epsilon)).$$

Da Q_j ein Würfel der Kantenlänge δ_0 ist, ist $Q_j + B(0, \delta\epsilon)$ in einem Würfel der Kantenlänge

$$\delta_0 + 2\delta\epsilon = \delta_0 \left(1 + \frac{2\delta}{\delta_0}\epsilon\right) = \delta_0(1 + 2\sqrt{n}\epsilon)$$

enthalten. Daraus folgt:

$$\lambda(T(Q_j)) \leq |\det DT(q_j)| \delta_0^n (1 + 2\sqrt{n}\epsilon)^n = |\det DT(q_j)| \cdot \lambda(Q_j) \cdot (1 + 2\sqrt{n}\epsilon)^n.$$

Summation über alle Q_j liefert

$$\begin{aligned} \lambda(T(Q)) &\leq (1 + 2\sqrt{n}\epsilon)^n \sum_{j=1}^k |\det DT(q_j)| \lambda(Q_j) \leq (1 + 2\sqrt{n}\epsilon)^n \cdot \sum_{j=1}^k \int_{Q_j} |\det DT(x)| d\lambda(x) \\ &= (1 + 2\sqrt{n}\epsilon)^n \int_Q |\det DT(x)| d\lambda(x). \end{aligned}$$

Da dies für alle $\epsilon > 0$ gilt, folgt die Behauptung.

Bemerkung. Die Aussage von Schritt 1 lässt sich auch auf Quader übertragen.

2. Schritt:

Sei $A \in \mathcal{B}(X)$, so gilt:

$$\lambda(T(A)) \leq \int_A |\det DT(x)| d\lambda(x).$$

Beweis: Man zeigt die Ungleichung erst für offene Mengen. Ist $U \subset X$ offen, so existiert eine abzählbare Familie von abgeschlossenen Quadern $(Q_j)_{j \in \mathbb{N}}$ mit $\bigcup_{j=1}^{\infty} Q_j = U$. Man kann sogar zeigen, dass man die Quader disjunkt wählen kann. Dann gilt:

$$\begin{aligned} \lambda(T(U)) &\leq \sum_{j=1}^{\infty} \lambda(T(Q_j)) \leq \sum_{j=1}^{\infty} \int_{Q_j} |\det DT(x)| d\lambda(x) = \sum_{j=1}^{\infty} \int_U \chi_{Q_j}(x) |\det DT(x)| d\lambda(x) \\ &\stackrel{\text{B. Levi}}{=} \int_U \sum_{j=1}^{\infty} \chi_{Q_j}(x) |\det DT(x)| d\lambda(x) = \int_U 1 \cdot |\det DT(x)| d\lambda(x). \end{aligned}$$

Sei nun $A \in \mathcal{B}(X)$, so gilt (siehe z.B. das Buch von Elstrodt):

$$\lambda(A) = \inf\{\lambda(U) \mid U \text{ offen}, A \subset U\}.$$

Daher existiert eine Folge $U_1 \supset U_2 \supset \dots \supset U_k \supset \dots \supset A$ offener Mengen mit $\lambda(A) = \lambda\left(\bigcap_{k=1}^{\infty} U_k\right)$.

Nehme zunächst an, dass $\int_X |\det DT(x)| d\lambda(x) < \infty$. Dann folgt mit dem Satz 18.6.2 von Lebesgue :

$$\begin{aligned} \lambda(T(A)) &\leq \lim_{k \rightarrow \infty} \int_{U_k} |\det DT(x)| d\lambda(x) = \lim_{k \rightarrow \infty} \int_X |\det DT(x)| \chi_{U_k}(x) d\lambda(x) \\ &= \int_X |\det DT(x)| \chi_A(x) d\lambda(x). \end{aligned}$$

Ist $\int_X |\det DT(x)| d\lambda(x) = \infty$, so betrachte die offenen Mengen

$$X_n = B(0, n) \cap \{x \in X \mid |\det DT(x)| < n\}.$$

Insbesondere ist $(X_n)_{n \in \mathbb{N}}$ eine monoton steigende Folge von Mengen mit $X = \bigcup_{n=1}^{\infty} X_n$ und $\int_{X_n} |\det DT(x)| d\lambda(x) < \infty$. Dann gilt:

$$\begin{aligned} \lambda(T(A)) &= \lambda\left(T\left(\bigcup_{n=1}^{\infty} A \cap X_n\right)\right) = \lambda\left(\bigcup_{n=1}^{\infty} T(A \cap X_n)\right) \stackrel{18.1,11}{=} \lim_{n \rightarrow \infty} \lambda(T(A \cap X_n)) \\ &\leq \lim_{n \rightarrow \infty} \int_X |\det DT(x)| \chi_{A \cap X_n} d\lambda(x) \leq \int_X |\det DT(x)| \chi_A(x) d\lambda(x) \\ &= \int_A |\det DT(x)| d\lambda(x). \end{aligned}$$

3. Schritt:

Für alle $f \in M^+(Y, \mathcal{B}(Y))$ gilt:

$$\int_Y f(y) d\lambda(y) \leq \int_X f \circ T(x) |\det DT(x)| d\lambda(x).$$

Beweis: Sei $f = \chi_B$ mit $B \in \mathcal{B}(Y)$ und $A := T^{-1}(B) \in \mathcal{B}(X)$. Dann gilt:

$$\int_Y f(y) d\lambda(y) = \lambda(B) = \lambda(T(A)) \leq \int_X \chi_A |\det DT(x)| d\lambda(x).$$

Da $\chi_B \circ T = \chi_{T^{-1}(B)} = \chi_A$, folgt die Behauptung für $f = \chi_B$. Damit folgt der 3. Schritt auch für alle Treppenfunktionen $f \in T^+(Y, \mathcal{B})$.

Für beliebiges $f \in M^+(Y, \mathcal{B})$ wähle eine monoton wachsende Folge $(f_n)_{n \in \mathbb{N}} \in T^+(Y, \mathcal{B}(X))$ mit $\lim_{n \rightarrow \infty} f_n(x) = f(x)$ für alle $x \in X$. Dann gilt:

$$\begin{aligned} \int_Y f d\lambda &= \lim_{n \rightarrow \infty} \int_Y f_n(y) d\lambda(y) \leq \lim_{n \rightarrow \infty} \int_X f_n \circ T(x) |\det DT(x)| d\lambda(x) \\ &= \int_X f \circ T(x) |\det DT(x)| d\lambda(x). \end{aligned}$$

4. Schritt:

Für alle $f \in M^+(Y, \mathcal{B}(Y))$ gilt:

$$\int_Y f(y) d\lambda(y) = \int_X f \circ T(x) |\det DT(x)| d\lambda(x).$$

Beweis: Betrachte den C^1 -Diffeomorphismus $T^{-1} : Y \rightarrow X$, so gilt für $g \in M^+(X, \mathcal{B}(X))$:

$$\int_X g(x) d\lambda(x) \leq \int_Y g \circ T^{-1}(y) |\det DT^{-1}(y)| d\lambda(y).$$

Setze $g(x) = f \circ T(x) |\det DT(x)|$, so folgt:

$$\int_X f \circ T(x) |\det DT(x)| d\lambda(x) \leq \int_Y f(y) |\det DT(T^{-1}(y))| \cdot |\det DT^{-1}(y)| d\lambda(y).$$

Da

$$\text{id} = D(T \circ T^{-1})(y) = DT(T^{-1}(y)) \circ DT^{-1}(y),$$

ist $|\det DT(T^{-1}(y)) \cdot \det DT^{-1}(y)| = |\det DT(T^{-1}(y)) \circ DT^{-1}(y)| = 1$ und die Behauptung folgt mit Schritt 3.

5. Schritt:

Ist $f : Y \rightarrow \overline{\mathbb{R}}$ integrierbar, so zerlege $f = f^+ - f^-$ und wende Schritt 4. auf f^+ und f^- an. \square

Beispiel. Betrachte die Transformation $T : (0, \infty) \times (0, 2\pi) \rightarrow \mathbb{R}^2$ mit $T(r, \varphi) = r \cdot (\cos \varphi, \sin \varphi)$. Dann ist T ein Diffeomorphismus auf sein Bild

$$\text{Bild } T = T((0, \infty) \times (0, 2\pi)) = \mathbb{R}^2 \setminus \{(x, 0) \mid x \geq 0\}.$$

Insbesondere ist Bild T fast ganz \mathbb{R}^2 bis auf eine Menge vom Maß 0 (bzgl. λ^2). Es gilt:

$$|\det DT(r, \varphi)| = \left| \det \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix} \right| = r.$$

Ist $f \in M^+(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$ oder f integrierbar, so folgt:

$$\int_{\mathbb{R}^2} f(x) d\lambda^2(x) = \int_{\mathbb{R} \times (0, 2\pi)} f(r \cos \varphi, r \sin \varphi) r d\lambda^2(r, \varphi) = \int_0^\infty \int_0^{2\pi} f(r \cos \varphi, r \sin \varphi) r d\varphi dr.$$

Wir benutzen diese Formel, um $\int_{-\infty}^{+\infty} e^{-x^2} dx$ zu berechnen. Mit Hilfe des Satzes von Fubini gilt

$$\int_{\mathbb{R}^2} e^{-x^2 - y^2} d\lambda^2(x, y) = \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} e^{-x^2} e^{-y^2} dx \right) dy = \int_{-\infty}^{+\infty} e^{-y^2} \left(\int_{-\infty}^{+\infty} e^{-x^2} dx \right) dy = \left(\int_{-\infty}^{+\infty} e^{-x^2} dx \right)^2.$$

Aus obiger Formel folgt andererseits mit $x^2 + y^2 = r^2$:

$$\begin{aligned} \int_{\mathbb{R}^2} e^{-x^2 - y^2} d\lambda^2(x, y) &= \int_0^\infty \int_0^{2\pi} e^{-r^2} r d\varphi dr = 2\pi \int_0^\infty e^{-r^2} r dr = 2\pi \int_0^\infty \lim_{n \rightarrow \infty} e^{-r^2} r \cdot \chi_{[0, n]}(r) dr \\ &\stackrel{\text{Levi}}{=} 2\pi \lim_{n \rightarrow \infty} \int_0^n e^{-r^2} r dr = 2\pi \lim_{n \rightarrow \infty} \frac{-e^{-r^2}}{2} \Big|_0^n = \pi \lim_{n \rightarrow \infty} (-e^{-n^2} + e^0) = \pi. \end{aligned}$$

Also erhält man

$$\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}.$$

18.9 L^p -Räume

Eine fundamentale Rolle spielen in der Physik – und hier insbesondere in der Quantenmechanik – die L^p Räume. Dabei handelt es sich um vollständige normierte Funktionenvektorräume (Banachräume), wobei die Norm mit Hilfe des Integrals definiert wird. Im Falle $p = 2$ erhält man den Hilbertraum der quadratintegrierbaren Funktionen.

Die Integrationstheorie lässt sich leicht auf komplexwertige Funktionen ausdehnen. Sei (X, \mathcal{A}, μ) ein Maßraum. Eine Funktion $f : X \rightarrow \mathbb{C}$ mit $f(x) = f_1(x) + if_2(x)$ heißt *messbar*, falls $f_1, f_2 : X \rightarrow \mathbb{R}$ messbar sind. Ist f messbar, so auch $|f| = \sqrt{f_1^2 + f_2^2}$. f heißt *integrierbar*, falls $f_1, f_2 : X \rightarrow \mathbb{R}$ integrierbar sind. Ist f integrierbar, so heißt

$$\int_X f d\mu := \int_X f_1 d\mu + i \int_X f_2 d\mu \in \mathbb{C}$$

das *Integral* von f . Seien $f, g : X \rightarrow \mathbb{C}$ integrierbar und $\alpha \in \mathbb{C}$. Dann sind $f + g$ und αf integrierbar und es gilt:

$$\alpha \int_X f d\mu = \int_X \alpha f d\mu \quad \text{und} \quad \int_X (f + g) d\mu = \int_X f d\mu + \int_X g d\mu.$$

Lemma 18.9.1. *Ist $f : X \rightarrow \mathbb{C}$ integrierbar, so auch $|f| : X \rightarrow \mathbb{R}$ und es gilt:*

$$\left| \int_X f d\mu \right| \leq \int_X |f| d\mu.$$

Beweis. Sei $f = f_1 + if_2$ integrierbar. Dann ist wegen

$$|f| = \sqrt{|f_1|^2 + |f_2|^2} \leq |f_1| + |f_2|$$

$|f|$ messbar. Da aus der Integrierbarkeit von f_1 und f_2 auch die Integrierbarkeit von $|f_1|$ und $|f_2|$ folgt, erhalten wir mit $\int_X |f| d\mu \leq \int_X |f_1| d\mu + \int_X |f_2| d\mu < \infty$ auch die Integrierbarkeit von $|f|$. Sei nun $\alpha \in \mathbb{C}$ mit $|\alpha| = 1$ und

$$\left| \int_X f d\mu \right| = \alpha \int_X f d\mu = \int_X \alpha f d\mu \in \mathbb{R}.$$

Wegen $\operatorname{Re}(z) \leq |z| \forall z \in \mathbb{C}$ folgt somit:

$$\left| \int_X f d\mu \right| = \operatorname{Re} \int_X \alpha f d\mu = \int_X \operatorname{Re}(\alpha f) d\mu \leq \int_X |f| d\mu.$$

□

Definition 18.9.2. Für $f : X \rightarrow \mathbb{C}$ messbar und $p \geq 1$ definiere

$$\|f\|_p := \left(\int_X |f|^p(x) d\mu(x) \right)^{1/p}.$$

Dann heißt

$$\mathcal{L}^p = \mathcal{L}^p(X, \mathcal{A}, \mu) = \{f : X \rightarrow \mathbb{C} \mid f \text{ messbar}, \|f\|_p < \infty\}$$

der Raum der \mathcal{L}^p -Funktionen.

Bemerkung. \mathcal{L}^p ist ein \mathbb{C} -Vektorraum, denn ist $\alpha \in \mathbb{C}$ und $f, g \in \mathcal{L}^p$, so ist $\|\alpha f\|_p = |\alpha| \cdot \|f\|_p < \infty$ und da

$$|f + g|^p \leq (|f| + |g|)^p \leq (2 \max(|f|, |g|))^p = 2^p \max(|f|^p, |g|^p) \leq 2^p (|f|^p + |g|^p),$$

ist mit $f, g \in \mathcal{L}^p$ stets auch $f + g \in \mathcal{L}^p$.

Wir wollen nun zeigen, dass $\|\cdot\|_p$ die Dreiecksungleichung erfüllt. Dazu zeigen wir zunächst die folgende fundamentale Ungleichung.

Satz 18.9.3 (Höldersche Ungleichung). *Es seien $p, q > 1$ reelle Zahlen mit*

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Sind $f, g : X \rightarrow \mathbb{C}$ messbare Funktionen, so gilt:

$$\|f \cdot g\|_1 \leq \|f\|_p \cdot \|g\|_q.$$

Bemerkung. Betrachten wir Funktionen mit Bildraum \mathbb{R} statt \mathbb{C} , so könnten wir auch zulassen, dass f den Wert ∞ annimmt.

Beweis. Wir können $0 < \|f\|_p < \infty$ und $0 < \|g\|_q < \infty$ annehmen. Denn ist $\|f\|_p = (\int_X |f|^p)^{1/p} = 0$, so ist $|f|^p = 0$ und somit auch $|f| = 0$ μ -fast überall. Dann ist auch $f = 0$ μ -fast überall und daher $\int_X |f \cdot g| d\mu = 0$. Für $\|f\|_p = \infty$ oder $\|g\|_q = \infty$ ist die Aussage klar.

Aus der Anwendung der Jensenschen Ungleichung (Übungsblatt 7, Aufgabe 4.e) folgt

$$x_1^{\lambda_1} \cdot \dots \cdot x_n^{\lambda_n} \leq \lambda_1 x_1 + \dots + \lambda_n x_n$$

für alle $x_j \geq 0$ und $\lambda_j > 0$ mit $\sum_{j=1}^n \lambda_j = 1$.

Insbesondere für $n = 2$ erhalten wir mit $\lambda_1 = \frac{1}{p}$ und $\lambda_2 = \frac{1}{q}$

$$x_1^{\frac{1}{p}} x_2^{\frac{1}{q}} \leq \frac{1}{p} x_1 + \frac{1}{q} x_2.$$

Setzen wir $x_1 = a^p$ und $x_2 = b^q$ für $a, b \geq 0$, so ergibt sich

$$a \cdot b \leq \frac{1}{p} a^p + \frac{1}{q} b^q.$$

Also erhalten wir für $a = \frac{|f(x)|}{\|f\|_p}$ und $b = \frac{|g(x)|}{\|g\|_q}$ die Abschätzung

$$\frac{|f(x)|}{\|f\|_p} \cdot \frac{|g(x)|}{\|g\|_q} \leq \frac{1}{p} \frac{|f(x)|^p}{\|f\|_p^p} + \frac{1}{q} \frac{|g(x)|^q}{\|g\|_q^q}.$$

Da $\|f\|_p^p = \int_X |f(x)|^p d\mu$ und $\|g\|_q^q = \int_X |g(x)|^q d\mu$, folgt durch Integration der obigen Ungleichung:

$$\frac{1}{\|f\|_p \|g\|_q} \int_X |f(x)| |g(x)| d\mu(x) \leq \frac{1}{p} \int_X \frac{|f(x)|^p}{\|f\|_p^p} d\mu(x) + \frac{1}{q} \int_X \frac{|g(x)|^q}{\|g\|_q^q} d\mu(x) = \frac{1}{p} + \frac{1}{q} = 1.$$

□

Bemerkung. Insbesondere erhalten wir für $p = q = 2$ die *Cauchy-Schwarzsche Ungleichung*

$$\int_X |f \cdot g| d\mu \leq \left(\int_X |f|^2 d\mu \right)^{1/2} \left(\int_X |g|^2 d\mu \right)^{1/2}.$$

Aus der Hölderschen Ungleichung lässt sich die Dreiecksungleichung (Minkowskische Ungleichung) ableiten.

Satz 18.9.4 (*Minkowskische Ungleichung*). *Es seien $f, g : X \rightarrow \mathbb{C}$ messbar und $p \geq 1$. Dann gilt:*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p.$$

Beweis. Für $p = 1$ folgt die Ungleichung aus $|f + g| \leq |f| + |g|$ durch Integration. Es sei also o.B.d.A. $p > 1$ und $\|f + g\|_p > 0$ sowie $\|f\|_p, \|g\|_p < \infty$. Betrachte $q > 1$ mit $\frac{1}{p} + \frac{1}{q} = 1$. Dann gilt:

$$\begin{aligned} \int_X |f + g|^p d\mu &= \int_X |f + g| \cdot |f + g|^{p-1} d\mu \leq \int_X |f| \cdot |f + g|^{p-1} d\mu + \int_X |g| \cdot |f + g|^{p-1} d\mu \\ &= \left\| |f| \cdot |f + g|^{p-1} \right\|_1 + \left\| |g| \cdot |f + g|^{p-1} \right\|_1 \\ &\stackrel{18.9.3}{\leq} \|f\|_p \cdot \left\| |f + g|^{p-1} \right\|_q + \|g\|_p \cdot \left\| |f + g|^{p-1} \right\|_q = (\|f\|_p + \|g\|_p) \left\| |f + g|^{p-1} \right\|_q. \end{aligned}$$

Da $\frac{1}{q} = 1 - \frac{1}{p} = \frac{p-1}{p}$, folgt

$$\begin{aligned} \left\| |f + g|^{p-1} \right\|_q &= \left(\int_X |f + g|^{(p-1)q} d\mu \right)^{1/q} = \left(\int_X |f + g|^p d\mu \right)^{\frac{p-1}{p}} \\ &= \|f + g\|_p^{p-1} \end{aligned}$$

und weiter

$$\|f + g\|_p^p = \int_X |f + g|^p d\mu \leq (\|f\|_p + \|g\|_p) \cdot \|f + g\|_p^{p-1}$$

und somit nach Division durch $\|f + g\|_p^{p-1}$ die Behauptung. □

Dies zeigt, dass $\|\cdot\|_p$ eine Halbnorm auf \mathcal{L}^p definiert.

Definition 18.9.5. Sei V ein \mathbb{C} -Vektorraum, so heißt $\|\cdot\| : V \rightarrow \mathbb{R}$ eine *Halbnorm*, falls

1. $\|f\| \geq 0$ für alle $f \in V$,
2. $\|\alpha f\| = |\alpha| \|f\|$ für alle $\alpha \in \mathbb{C}, f \in V$ und
3. $\|f + g\| \leq \|f\| + \|g\|$ für alle $f, g \in V$.

Bemerkung. Hingegen muss aus $\|f\| = 0$ nicht $f = 0$ folgen. Ist dies der Fall, so ist $\|\cdot\|$ eine Norm (siehe Definition 13.2.3).

Korollar 18.9.6. *Auf $\mathcal{L}^p(X, \mathcal{A}, \mu)$ wird durch $\|\cdot\|_p$ eine Halbnorm erklärt.*

Beweis. Folgt aus Satz 18.9.4. □

Bemerkungen.

- (a) Es ist $\|f\|_p = 0$ genau dann, falls $f = 0$ μ -fast überall.
 (b) Man kann aus \mathcal{L}^p wie folgt einen normierten Vektorraum machen:
 Sei

$$N = \{f \in \mathcal{L}^p \mid f(x) = 0 \text{ für } \mu\text{-fast alle } x \in X\}.$$

Dann ist N ein Unterraum von \mathcal{L}^p . Wir teilen nun die Funktionen in \mathcal{L}^p in *Äquivalenzklassen* ein: Wir nennen zwei Funktionen $f, g \in \mathcal{L}^p$ *äquivalent* (in Zeichen $f \sim g$), falls $f - g =: h \in N$. Ist $f \in \mathcal{L}^p$, so bezeichnen wir mit

$$[f] = \{f' \mid f' \sim f\} = \{f + h \mid h \in N\}$$

die Menge der zur Funktion f äquivalenten Funktionen. Insbesondere gilt für $f, g \in \mathcal{L}^p$

$$[f] = [g] \iff f - g = h \in N.$$

Eine Funktion in der Klasse $[f]$ heißt auch *Repräsentant* der Klasse $[f]$. Die Menge aller Äquivalenzklassen bezeichnen wir auch mit

$$L^p := \mathcal{L}^p / N.$$

Wir erklären nun eine Addition bzw. skalare Multiplikation auf L^p , indem wir Repräsentanten addieren und dann zur Äquivalenzklasse übergehen. Genauer definiere für $[f], [g] \in L^p$ und $\alpha \in \mathbb{C}$

- (i) $[f] + [g] = [f + g]$,
 (ii) $\alpha[f] = [\alpha f]$.

Diese Definitionen sind unabhängig von der Wahl der Repräsentanten, denn ist $[f'] = [f]$ und $[g'] = [g]$, so ist $f' = f + h_1$ und $g' = g + h_2$ mit $h_1, h_2 \in N$. Wegen $f' + g' = f + g + h$ mit $h = h_1 + h_2 \in N$ liegen $f' + g'$ und $f + g$ in der gleichen Klasse, d.h.

$$[f' + g'] = [f + g].$$

Genauso folgt

$$[\alpha f] = [\alpha f'],$$

falls $[f] = [f']$, denn dann ist auch $\alpha(f - f') \in N$. Das Nullelement ist durch

$$[0] = \{0 + h \mid h \in N\} = N$$

gegeben. Auch die Halbnorm $\|\cdot\|_p$ ist auf L^p definiert durch

$$\|[f]\|_p := \|f\|_p.$$

Denn ist $[f'] = [f]$, also $f' = f + h$ mit $h \in N$, so ist $\|f\|_p = \|f'\|_p$ wegen Korollar 18.5.16. Außerdem definiert $\|\cdot\|_p$ eine Norm und nicht nur eine Halbnorm auf L^p , denn ist

$$\|[f]\|_p := \|f\|_p = 0,$$

so ist $f = 0$ μ -fast überall, d.h. $[f] = 0$.

Wir werden in Zukunft die Elemente in L^p und \mathcal{L}^p meist mit den gleichen Symbolen bezeichnen.

(c) Für $p = 2$ ist die Norm $\| \cdot \|_2$ durch das Skalarprodukt

$$\langle [f], [g] \rangle = \int_X f(x) \bar{g}(x) d\mu(x),$$

mit $[f], [g] \in L^p$ induziert (wobei \bar{g} die komplex Konjugierte bezeichnet), d.h. es gilt

$$\|f\|_2 = \langle f, f \rangle^{1/2}.$$

Genauso wie im Falle der Norm, hängt die Definition des Skalarproduktes nicht von der Wahl des Repräsentanten ab (siehe 13.1.1 für die Definition des Skalarproduktes auf einem komplexen Vektorraum).

Satz 18.9.7 (Riesz-Fischer). *($L^p, \| \cdot \|_p$) ist ein vollständiger normierter Vektorraum für alle $p \geq 1$. Insbesondere ist $(L^2, \| \cdot \|_2)$ ein Hilbertraum.*

Beweis. Sei $([f_n])_{n \in \mathbb{N}}$ eine Cauchyfolge in L^p . Dann bilden auch die Repräsentanten $(f_n)_{n \in \mathbb{N}}$ eine Cauchyfolge in \mathcal{L}^p , d.h.

$$\forall \epsilon > 0 \exists n_0 \in \mathbb{N} \quad \forall n, m \geq n_0 : \|f_n - f_m\|_p < \epsilon.$$

Insbesondere existiert zu jedem $k \in \mathbb{N}$ und $\epsilon = \frac{1}{2^k}$ ein $n_k \in \mathbb{N}$ mit $\|f_{n_k} - f_m\|_p \leq \frac{1}{2^k}$ für alle $m \geq n_k$. Außerdem können wir die Folge $(n_k)_{k \in \mathbb{N}}$ streng monoton wachsend wählen. Setze $g_k := f_{n_k} - f_{n_{k+1}}$, so gilt:

$$\left\| \sum_{k=1}^{\ell} |g_k| \right\|_p \stackrel{\text{Mink.}}{\leq} \sum_{k=1}^{\ell} \|g_k\|_p \leq \sum_{k=1}^{\ell} \frac{1}{2^k} \leq 1,$$

d.h. $\int_X \left(\sum_{k=1}^{\ell} |g_k| \right)^p d\mu \leq 1$ für alle $\ell \in \mathbb{N}$. Mit dem Satz von B. Levi folgt:

$$\int_X \left(\sum_{k=1}^{\infty} |g_k| \right)^p d\mu \leq 1.$$

Also ist die Funktion $\left(\sum_{k=1}^{\infty} |g_k| \right)^p$ integrierbar und wegen Korollar 18.5.12 gilt $\sum_{k=1}^{\infty} |g_k(x)| < \infty$

für μ -fast alle $x \in X$. Insbesondere konvergiert wegen Satz 2.5.8 die Reihe $\sum_{k=1}^{\infty} g_k(x)$ für μ -fast alle $x \in X$. Da

$$\sum_{k=1}^{\ell} g_k = f_{n_1} - f_{n_{\ell+1}},$$

existiert $f(x) := \lim_{\ell \rightarrow \infty} f_{n_{\ell}}(x)$ für μ -fast alle $x \in X$. Setze diese Funktion zu einer messbaren Funktion $f : X \rightarrow \mathbb{C}$ fort (z.B. setze $f(x) = 0$, falls $\lim_{\ell \rightarrow \infty} f_{n_{\ell}}(x)$ nicht existiert).

Wir müssen noch zeigen: $\|f - f_m\|_p \rightarrow 0$ für $m \rightarrow \infty$ und $f \in \mathcal{L}^p$. Dazu benutzen wir nochmal, dass $(f_n)_{n \in \mathbb{N}}$ eine Cauchyfolge in \mathcal{L}^p darstellt. Also existiert zu jedem $\epsilon > 0$ ein $n_0 \in \mathbb{N}$ mit

$$\|f_{\ell} - f_m\|_p < \epsilon \quad \text{für alle } \ell, m \geq n_0.$$

Aus dem Lemma 18.6.1 von Fatou folgt:

$$\int_X |f - f_m|^p d\mu = \int_X \lim_{k \rightarrow \infty} |f_{n_k} - f_m|^p d\mu \leq \liminf_{k \rightarrow \infty} \int_X |f_{n_k} - f_m|^p d\mu = \liminf_{k \rightarrow \infty} \|f_{n_k} - f_m\|_p^p \leq \epsilon^p$$

und somit

$$\|f - f_m\|_p \leq \epsilon \text{ für alle } m \geq n_0. \quad (18.13)$$

Aus der Dreiecksungleichung folgt:

$$\|f\|_p = \|f - f_m + f_m\|_p \leq \|f - f_m\|_p + \|f_m\|_p < \infty.$$

Also ist $f \in \mathcal{L}^p$ und $[f] \in L^p$. Aus (18.13) erhalten wir

$$\lim_{m \rightarrow \infty} \|[f] - [f_m]\|_p = \lim_{m \rightarrow \infty} \|[f - f_m]\|_p = \lim_{m \rightarrow \infty} \|f - f_m\|_p = 0$$

und somit ist der normierte Vektorraum $(L^p, \|\cdot\|_p)$ vollständig für alle $p \geq 1$. \square

Bemerkung. Der Beweis des Satzes von Riesz- Fischer zeigt auch: ist $(f_n)_{n \in \mathbb{N}}$ eine Cauchyfolge in \mathcal{L}^p , so existiert eine Teilfolge f_{n_ℓ} und eine Funktion $f \in \mathcal{L}^p$ mit

$$f(x) := \lim_{\ell \rightarrow \infty} f_{n_\ell}(x)$$

für μ -fast alle $x \in X$. Ist insbesondere $(f_n)_{n \in \mathbb{N}}$ eine Folge in \mathcal{L}^p und $f \in \mathcal{L}^p$ mit

$$\lim_{n \rightarrow \infty} \|f_n - f\|_p = 0,$$

so existiert eine Teilfolge f_{n_ℓ} mit

$$f(x) := \lim_{\ell \rightarrow \infty} f_{n_\ell}(x)$$

für μ -fast alle $x \in X$. Wie das folgende Beispiel zeigt, ist es im allgemeinen notwendig, für die punktweise Konvergenz eine Teilfolge zu betrachten.

Beispiel. Sei $X = [0, 1]$ und λ das Lebesguemaß auf $[0, 1]$. Betrachte die Folge der Intervalle

$$[0, 1], \left[0, \frac{1}{2}\right], \left[\frac{1}{2}, 1\right], \left[0, \frac{1}{3}\right], \left[\frac{1}{3}, \frac{2}{3}\right], \left[\frac{2}{3}, 1\right], \left[0, \frac{1}{4}\right], \left[\frac{1}{4}, \frac{2}{4}\right], \dots$$

Sei $I_n \subset I$ das n -te Intervall dieser Aufzählung, so besitzt für kein $x \in [0, 1]$ die Folge $\chi_{I_n}(x)$ einen Grenzwert. Denn für jeweils unendlich viele $n \in \mathbb{N}$ hat die Folge den Wert 1 und 0. Auf der anderen Seite gilt

$$\|\chi_{I_n}\|_p = \lambda(I_n) \rightarrow 0 \text{ für } n \rightarrow \infty.$$

Kapitel 19

Differentialformen und Integralsätze

Ziel dieses Kapitels ist eine Verallgemeinerung des Kurvenintegrals, das wir in Abschnitt 17.2 behandelt haben. Es ermöglicht uns, Integrale nicht nur über Kurven, sondern auch über höherdimensionale Objekte, wie z.B. Untermannigfaltigkeiten, zu definieren.

19.1 Differentialformen

In Abschnitt 17.2 haben wir 1-Formen über Kurven integriert. Statt über Kurven wollen wir nun auch über k -dimensionale Untermannigfaltigkeiten integrieren. Die geeigneten Integranden werden Differentialformen von Grade k sein. Diese ordnen jedem $x \in U \subset \mathbb{R}^n$ eine alternierende k -Form zu.

Definition 19.1.1. Sei V ein Vektorraum der Dimension n über \mathbb{R} und $V^k := \underbrace{V \times \dots \times V}_{k\text{-mal}}$.

Eine Abbildung

$$\alpha : V^k \rightarrow \mathbb{R}$$

heißt *alternierende* oder *schiefssymmetrische k -Form*, falls

- (i) sie *multilinear* ist, d.h. für jedes $1 \leq j \leq k$

$$\alpha(\dots, \lambda v_j + \mu v'_j, \dots) = \lambda \alpha(\dots, v_j, \dots) + \mu \alpha(\dots, v'_j, \dots)$$

gilt, wobei $\lambda, \mu \in \mathbb{R}$, $v_j, v'_j \in V$ und alle anderen Variablen festgehalten werden;

- (ii) sie *alternierend* ist, d.h.

$$\alpha(\dots, v_i, \dots, v_j, \dots) = -\alpha(\dots, v_j, \dots, v_i, \dots)$$

für alle $v_i, v_j \in V$.

Der Dualraum $V^* = L(V, \mathbb{R})$ zu V ist der Vektorraum aller alternierenden 1-Formen. Die Menge aller k -Formen auf V bezeichnen wir mit $\Lambda^k(V^*)$. Insbesondere ist $\Lambda^1 V^* = V^*$. Wir setzen $\Lambda^0(V^*) := \mathbb{R}$.

Bemerkungen.

- (a) $\Lambda^k(V^*)$ ist ein Vektorraum über \mathbb{R} . Ist $\alpha \in \Lambda^k(V^*)$ und S_k (siehe Seite 205) die Permutationsgruppe der Menge $\{1, \dots, k\}$, so gilt für jedes $\sigma \in S_k$

$$\alpha(v_{\sigma(1)}, \dots, v_{\sigma(k)}) = \text{sign } \sigma \alpha(v_1, \dots, v_k),$$

wobei $\text{sign } \sigma$ das Vorzeichen der Permutation bezeichnet.

- (b) Die Determinante ist eine alternierende n -Form auf \mathbb{R}^n und somit ist $\det \in \Lambda^n((\mathbb{R}^n)^*)$.
 (c) Ist $v \in \mathbb{R}^3$ und ist $\alpha : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}$ definiert durch

$$\alpha(v_1, v_2) = \det(v, v_1, v_2),$$

so ist $\alpha \in \Lambda^2((\mathbb{R}^3)^*)$

Physikalische Interpretation von α : Beschreibt v das Geschwindigkeitsfeld einer konstanten Strömung, so lässt sich $\alpha(v_1, v_2)$ als der Fluss durch das durch v_1, v_2 aufgespannte, orientierte Parallelogramm ansehen. Die Reihenfolge der Vektoren v_1, v_2 gibt die Orientierung des Parallelogrammes an. Sie bestimmt, ob die Strömung das Parallelogramm in positiver oder negativer Richtung durchläuft.

- (d) Bedingung (ii) ist äquivalent zu

$$\alpha(\dots, v, \dots, v, \dots) = 0$$

für alle $v \in V$, wie man durch Anwendung von (i) für $v = v_i + v_j$ nachrechnet.

Definition 19.1.2 (*Äußeres Produkt (Dachprodukt) von 1-Formen*). Es seien V ein reeller Vektorraum und $\alpha_1, \dots, \alpha_k \in \Lambda^1(V^*) = V^*$ 1-Formen. Dann heißt

$$\alpha_1 \wedge \dots \wedge \alpha_k : V^k \rightarrow \mathbb{R}$$

mit

$$\alpha_1 \wedge \dots \wedge \alpha_k(v_1, \dots, v_k) = \det \begin{pmatrix} \alpha_1(v_1) & \dots & \alpha_1(v_k) \\ \vdots & & \vdots \\ \alpha_k(v_1) & \dots & \alpha_k(v_k) \end{pmatrix}$$

das *äußere Produkt* oder *Dachprodukt* des k -Tupels von 1-Formen $(\alpha_1, \dots, \alpha_k)$.

Bemerkungen.

- (a) Aus der Definition der Determinante folgt:

$$\alpha_1 \wedge \dots \wedge \alpha_k \in \Lambda^k(V^*).$$

- (b) Betrachte die lineare Abbildung $L : V \rightarrow \mathbb{R}^k$ mit $L(v) = (\alpha_1(v), \dots, \alpha_k(v))$. Dann beschreibt $\alpha_1 \wedge \dots \wedge \alpha_k(v_1, \dots, v_k)$ das "orientierte" Volumen des durch die Vektoren

$$L(v_1), \dots, L(v_k) \text{ aufgespannten Parallelotops } P = \left\{ \sum_{i=1}^k t_i L(v_i) \mid 0 \leq t_i \leq 1 \right\}.$$

- (c) Das Dachprodukt besitzt folgende Eigenschaften:

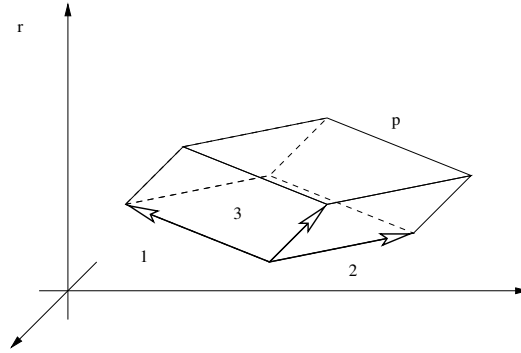


Abbildung 19.1: Parallelotop im \mathbb{R}^3

(i) Seien $\alpha'_i, \alpha_1, \dots, \alpha_k \in \Lambda^1(V^*) = V^*$ 1-Formen und $\lambda, \mu \in \mathbb{R}$, so gilt

$$\alpha_1 \wedge \dots \wedge (\lambda \alpha_i + \mu \alpha'_i) \wedge \dots \wedge \alpha_k = \lambda \alpha_1 \wedge \dots \wedge \alpha_i \wedge \dots \wedge \alpha_k + \mu \alpha_1 \wedge \dots \wedge \alpha'_i \wedge \dots \wedge \alpha_k.$$

(ii) Ist $\sigma \in S_k$, so gilt

$$\alpha_{\sigma(1)} \wedge \dots \wedge \alpha_{\sigma(k)} = \text{sign } \sigma \alpha_1 \wedge \dots \wedge \alpha_k.$$

Sei $e_1, \dots, e_n \in V$ eine Basis von V und $e_j^* \in V^*$ die Linearform definiert durch $e_j^*(e_i) = \delta_{ij}$, so sind e_1^*, \dots, e_n^* eine Basis von V^* (siehe Satz 8.3.4). Sie heißt die zu e_1, \dots, e_n *duale Basis*. Mit Hilfe der dualen Basis erhält man wie folgt eine Basis für $\Lambda^k(V^*)$:

Satz 19.1.3. Sei $e_1, \dots, e_n \in V$ Basis von V und $e_1^*, \dots, e_n^* \in \Lambda^1(V^*)$ die duale Basis. Dann ist

$$\{e_{i_1}^* \wedge \dots \wedge e_{i_k}^* \mid 1 \leq i_1 < i_2 < \dots < i_k \leq n\}$$

eine Basis von $\Lambda^k(V^*)$.

Beweis. Die Vektoren $e_{i_1}^* \wedge \dots \wedge e_{i_k}^* \in \Lambda^k(V^*)$ sind linear unabhängig: Dazu nehme man an, es sei

$$\alpha := \sum_{1 \leq i_1 < \dots < i_k \leq n} a_{i_1 \dots i_k} e_{i_1}^* \wedge \dots \wedge e_{i_k}^* = 0 \tag{19.1}$$

für gewisse $a_{i_1 \dots i_k} \in \mathbb{R}$.

Für das geordnete k -Tupel $(e_{j_1}, \dots, e_{j_k}) \in V^k$ mit $1 \leq j_1 < j_2 < \dots < j_k \leq n$ gilt:

$$e_{i_1}^* \wedge \dots \wedge e_{i_k}^*(e_{j_1}, \dots, e_{j_k}) = \det \begin{pmatrix} e_{i_1}^*(e_{j_1}) & \dots & e_{i_1}^*(e_{j_k}) \\ \vdots & & \vdots \\ e_{i_k}^*(e_{j_1}) & \dots & e_{i_k}^*(e_{j_k}) \end{pmatrix} = \begin{cases} 1, & \text{falls } i_1 = j_1, \dots, i_k = j_k \\ 0, & \text{sonst.} \end{cases}$$

Denn ist die Determinante ungleich null, so ist die erste Spalte der obigen Matrix von null verschieden. Daher existiert ein $\ell \in \{1, \dots, k\}$ mit $e_{i_\ell}^*(e_{j_1}) = 1$, d.h. $i_\ell = j_1$. Da die erste Zeile ebenfalls von null verschieden ist, muss $\ell = 1$ gelten. Denn wäre $\ell \geq 2$, so ist $i_1 < i_\ell = j_1$ und somit ist $i_1 \neq j_m$ für alle $m \in \{1, \dots, k\}$, d.h. $e_{i_1}^*(e_{j_m}) = 0$. Also ist $i_1 = j_1$. Genauso folgt:

$$i_2 = j_2, \dots, i_m = j_m.$$

Wenden wir nun α aus (19.1) auf $(e_{j_1}, \dots, e_{j_k}) \in V^k$ mit $1 \leq j_1 < j_2 < \dots < j_k \leq n$ an, so folgt also $a_{j_1, \dots, j_k} = 0$.

Auf der anderen Seite spannen die Vektoren

$$\{e_{i_1}^* \wedge \dots \wedge e_{i_k}^* \mid 1 \leq i_1 < i_2 < \dots < i_k \leq n\}$$

den Vektorraum $\Lambda^k(V^*)$ auf. Ist nämlich $\alpha \in \Lambda^k(V^*)$, so folgt:

$$\alpha = \sum_{1 \leq i_1 < \dots < i_k \leq n} \alpha(e_{i_1}, \dots, e_{i_k}) e_{i_1}^* \wedge \dots \wedge e_{i_k}^*,$$

denn beide Seiten stimmen auf allen k -Tupeln $(e_{j_1}, \dots, e_{j_k}) \in V^k$ mit $1 \leq j_1 < \dots < j_k \leq n$ überein. Da durch die Werte auf den geordneten k -Tupeln $(e_{j_1}, \dots, e_{j_k}) \in V^k$ eine k -Form eindeutig bestimmt wird (Beweis!), folgt die Behauptung. \square

Bemerkung. Ist $\dim V = n$, so folgt $\dim \Lambda^k(V^*) = \binom{n}{k}$. Insbesondere ist $\dim \Lambda^n(V^*) = 1$ und $\dim \Lambda^k(V^*) = 0$ für $k > n$.

Das äußere Produkt von 1-Formen lässt sich wie folgt zu einem äußeren Produkt von k -Formen mit ℓ -Formen ausdehnen:

Satz 19.1.4. *Es gibt genau eine Abbildung*

$$\begin{aligned} \Lambda^k(V^*) \times \Lambda^\ell(V^*) &\rightarrow \Lambda^{k+\ell}(V^*), \\ (\alpha, \beta) &\mapsto \alpha \wedge \beta \end{aligned}$$

mit folgenden Eigenschaften:

(i) $\alpha \wedge \beta$ ist linear in jedem Faktor, d.h. für alle $\alpha_1, \alpha_2, \alpha \in \Lambda^k(V^*)$ und $\beta_1, \beta_2, \beta \in \Lambda^\ell(V^*)$, sowie $\lambda \in \mathbb{R}$ gilt:

$$(\alpha_1 + \alpha_2) \wedge \beta = \alpha_1 \wedge \beta + \alpha_2 \wedge \beta \quad \text{und} \quad \alpha \wedge (\beta_1 + \beta_2) = \alpha \wedge \beta_1 + \alpha \wedge \beta_2$$

sowie

$$\lambda(\alpha \wedge \beta) = (\lambda\alpha) \wedge \beta = \alpha \wedge (\lambda\beta).$$

(ii) Sind $\alpha_1, \dots, \alpha_k, \beta_1, \dots, \beta_\ell \in \Lambda^1(V^*)$, so gilt:

$$(\alpha_1 \wedge \dots \wedge \alpha_k) \wedge (\beta_1 \wedge \dots \wedge \beta_\ell) = \alpha_1 \wedge \dots \wedge \alpha_k \wedge \beta_1 \wedge \dots \wedge \beta_\ell.$$

Beweis.

(a) Existenz: Sei $e_1, \dots, e_n \in V$ eine Basis und $e_1^*, \dots, e_n^* \in \Lambda^1(V^*) = V^*$ die duale Basis. Seien

$$\alpha = \sum_{i_1 < \dots < i_k} a_{i_1 \dots i_k} e_{i_1}^* \wedge \dots \wedge e_{i_k}^* \in \Lambda^k(V^*) \quad \text{und} \quad \beta = \sum_{j_1 < \dots < j_\ell} b_{j_1, \dots, j_\ell} e_{j_1}^* \wedge \dots \wedge e_{j_\ell}^* \in \Lambda^\ell(V^*),$$

so definiere

$$\alpha \wedge \beta := \sum_{\substack{i_1 < \dots < i_k \\ j_1 < \dots < j_\ell}} a_{i_1 \dots i_k} b_{j_1, \dots, j_\ell} e_{i_1}^* \wedge \dots \wedge e_{i_k}^* \wedge e_{j_1}^* \wedge \dots \wedge e_{j_\ell}^*. \quad (19.2)$$

Diese Definition erfüllt die verlangten Eigenschaften, was man durch Nachrechnen überprüft.

- (b) Eindeutigkeit: Ist \wedge ein Dachprodukt mit den Eigenschaften (i) und (ii), so ist " \wedge " nach Wahl einer Basis $e_1, \dots, e_n \in V$ von der Form (19.2).

□

Bemerkung.

- (a) Wir lassen auch den Fall $k = 0$ bzw. $\ell = 0$ zu. Für $\lambda \in \mathbb{R} = \Lambda^0(\mathbb{R})$ und $\alpha \in \Lambda^k(V)$ setze

$$\lambda \wedge \alpha := \alpha \wedge \lambda := \lambda \cdot \alpha.$$

- (b) Es gelten folgende weitere Rechenregeln für $\alpha \in \Lambda^k(V^*)$, $\beta \in \Lambda^\ell(V^*)$ und $\gamma \in \Lambda^m(V^*)$:

$$\alpha \wedge (\beta \wedge \gamma) = (\alpha \wedge \beta) \wedge \gamma$$

und

$$\alpha \wedge \beta = (-1)^{k \cdot \ell} \beta \wedge \alpha.$$

Der Beweis sei als Übung überlassen.

Nun wollen wir Differentialformen auf offenen Teilmengen des \mathbb{R}^n definieren.

Definition 19.1.5. Sei $U \subset \mathbb{R}^n$ offen. Eine Abbildung $\omega : U \rightarrow \Lambda^k((\mathbb{R}^n)^*)$ heißt *Differentialform vom Grade k (k -Form)*. Eine k -Form $\omega : U \rightarrow \Lambda^k((\mathbb{R}^n)^*)$ heißt *m -mal stetig differenzierbar*, falls für alle Vektoren $v_1, \dots, v_k \in \mathbb{R}^n$ die reellwertige Abbildung

$$p \mapsto \omega(p)(v_1, \dots, v_k)$$

auf U m -mal stetig differenzierbar ist. Wir bezeichnen mit $\Omega_m^k(U)$ die Menge der m -mal stetig differenzierbaren k -Formen auf U . Die Menge aller unendlich oft differenzierbaren k -Formen auf U wird mit $\Omega^k(U)$ bezeichnet.

Bemerkungen.

- (a) Da $\Lambda^0((\mathbb{R}^n)^*) = \mathbb{R}$, entspricht $\Omega_m^0(U)$ der Menge der m -mal stetig differenzierbaren Funktionen $C^m(U)$ auf U und $\Omega^0(U)$ der Menge der unendlich oft differenzierbaren Funktionen $C^\infty(U)$.
- (b) Sind $x_i : U \rightarrow \mathbb{R}$ die kanonischen Koordinatenfunktionen auf $U \subset \mathbb{R}^n$ und e_1, \dots, e_n die kanonische Basis des \mathbb{R}^n , so folgt

$$dx_i(p)e_j = \frac{\partial x_i}{\partial x_j}(p) = \delta_{ij}.$$

Insbesondere ist $dx_1(p), \dots, dx_n(p)$ für jedes $p \in U$ die zu e_1, \dots, e_n duale Basis von $(\mathbb{R}^n)^* = \Lambda^1((\mathbb{R}^n)^*)$.

Ist ω eine k -Form, so existieren wegen Satz 19.1.3 Funktionen $a_{i_1 \dots i_k} : U \rightarrow \mathbb{R}$ mit

$$\omega(p) = \sum_{1 \leq i_1 < \dots < i_k \leq n} a_{i_1 \dots i_k}(p) dx_{i_1}(p) \wedge \dots \wedge dx_{i_k}(p).$$

Wegen $\omega(p)(e_{i_1}, \dots, e_{i_k}) = a_{i_1 \dots i_k}(p)$ ist ω genau dann m -mal stetig differenzierbar, falls alle Koeffizienten $a_{i_1 \dots i_k} : U \rightarrow \mathbb{R}$ m -mal stetig differenzierbar sind.

Definition 19.1.6. Sei $U \subset \mathbb{R}^n$ offen und ω, η k -Formen auf U und $\lambda \in \mathbb{R}$. Dann definiere das Produkt $\lambda\omega : U \rightarrow \Lambda^k((\mathbb{R}^n)^*)$ und die Summe $\omega + \eta : U \rightarrow \Lambda^k((\mathbb{R}^n)^*)$ durch

$$(\lambda\omega)(p) = \lambda\omega(p) \quad \text{und} \quad (\omega + \eta)(p) = \omega(p) + \eta(p).$$

Sind ω eine k - und η eine ℓ -Form, so definiere ihr Dachprodukt $\omega \wedge \eta : U \rightarrow \Lambda^{k+\ell}((\mathbb{R}^n)^*)$ durch

$$(\omega \wedge \eta)(p) = \omega(p) \wedge \eta(p).$$

Bemerkung. Sind $\omega, \eta \in \Omega_m^k(U)$ und $\lambda \in \mathbb{R}$, so ist auch $\lambda\omega + \eta \in \Omega_m^k(U)$, d.h. $\Omega_m^k(U)$ und $\Omega^k(U)$ sind bezüglich der Addition und skalaren Multiplikation Vektorräume über \mathbb{R} . Sind $\omega \in \Omega_m^k(U), \eta \in \Omega_m^\ell(U)$, so ist $\omega \wedge \eta \in \Omega_m^{k+\ell}(U)$.

Das Differential von k -Formen kann wie folgt definiert werden:

Definition 19.1.7. Sei $\omega \in \Omega_1^k(U)$ gegeben durch

$$\omega = \sum_{1 \leq i_1 < \dots < i_k \leq n} a_{i_1 \dots i_k} dx_{i_1} \wedge \dots \wedge dx_{i_k}.$$

Dann heißt

$$d\omega := \sum_{1 \leq i_1 < \dots < i_k \leq n} da_{i_1 \dots i_k} \wedge dx_{i_1} \wedge \dots \wedge dx_{i_k}$$

die *äußere Ableitung* oder das *Differential* von ω .

Bemerkungen.

(a) $d\omega$ ist eine $k+1$ -Form. Ist $m \geq 1$ und $\omega \in \Omega_m^k(U)$, so ist $d\omega \in \Omega_{m-1}^{k+1}(U)$.

(b) Ist $\omega = f : U \rightarrow \mathbb{R}$ stetig differenzierbar, dann ist $d\omega = df$ eine 1-Form auf U . In Kapitel 15 hatten wir das Differential einer Funktion f mit Df bezeichnet. Um eine einheitliche Schreibweise zu gewährleisten, bevorzugen wir im Kontext der Differentialformen die Bezeichnung df für \mathbb{R} -wertige Funktionen.

(c) Ist $\omega \in \Omega_1^1(U)$ mit $\omega(x) = \sum_{i=1}^n a_i(x) dx_i$, so ist $da_i = \sum_{j=1}^n \frac{\partial}{\partial x_j} a_i dx_j$ und somit

$$\begin{aligned} d\omega &= \sum_{i,j=1}^n \frac{\partial}{\partial x_j} a_i dx_j \wedge dx_i = \sum_{i < j} \frac{\partial}{\partial x_j} a_i dx_j \wedge dx_i + \sum_{j < i} \frac{\partial}{\partial x_j} a_i dx_j \wedge dx_i \\ &= -\sum_{i < j} \frac{\partial}{\partial x_j} a_i dx_i \wedge dx_j + \sum_{j < i} \frac{\partial}{\partial x_j} a_i dx_j \wedge dx_i \\ &= \sum_{i < j} \frac{\partial}{\partial x_i} a_j dx_i \wedge dx_j - \sum_{i < j} \frac{\partial}{\partial x_j} a_i dx_i \wedge dx_j \quad (\text{durch Umbenennung von } i \text{ und } j) \\ &= \sum_{i < j} \left(\frac{\partial}{\partial x_i} a_j - \frac{\partial}{\partial x_j} a_i \right) dx_i \wedge dx_j. \end{aligned}$$

Insbesondere ist für eine zweimal stetig differenzierbare Funktion $f : U \rightarrow \mathbb{R}$ mit $\omega = df = \sum_{i=1}^n \frac{\partial f}{\partial x_i} dx_i$:

$$ddf = \sum_{i < j} \left(\frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} f - \frac{\partial}{\partial x_j} \frac{\partial}{\partial x_i} f \right) dx_i \wedge dx_j = 0.$$

- (d) Sei $U \subset \mathbb{R}^n$ und $\omega : U \rightarrow \Lambda^k((\mathbb{R}^n)^*)$ eine differenzierbare k -Form. Ihre äußere Ableitung $d\omega$ lässt sich auch koordinatenfrei definieren. Für Vektoren $v_1, \dots, v_k \in \mathbb{R}^n$ betrachte die reellwertige Abbildung

$$p \mapsto \omega(p)(v_1, \dots, v_k).$$

Definiere nun

$$\eta(p)(v_1, \dots, v_{k+1}) = \sum_{i=1}^{k+1} (-1)^{i-1} D(\omega(v_1, \dots, \hat{v}_i, \dots, v_{k+1}))(p)(v_i),$$

wobei \hat{v}_i bedeutet, dass v_i ausgelassen werden soll. Dann gilt: η ist eine $k+1$ -Form und $d\omega = \eta$. Überprüfen Sie diesen Sachverhalt für $k = 1$.

Satz 19.1.8. (a) Seien $\omega, \eta \in \Omega_1^k(U)$ und $\lambda, \mu \in \mathbb{R}$, so ist

$$d(\lambda\omega + \mu\eta) = \lambda d\omega + \mu d\eta.$$

Insbesondere definiert für $m \geq 1$ die äußere Ableitung $d : \Omega_m^k(U) \rightarrow \Omega_{m-1}^{k+1}(U)$ einen linearen Operator.

- (b) Ist $\omega \in \Omega_1^k(U)$ und $\eta \in \Omega_1^\ell(U)$, so ist

$$d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^k \omega \wedge d\eta \in \Omega_0^{k+\ell}(U).$$

- (c) Ist $\omega \in \Omega_2^k(U)$, so ist

$$dd\omega = 0.$$

Beweis.

- (a) ist trivial.

- (b) Wegen der Linearität der äußeren Ableitung reicht es aus, (b) für ω, η von der Form

$$\omega = f dx_{i_1} \wedge \dots \wedge dx_{i_k} =: f \cdot dx_I \quad \text{und} \quad \eta = g dx_{j_1} \wedge \dots \wedge dx_{j_\ell} =: g \cdot dx_J$$

mit $f, g \in C^1(U)$ zu beweisen (solche Differentialformen heißen auch *Monome*):

$$\begin{aligned} d(\omega \wedge \eta) &= d(fg dx_I \wedge dx_J) = (g \cdot df + f \cdot dg) \wedge dx_I \wedge dx_J \\ &= (df \wedge dx_I) \wedge (g dx_J) + (-1)^k (f dx_I) \wedge (dg \wedge dx_J) \\ &= d\omega \wedge \eta + (-1)^k \omega \wedge d\eta. \end{aligned}$$

- (c) Wegen der Linearität von d genügt es wieder, ω von der Form

$$\omega = f dx_{i_1} \wedge \dots \wedge dx_{i_k} =: f dx_I$$

zu betrachten. Dann ist $d\omega = df \wedge dx_I$ und wegen Teil (b) folgt mit Bemerkung (c) nach Definition 19.1.7:

$$dd\omega = ddf \wedge dx_I - df \wedge d(1 \cdot dx_I) = ddf \wedge dx_I - df \wedge d(1) \wedge dx_I = 0.$$

□

Beispiel. In diesem Beispiel wollen wir Bezüge zu der in der Physik für \mathbb{R}^3 gebräuchlichen Vektoranalysis aufzeigen. Sei $U \subset \mathbb{R}^3$ eine offene Menge und $F : U \rightarrow \mathbb{R}^3$ ein stetig differenzierbares Vektorfeld. Dann können wir F eine 1-Form $\omega_F^1 \in \Omega_1^1(U)$ und eine 2-Form $\omega_F^2 \in \Omega_1^2(U)$ zuordnen: Setze

$$\omega_F^1(p)(v) = \langle F(p), v \rangle \quad \text{und} \quad \omega_F^2(p)(v_1, v_2) := \det(F(p), v_1, v_2) = \langle F(p), (v_1 \times v_2) \rangle$$

mit dem Kreuzprodukt \times in \mathbb{R}^3 . Sind $F = (F_1, F_2, F_3)$ die Komponenten des Vektorfeldes, so gilt:

$$\omega_F^1 = F_1 dx_1 + F_2 dx_2 + F_3 dx_3 \quad \text{sowie} \quad \omega_F^2 = F_1 dx_2 \wedge dx_3 + F_2 dx_3 \wedge dx_1 + F_3 dx_1 \wedge dx_2.$$

Umgekehrt lässt sich jeder 1-Form und 2-Form auf $U \subset \mathbb{R}^3$ eine Vektorfeld zuordnen.

Man rechnet nun leicht nach:

(a) $df = \omega_{\text{grad } f}^1$ für jede C^1 -Funktion $f : U \rightarrow \mathbb{R}$,

(b) $d\omega_F^1 = \omega_{\text{rot } F}^2$,

(c) $d\omega_F^2 = \text{div } F dx_1 \wedge dx_2 \wedge dx_3$.

Dabei folgt (a) aus $df(p)(v) = \langle \text{grad } f(p), v \rangle \forall v \in \mathbb{R}^3$. Beweis zu (b):

$$\begin{aligned} d\omega_F^1 &= \frac{\partial F_1}{\partial x_2} dx_2 \wedge dx_1 + \frac{\partial F_1}{\partial x_3} dx_3 \wedge dx_1 + \frac{\partial F_2}{\partial x_1} dx_1 \wedge dx_2 + \frac{\partial F_2}{\partial x_3} dx_3 \wedge dx_2 \\ &\quad + \frac{\partial F_3}{\partial x_1} dx_1 \wedge dx_3 + \frac{\partial F_3}{\partial x_2} dx_2 \wedge dx_3 \\ &= \left(\frac{\partial F_3}{\partial x_2} - \frac{\partial F_2}{\partial x_3} \right) dx_2 \wedge dx_3 + \left(\frac{\partial F_1}{\partial x_3} - \frac{\partial F_3}{\partial x_1} \right) dx_3 \wedge dx_1 + \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \right) dx_1 \wedge dx_2 \end{aligned}$$

Beweis zu (c):

$$\begin{aligned} d\omega_F^2 &= \frac{\partial F_1}{\partial x_1} dx_1 \wedge dx_2 \wedge dx_3 + \frac{\partial F_2}{\partial x_2} dx_2 \wedge dx_3 \wedge dx_1 + \frac{\partial F_3}{\partial x_3} dx_3 \wedge dx_1 \wedge dx_2 \\ &= \left(\frac{\partial F_1}{\partial x_1} + \frac{\partial F_2}{\partial x_2} + \frac{\partial F_3}{\partial x_3} \right) dx_1 \wedge dx_2 \wedge dx_3. \end{aligned}$$

Da $dd\omega = 0$, so folgt für jede C^2 -Funktion $f : U \rightarrow \mathbb{R}$:

$$0 =ddf = d\omega_{\text{grad } f}^1 = \omega_{\text{rot grad } f}^2$$

und somit $\text{rot grad } f = 0$. Außerdem folgt für jedes C^2 -Vektorfeld $F : U \rightarrow \mathbb{R}^3$

$$0 = dd\omega_F^1 = d\omega_{\text{rot } F}^2 = \text{div rot } F dx_1 \wedge dx_2 \wedge dx_3$$

und somit $\text{div rot } F = 0$.

Vektorfelder und 1-Formen lassen sich in allen Dimensionen identifizieren. Ab Dimension 4 existieren jedoch keine Beziehungen zwischen Vektorfeldern und 2-Formen. Ist $U \subset \mathbb{R}^4$ offen und $\omega \in \Omega_1^2(U)$, so benötigt man eine Abbildung $F : U \rightarrow \mathbb{R}^6$, um ω darzustellen. Da 2-Formen auf $U \subset \mathbb{R}^4$ in der relativistischen Elektrodynamik eine wichtige Rolle spielen, führt man, wenn man die Sprache der Differentialformen umgehen will, so genannte ‘‘Sechser’’-Vektoren ein.

19.2 Zurückholen (pull-back) von Differentialformen

Wir wollen nun das Integral von k -Formen definieren. Dazu benötigen wir die Operation des Zurückholens von Differentialformen.

Definition 19.2.1. Sei $U \subset \mathbb{R}^n$ und $V \subset \mathbb{R}^m$ offen, $\varphi : V \rightarrow U$ stetig differenzierbar und ω eine k -Form auf U . Dann heißt die auf V durch

$$(\varphi^*\omega)(p)(v_1, \dots, v_k) := \omega(\varphi(p))(D\varphi(p)(v_1), \dots, D\varphi(p)(v_k)), \quad p \in V, v_1, \dots, v_k \in \mathbb{R}^m$$

definierte k -Form die mittels φ zurückgeholte k -Form und wird mit $\varphi^*\omega$ notiert. Ist ω eine 0-Form, d.h. eine Funktion f auf U , so ist

$$(\varphi^*\omega)(p) = \varphi^*f(p) = f \circ \varphi(p).$$

Es gelten folgende Rechenregeln:

Lemma 19.2.2. Seien $V \subset \mathbb{R}^m, U \subset \mathbb{R}^n$ offen und $\varphi : V \rightarrow U$ eine C^1 -Abbildung. Dann gilt:

(a) Sind ω_1, ω_2 k -Formen auf U und $f : U \rightarrow \mathbb{R}$ eine Funktion, so folgt

$$\varphi^*(\omega_1 + \omega_2) = \varphi^*\omega_1 + \varphi^*\omega_2$$

und

$$\varphi^*(f\omega_1) = (f \circ \varphi)\varphi^*(\omega_1) = \varphi^*f \wedge \varphi^*(\omega_1).$$

(b) Sind $f_1, \dots, f_k : U \rightarrow \mathbb{R}$ C^1 -Funktionen, so gilt:

$$\varphi^*(df_1 \wedge \dots \wedge df_k) = d(f_1 \circ \varphi) \wedge \dots \wedge d(f_k \circ \varphi).$$

(c) Ist ω eine k -Form auf U und η eine ℓ -Form auf U , so gilt:

$$\varphi^*(\omega \wedge \eta) = \varphi^*\omega \wedge \varphi^*\eta.$$

(d) Ist $W \subset \mathbb{R}^\ell$ offen, $\psi : W \rightarrow V$ eine C^1 -Abbildung und ω eine k -Form auf U , so gilt:

$$(\varphi \circ \psi)^*\omega = \psi^*(\varphi^*\omega).$$

Beweis.

(a) folgt unmittelbar aus der Definition von φ^* .

(b) Betrachte Vektoren $v_1, \dots, v_k \in \mathbb{R}^m$ und $p \in V$. Dann gilt:

$$\begin{aligned} \varphi^*(df_1 \wedge \dots \wedge df_k)(p)(v_1, \dots, v_k) &= (df_1 \wedge \dots \wedge df_k)(\varphi(p))(D\varphi(p)(v_1), \dots, D\varphi(p)(v_k)) \\ &= df_1(\varphi(p)) \wedge \dots \wedge df_k(\varphi(p))(D\varphi(p)(v_1), \dots, D\varphi(p)(v_k)) \\ &= \det \begin{pmatrix} df_1(\varphi(p))D\varphi(p)(v_1) & \cdots & df_1(\varphi(p))D\varphi(p)(v_k) \\ \vdots & & \vdots \\ df_k(\varphi(p))D\varphi(p)(v_1) & \cdots & df_k(\varphi(p))D\varphi(p)(v_k) \end{pmatrix} \\ &= \det \begin{pmatrix} d(f_1 \circ \varphi)(p)(v_1) & \cdots & d(f_1 \circ \varphi)(p)(v_k) \\ \vdots & & \vdots \\ d(f_k \circ \varphi)(p)(v_1) & \cdots & d(f_k \circ \varphi)(p)(v_k) \end{pmatrix} \\ &= d(f_1 \circ \varphi)(p) \wedge \dots \wedge d(f_k \circ \varphi)(p)(v_1, \dots, v_k) \\ &= (d(f_1 \circ \varphi) \wedge \dots \wedge d(f_k \circ \varphi))(p)(v_1, \dots, v_k). \end{aligned}$$

(c) Ist $k = 0$, d.h. $\omega = f$ eine Funktion und η eine ℓ -Form auf U , so gilt:

$$\varphi^*(f\eta) = (f \circ \varphi)\varphi^*\eta = \varphi^*f \wedge \varphi^*\eta.$$

Wegen (a) genügt es, (c) für ω und η der Form

$$\omega = f dx_{i_1} \wedge \dots \wedge dx_{i_k} \text{ und } \eta = g dx_{j_1} \wedge \dots \wedge dx_{j_\ell}$$

mit $f, g : U \rightarrow \mathbb{R}$ zu zeigen, wobei $x_1, \dots, x_n : U \rightarrow \mathbb{R}$ die kanonischen Koordinatenfunktionen bezeichnen. Dann folgt aus (a) und (b) :

$$\varphi^*\omega = (f \circ \varphi)d\varphi_{i_1} \wedge \dots \wedge d\varphi_{i_k} \text{ und } \varphi^*\eta = (g \circ \varphi)d\varphi_{j_1} \wedge \dots \wedge d\varphi_{j_\ell}$$

mit $\varphi_j = x_j \circ \varphi$. Wegen

$$\omega \wedge \eta = (f \cdot g)dx_{i_1} \wedge \dots \wedge dx_{i_k} \wedge dx_{j_1} \wedge \dots \wedge dx_{j_\ell}$$

folgt daher mit (b):

$$\varphi^*(\omega \wedge \eta) = (f \circ \varphi) \cdot (g \circ \varphi)d\varphi_{i_1} \wedge \dots \wedge d\varphi_{i_k} \wedge d\varphi_{j_1} \wedge \dots \wedge d\varphi_{j_\ell} = \varphi^*\omega \wedge \varphi^*\eta.$$

(d) Sei als Übung überlassen (siehe Aufgabenblatt 11, Aufg 3).

□

Bemerkung. Sei $\varphi : V \rightarrow U$ ein $m + 1$ -mal stetig differenzierbare Abbildung und $\omega \in \Omega_j^k(U)$ mit $j \leq m$. Dann ist $\varphi^*\omega \in \Omega_j^k(V)$. Also definiert

$$\varphi^* : \Omega_j^k(U) \rightarrow \Omega_j^k(V)$$

wegen 19.2.2(a) einen linearen Operator.

Beispiel. Auf $U = \mathbb{R}^2$ mit den Standard-Koordinaten x, y sei die 2-Form $\omega(x, y) = dx \wedge dy$ gegeben. Um diese in Polarkoordinaten darzustellen, soll ω mittels $\varphi : V := (0, \infty) \times (0, 2\pi) \rightarrow U, \varphi(r, \phi) = (r \cos \phi, r \sin \phi)$ zurückgezogen werden. Es gilt

$$d\varphi_1(r, \phi) := d(x \circ \varphi)(r, \phi) = \frac{\partial \varphi_1}{\partial r}(r, \phi) dr + \frac{\partial \varphi_1}{\partial \phi}(r, \phi) d\phi = \cos \phi dr - r \sin \phi d\phi$$

sowie

$$d\varphi_2(r, \phi) := d(y \circ \varphi)(r, \phi) = \frac{\partial \varphi_2}{\partial r}(r, \phi) dr + \frac{\partial \varphi_2}{\partial \phi}(r, \phi) d\phi = \sin \phi dr + r \cos \phi d\phi$$

und somit nach Lemma 19.2.2 (b):

$$\begin{aligned} \varphi^*\omega(r, \phi) &= d\varphi_1 \wedge d\varphi_2(r, \phi) = (\cos \phi dr - r \sin \phi d\phi) \wedge (\sin \phi dr + r \cos \phi d\phi) \\ &= r(\cos^2 \phi + \sin^2 \phi)dr \wedge d\phi = r dr \wedge d\phi. \end{aligned}$$

Die äußere Ableitung kommutiert mit dem Zurückziehen von Differentialformen, d.h. es gilt:

Lemma 19.2.3. Seien $V \subset \mathbb{R}^m, U \subset \mathbb{R}^n$ offen, $\varphi : V \rightarrow U$ eine C^2 -Abbildung und $\omega \in \Omega_1^k(U)$. Dann folgt

$$d(\varphi^*\omega) = \varphi^*d\omega.$$

Beweis. Wegen Satz 19.1.8(a) und obiger Bemerkung sind $d\varphi^*, \varphi^*d : \Omega_1^k(U) \rightarrow \Omega_0^{k+1}(V)$ lineare Operatoren. Also genügt es wieder, die Aussage für Elemente $\omega \in \Omega_1^k(U)$ der Form

$$\omega = f dx_{i_1} \wedge \dots \wedge dx_{i_k}$$

zu beweisen. Dann gilt

$$\varphi^*\omega = (f \circ \varphi) d\varphi_{i_1} \wedge \dots \wedge d\varphi_{i_k}$$

und wegen Satz 19.1.8 (b) folgt

$$d(\varphi^*\omega) = d(f \circ \varphi) \wedge d\varphi_{i_1} \wedge \dots \wedge d\varphi_{i_k} + f \circ \varphi d(d\varphi_{i_1} \wedge \dots \wedge d\varphi_{i_k}).$$

Durch Induktion über k zeigt man mit Satz 19.1.8(b)

$$d(d\varphi_{i_1} \wedge \dots \wedge d\varphi_{i_k}) = 0,$$

denn

$$d(d\varphi_{i_1} \wedge \dots \wedge d\varphi_{i_k}) = d(d\varphi_{i_1} \wedge \dots \wedge d\varphi_{i_{k-1}}) \wedge d\varphi_{i_k} + (-1)^{k-1} d\varphi_{i_1} \wedge \dots \wedge d\varphi_{i_{k-1}} \wedge dd\varphi_{i_k} = 0 + 0 = 0.$$

Außerdem ist $d(f \circ \varphi) = (df \circ \varphi)D\varphi = \varphi^*df$, denn aus der Definition von φ^* und der Kettenregel folgt

$$\varphi^*df(p)(v) = df(\varphi(p))(D\varphi(p)(v)) = d(f \circ \varphi)(p)(v)$$

für alle $p \in U$ und $v \subset \mathbb{R}^n$. Daher folgt aus Lemma 19.2.2 (b) und (c)

$$d(\varphi^*\omega) = \varphi^*df \wedge \varphi^*(dx_{i_1} \wedge \dots \wedge dx_{i_k}) = \varphi^*(df \wedge dx_{i_1} \wedge \dots \wedge dx_{i_k}) = \varphi^*d\omega.$$

□

Nun wollen wir der Frage nachgehen, wann eine k -Form $\omega \in \Omega^k(U)$ eine Stammform $\eta \in \Omega^{k-1}(U)$ besitzt, d.h. wann die Gleichung $d\eta = \omega$ eine Lösung besitzt. Wegen Satz 19.1.7 (iii) ist $d\omega = 0$ ein notwendiges Kriterium dafür.

Definition 19.2.4. Eine Differentialform $\omega \in \Omega_1^k(U)$ heißt *exakt*, falls sie eine Stammform besitzt. Sie heißt *geschlossen*, falls $d\omega = 0$ gilt.

Bemerkungen.

- (a) Die geschlossenen sowie die exakten k -Formen bilden Untervektorräume von $\Omega^k(U)$, denn es gilt:

$$\{\text{geschlossene } k\text{-Formen}\} = \ker(d : \Omega^k(U) \rightarrow \Omega^{k+1}(U)),$$

$$\{\text{exakte } k\text{-Formen}\} = \text{Bild}(d : \Omega^{k-1}(U) \rightarrow \Omega^k(U)).$$

Wegen Satz 19.1.8 (c) sind exakte Formen geschlossen. Der Quotientenraum

$$H^k(U) := \{\text{geschlossene } k\text{-Formen}\} / \{\text{exakte } k\text{-Formen}\}$$

heißt k -te de Rham'sche Kohomologiegruppe. Für $k = 0$ definiert man:

$$H^0(U) := \{\text{geschlossene 0-Formen}\} = \{f : U \rightarrow \mathbb{R} \mid df = 0\}.$$

Ist U zusammenhängend, so ist $H^0(U)$ die Menge der auf U konstanten Funktionen. Also folgt in diesem Falle $H^0(U) \cong \mathbb{R}$.

Für $k = 1$ und U sternförmig gilt: $H^1(U) = \{0\}$, denn wegen des Lemmas von Poincaré für 1-Formen ist jede geschlossene 1-Form auf U auch exakt.

- (b) Es gilt: $H^1(\mathbb{R}^2 \setminus \{0\}) \neq \{0\}$. Betrachte die Windungsform (siehe Kap. 17.2) $\omega = \frac{-ydx + xdy}{x^2 + y^2}$. Dann ist $d\omega = 0$, aber ω ist nicht exakt. Man kann zeigen: $H^1(\mathbb{R}^2 \setminus \{0\}) \cong \mathbb{R}$.

Nun wollen wir das für 1-Formen bewiesene Lemma von Poincaré (siehe Satz 17.2.12) auf k -Formen verallgemeinern, d.h. wir zeigen, dass jede geschlossene k -Form auf einer sternförmigen Menge exakt ist.

Zu einer offenen Menge $U \subset \mathbb{R}^n$ betrachte zunächst die Zylindermenge

$$[0, 1] \times U = \{(t, x) \mid t \in [0, 1], x \in U\}.$$

Sei $V \subset \mathbb{R}^{n+1}$ eine offene Menge, die $[0, 1] \times U$ enthält. Jede k -Form η auf V lässt sich in eindeutiger Weise schreiben als

$$\eta(t, x) = \sum_{1 \leq i_1 < \dots < i_{k-1} \leq n} a_{i_1 \dots i_{k-1}}(t, x) dt \wedge dx_{i_1} \wedge \dots \wedge dx_{i_{k-1}} + \tilde{\eta}(t, x),$$

wobei

$$\tilde{\eta}(t, x) = \sum_{1 \leq j_1 < \dots < j_k \leq n} b_{j_1 \dots j_k}(t, x) dx_{j_1} \wedge \dots \wedge dx_{j_k},$$

d.h. $\tilde{\eta}$ besteht aus allen Monomen, die dt nicht enthalten. Definiere für $k \geq 1$ die Abbildung K vom Raum der stetigen k -Formen auf V in den Raum der $k-1$ -Formen auf U durch

$$(K(\eta))(x) := \sum_{1 \leq i_1 < \dots < i_{k-1} \leq n} \left(\int_0^1 a_{i_1 \dots i_{k-1}}(t, x) dt \right) dx_{i_1} \wedge \dots \wedge dx_{i_{k-1}}$$

für $x \in U$. Damit ist $K : \Omega_0^k(V) \rightarrow \Omega_0^{k-1}(U)$ ein linearer Operator.

Lemma 19.2.5. Sei $U \subset \mathbb{R}^n$ offen und $V \subset \mathbb{R}^{n+1}$ eine offene Teilmenge mit $[0, 1] \times U \subset V$. Seien $\psi_0, \psi_1 : U \rightarrow V$ definiert durch

$$\psi_0(x) = (0, x) \quad \text{und} \quad \psi_1(x) = (1, x).$$

Dann gilt für $\eta \in \Omega_1^k(V)$:

$$d(K(\eta)) + K(d\eta) = \psi_1^* \eta - \psi_0^* \eta. \quad (19.3)$$

Beweis. Da beide Seiten in (19.3) linear von η abhängen, genügt es, beide Seiten für Monome zu überprüfen.

1. Fall: Zunächst sei η von der Form:

$$\eta(t, x) = f(t, x) dt \wedge dx_{i_1} \wedge \dots \wedge dx_{i_{k-1}} = f(t, x) dt \wedge dx_I.$$

Dann folgt aus $df(t, x) = \frac{\partial f}{\partial t}(t, x)dt + \sum_{i=1}^n \frac{\partial f}{\partial x_i}(t, x)dx_i$ wegen $dt \wedge dt = 0$:

$$d\eta(t, x) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(t, x)dx_i \wedge dt \wedge dx_I = - \sum_{i=1}^n \frac{\partial f}{\partial x_i}(t, x)dt \wedge dx_i \wedge dx_I.$$

Dies impliziert

$$K(d\eta)(x) = - \sum_{i=1}^n \left(\int_0^1 \frac{\partial f}{\partial x_i}(t, x)dt \right) dx_i \wedge dx_I.$$

Mittels Differentiation unter dem Integral gilt

$$d(K(\eta))(x) = d \left(\left(\int_0^1 f(t, x)dt \right) dx_I \right) = \sum_{i=1}^n \left(\int_0^1 \frac{\partial f}{\partial x_i}(t, x)dt \right) dx_i \wedge dx_I.$$

Wir erhalten:

$$d(K(\eta)) + K(d\eta) = 0$$

Auf der anderen Seite folgt aus Lemma 19.2.2 (b)

$$\psi_0^* \eta = (f \circ \psi_0) d(t \circ \psi_0) \wedge d(x_{i_1} \circ \psi_0) \wedge \dots \wedge d(x_{i_{k-1}} \circ \psi_0)$$

und

$$\psi_1^* \eta = (f \circ \psi_1) d(t \circ \psi_1) \wedge d(x_{i_1} \circ \psi_1) \wedge \dots \wedge d(x_{i_{k-1}} \circ \psi_1).$$

Da $t : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}$ die Projektion auf die erste Koordinate ist, d.h. $t(s, x) = s$, so folgt

$$t \circ \psi_0(x) = 0 \text{ und } t \circ \psi_1(x) = 1 \text{ und somit } d(t \circ \psi_i) = 0 \text{ f\u00fcr } i \in \{0, 1\}.$$

Also erhalten wir

$$\psi_0^* \eta = \psi_1^* \eta = 0$$

und daher ist die Behauptung im 1. Fall bewiesen.

2. Fall:

Sei nun η ein Monom der Form $\eta = f \cdot dx_I$, wobei $dx_I = dx_{i_1} \wedge \dots \wedge dx_{i_k}$. Dann ist $K(\eta) = 0$ und somit auch $dK(\eta) = 0$. Da

$$d\eta(x, t) = \frac{\partial f}{\partial t}(t, x)dt \wedge dx_I + \sum_{i=1}^n \frac{\partial f}{\partial x_i}(t, x)dx_i \wedge dx_I,$$

folgt aus der Definition des Operators K :

$$K(d\eta)(x) = \left(\int_0^1 \frac{\partial f}{\partial t}(t, x)dt \right) dx_I = f(1, x)dx_I - f(0, x)dx_I.$$

Aus Lemma 19.2.2 (b) folgt

$$\psi_1^* \eta(x) = f \circ \psi_1(x) d(x_{i_1} \circ \psi_1) \wedge \dots \wedge d(x_{i_k} \circ \psi_1) = f(1, x) dx_{i_1} \wedge \dots \wedge dx_{i_k} = f(1, x) dx_I$$

sowie analog

$$\psi_0^* \eta(x) = f(0, x) dx_I.$$

Daher ist die Behauptung auch im 2. Fall bewiesen. \square

Sei nun $U \subset \mathbb{R}^n$ eine bezüglich $0 \in \mathbb{R}^n$ sternförmige Menge und $\varphi : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ die Abbildung mit $\varphi(t, x) = t \cdot x$. Dann enthält $V := \varphi^{-1}(U)$ die Menge $[0, 1] \times U$. Obiges Lemma impliziert somit das folgende Korollar.

Korollar 19.2.6. *Sei $U \subset \mathbb{R}^n$ sternförmig bezüglich $0 \in \mathbb{R}^n$ und $\omega \in \Omega_1^k(U)$. Dann ist $\varphi^*\omega \in \Omega_1^k(V)$ und es gilt:*

$$d(K(\varphi^*\omega)) + K(d\varphi^*\omega) = \omega.$$

Beweis. Sei $\omega \in \Omega_1^k(U)$ und $\eta := \varphi^*\omega$. Dann ist wegen Lemma 19.2.5

$$d(K(\eta)) + K(d\eta) = \psi_1^*\eta - \psi_0^*\eta,$$

wobei $\psi_0, \psi_1 : U \rightarrow V$ mit $\psi_0(x) = (0, x)$ und $\psi_1(x) = (1, x)$. Da $\varphi \circ \psi_1(x) = 1 \cdot x = x$ und $\varphi \circ \psi_0(x) = 0$, folgt

$$\psi_1^*\eta = \psi_1^*\varphi^*\omega = (\varphi \circ \psi_1)^*\omega = \text{id}^*\omega = \omega \quad \text{und} \quad \psi_0^*\eta = (\varphi \circ \psi_0)^*\omega = 0^*\omega = 0$$

und somit die Behauptung. \square

Aus diesem Korollar erhält man nun, dass jede geschlossene k -Form auf einer sternförmigen Menge eine Stammform besitzt.

Satz 19.2.7 (Lemma von Poincaré). *Sei $U \subset \mathbb{R}^n$ sternförmig und ω eine geschlossene k -Form auf U . Dann ist ω exakt.*

Beweis. Ohne Beschränkung der Annahmen sei U sternförmig bezüglich $0 \in \mathbb{R}^n$ und ω geschlossen. Dann folgt mit Lemma 19.2.3 $d\varphi^*\omega = \varphi^*d\omega = 0$ und eingesetzt in Korollar 19.2.6 somit $d(K(\varphi^*\omega)) = \omega$, d.h. ω ist exakt. \square

Das Lemma von Poincaré liefert auch eine Methode zur Berechnung von Stammformen für geschlossene Formen im Falle von sternförmigen Mengen.

Korollar 19.2.8. *Sei $U \subset \mathbb{R}^n$ eine sternförmige offene Menge und*

$$\omega = \sum_{1 \leq i_1 < \dots < i_k \leq n} a_{i_1 \dots i_k} dx_{i_1} \wedge \dots \wedge dx_{i_k} \in \Omega_m^k(U)$$

für ein $k, m \geq 1$. Sei

$$\eta(x) = \sum_{1 \leq i_1 < \dots < i_k \leq n} \sum_{\ell=1}^k (-1)^{\ell-1} \int_0^1 t^{k-1} a_{i_1 \dots i_k}(tx) dt x_{i_\ell} dx_{i_1} \wedge \dots \widehat{dx_{i_\ell}} \dots \wedge dx_{i_k},$$

wobei $\widehat{dx_{i_\ell}}$ die Auslassung von dx_{i_ℓ} anzeigt. Dann ist $\eta \in \Omega_m^{k-1}(U)$ und falls ω geschlossen ist, folgt $d\eta = \omega$.

Beweis. Aus der Definition von η folgt sofort: $\eta \in \Omega_m^{k-1}(U)$.

Wie im Beweis von Korollar 19.2.6 betrachte die Abbildung $\varphi : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ mit $\varphi(t, x) = t \cdot x$. Dann enthält $V := \varphi^{-1}(U)$ die Menge $[0, 1] \times U$. Man zeigt nun:

$$K(\varphi^*\omega) = \eta. \tag{19.4}$$

Dann folgt aus Korollar 19.2.6 $d\eta = \omega$, falls ω geschlossen ist. Wegen der Linearität des Operators $K\varphi^* : \Omega_m^k(U) \rightarrow \Omega_m^{k-1}(U)$ genügt es, die Identität (19.4) für Monome zu zeigen. Sei also $\omega \in \Omega_m^k(U)$ von der Form

$$\omega(x) = f(x)dx_{i_1} \wedge \dots \wedge dx_{i_k}.$$

Dann gilt:

$$\varphi^*\omega(t, x) = f(tx)d\varphi_{i_1} \wedge \dots \wedge d\varphi_{i_k}.$$

Da $\varphi_{i_\ell}(t, x) = tx_{i_\ell}$, folgt

$$d\varphi_{i_\ell} = x_{i_\ell}dt + tdx_{i_\ell}.$$

Man zeigt nun durch Induktion über k :

$$d\varphi_{i_1} \wedge \dots \wedge d\varphi_{i_k}(t, x) = t^k dx_{i_1} \wedge \dots \wedge dx_{i_k} + \sum_{\ell=1}^k (-1)^{\ell-1} t^{k-1} x_{i_\ell} dt \wedge dx_{i_1} \wedge \dots \wedge \widehat{dx_{i_\ell}} \wedge \dots \wedge dx_{i_k}.$$

Nach Definition des Operators K folgt

$$K(\varphi^*\omega)(x) = \sum_{\ell=1}^k (-1)^{\ell-1} \int_0^1 t^{k-1} f(tx) dt x_{i_\ell} dx_{i_1} \wedge \dots \wedge \widehat{dx_{i_\ell}} \wedge \dots \wedge dx_{i_k}$$

und daraus die Behauptung. □

Nun geben wir einige Anwendungen des Lemmas von Poincaré.

Bemerkungen. (*Anwendungen in der Vektoranalysis*)

- (a) Sei $\omega \in \Omega^k(U)$ eine unendlich oft differenzierbare k -Form auf einer offenen Teilmenge U des \mathbb{R}^n . Gesucht wird eine Lösung $\eta \in \Omega^{k-1}(U)$ der Differentialgleichung

$$d\eta = \omega.$$

Wegen $dd\eta = 0$ ist $d\omega = 0$ eine notwendige Bedingung für ihre Lösbarkeit. Ist nun U sternförmig, so ist diese notwendige Bedingung wegen des Lemmas von Poincaré auch hinreichend.

Wie sieht die Gesamtheit der Lösungen

$$\mathcal{L}_\omega := \{\eta \in \Omega^{k-1}(U) \mid d\eta = \omega\}$$

aus? Ist η_0 eine spezielle Lösung (diese kann mit Hilfe des Korollars 19.2.8 berechnet werden), so ist für jedes $\beta \in \Omega^{k-2}(U)$ auch $\eta = \eta_0 + d\beta$ wegen $dd\beta = 0$ eine Lösung. Ist umgekehrt η eine beliebige Lösung, so ist die Differenz $\eta - \eta_0 \in \Omega^{k-1}(U)$ geschlossen. Also existiert wieder wegen des Lemmas von Poincaré ein $\beta \in \Omega^{k-2}(U)$ mit $\eta - \eta_0 = d\beta$. Dies zeigt, dass \mathcal{L}_ω ein affiner Unterraum von $\Omega^{k-1}(U)$ ist (vgl. 16.8):

$$\mathcal{L}_\omega = \{\eta_0 + d\beta \mid \beta \in \Omega^{k-2}(U)\}.$$

- (b) Teil (a) hat folgende Anwendungen in der Vektoranalysis. Sei $U \subset \mathbb{R}^3$ eine offene und sternförmige Menge und $g : U \rightarrow \mathbb{R}$ eine C^∞ -Funktion. Dann existiert aus folgendem Grund ein Vektorfeld $F : U \rightarrow \mathbb{R}^3$ mit

$$\operatorname{div} F = g.$$

Betrachte die 3-Form $\omega = g dx_1 \wedge dx_2 \wedge dx_3$. Diese Differentialform ist – wie alle 3-Formen in \mathbb{R}^3 – geschlossen und damit existiert eine 2-Form η mit $d\eta = \omega$. Wegen des Beispiels am Ende von 19.1 existiert ein Vektorfeld $F : U \rightarrow \mathbb{R}^3$ mit $\omega_F^2 = \eta$. Also folgt $d\eta_F^2 = \operatorname{div} F dx_1 \wedge dx_2 \wedge dx_3$ und somit $\operatorname{div} F = g$. Ist $F_0 : U \rightarrow \mathbb{R}^3$ eine spezielle Lösung, so ist die Menge aller Lösungen von der Form

$$\{F_0 + \operatorname{rot} G \mid G : U \rightarrow \mathbb{R}^3 \text{ unendlich oft differenzierbar}\}.$$

Denn für jede beliebige Lösung $F : U \rightarrow \mathbb{R}^3$ gilt $\operatorname{div}(F - F_0) = 0 \Rightarrow d\omega_{F-F_0}^2 = 0$. Wie in (a) gezeigt, existiert eine 1-Form β mit $d\beta = \omega_{F-F_0}^2$. Sei $G : U \rightarrow \mathbb{R}^3$ das zu β gehörige Vektorfeld, also $\beta = \omega_G^1$. Dann ist $\omega_{\operatorname{rot} G}^2 = d\beta = \omega_{F-F_0}^2$ und somit ist $\operatorname{rot} G = F - F_0$.

Sei nun $H : U \rightarrow \mathbb{R}^3$ ein unendlich oft differenzierbares Vektorfeld. Gesucht ist die Menge aller Vektorfelder $F : U \rightarrow \mathbb{R}^3$ mit

$$\operatorname{rot} F = H.$$

Mit analogen Argumenten wie oben folgt unter Beachtung des oben zitierten Beispiels: Diese Gleichung hat genau dann eine Lösung, falls $\operatorname{div} H = 0$ gilt. Ist F_0 eine spezielle Lösung, so ist die Gesamtheit der Lösungen von der Form

$$\{F_0 + \operatorname{grad} f \mid f : U \rightarrow \mathbb{R} \text{ unendlich oft differenzierbar}\}.$$

Die Lösungen heißen auch *Vektorpotentiale* von H .

Bemerkungen. (*Anwendungen in der Topologie*)

- (a) Es gilt für alle $k \geq 1$

$$H^k(V) = 0, \quad \text{falls } V \text{ sternförmig.}$$

- (b) Dies zeigt, dass jede geschlossene Form $\eta \in \Omega_1^k(U)$ lokal exakt ist, denn zu jedem $p \in U$ gibt es ein $\epsilon > 0$ mit $B(p, \epsilon) \subset U$. Der Ball $B(p, \epsilon)$ ist konvex und somit insbesondere sternförmig.

19.3 Integration von Differentialformen

Wir wollen nun das Integral einer k -Form auf einer offenen Teilmenge des \mathbb{R}^n erklären. Zunächst definieren wir das Integral einer stetigen k -Form $\omega \in \Omega_0^k(U)$ auf einer offenen Teilmenge U des \mathbb{R}^k . Für jede solche k -Form ω existiert eine stetige Funktion $f : U \rightarrow \mathbb{R}$ mit

$$\omega(x) = f(x) dx_1 \wedge \dots \wedge dx_k.$$

Definition 19.3.1 (Integral von k -Formen auf dem \mathbb{R}^k). Sei $U \subset \mathbb{R}^k$ offen und $\omega \in \Omega^k(U)$ mit

$$\omega = f(x)dx_1 \wedge \dots \wedge dx_k.$$

Ist $A \subset U$ Lebesgue-messbar, so heißt ω integrierbar über A , falls das Lebesgue-Integral von f über A existiert. Man definiert:

$$\int_A \omega := \int_A f(x)d\lambda^k(x) := \int_{\mathbb{R}^k} f(x) \cdot \chi_A(x)d\lambda^k(x).$$

Bemerkung. Statt $\lambda^k(x)$ schreibt man auch d^kx oder dx . Ist $\omega \in \Omega_0^k(U)$ und A kompakt, also abgeschlossen und beschränkt, so ist ω über A integrierbar. Insbesondere gilt:

$$\int_A dx_1 \wedge \dots \wedge dx_k = \lambda^k(A).$$

Daher nennt man $dx_1 \wedge \dots \wedge dx_k \in \Omega^k(\mathbb{R}^k)$ auch die *Volumenform* des \mathbb{R}^k .

Genau wie wir 1-Formen auf dem \mathbb{R}^n über Kurven im \mathbb{R}^n integriert haben, lassen sich k -Formen auf dem \mathbb{R}^n über “ k -dimensionale Mengen” im \mathbb{R}^n integrieren.

Definition 19.3.2 (Integral von k -Formen auf dem \mathbb{R}^n). Seien $V \subset \mathbb{R}^k$ und $U \subset \mathbb{R}^n$ offene Mengen mit $k \leq n$. Seien $\omega \in \Omega_0^k(U)$ eine stetige k -Form und $\varphi : V \rightarrow U$ eine C^1 -Abbildung. Dann ist $\varphi^*\omega \in \Omega_0^k(V)$. Ist $A \subset V$ messbar und $\varphi^*\omega$ integrierbar über A , so definiere

$$\int_{(\varphi,A)} \omega := \int_A \varphi^*\omega.$$

Falls $A = V$, schreibe $\int_{\varphi} \omega$ statt $\int_{(\varphi,V)} \omega$.

Beispiel. Sei $U \subset \mathbb{R}^n$ offen und $\omega \in \Omega_0^1(U)$ eine stetige 1-Form. Ist $\gamma : I \rightarrow U$ eine differenzierbare Kurve, so ist

$$\int_{\gamma} \omega = \int_I \gamma^*\omega.$$

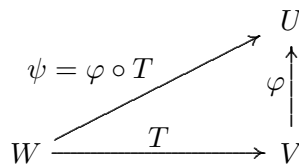
Da $\gamma^*\omega \in \Omega_0^1(I)$, ist $\gamma^*\omega(t) = f(t) dt$ für $t \in I$ und somit

$$f(t) = \gamma^*\omega(t)(1) = \omega(\gamma(t))D\gamma(t)(1) = \omega(\gamma(t))\dot{\gamma}(t).$$

Insbesondere stimmt diese Definition mit der Definition der Kurvenintegrale (siehe 17.2.3) überein.

Das Integral $\int_{\varphi} \omega$ kann unter einer Reparametrisierung von φ nur das Vorzeichen ändern.

Definition 19.3.3. Seien $V \subset \mathbb{R}^k, U \subset \mathbb{R}^n$ offen und $\varphi : V \rightarrow U$ eine C^1 -Abbildung. Ist $W \subset \mathbb{R}^k$ offen, so heißt eine Abbildung $\psi : W \rightarrow U$ *Reparametrisierung* von φ , falls es einen Diffeomorphismus $T : W \rightarrow V$ gibt mit $\psi = \varphi \circ T$.



Ist $\det DT(x) > 0$ für alle $x \in W$, so heißt die Reparametrisierung *orientierungserhaltend*.
Ist $\det DT(x) < 0$ für alle $x \in W$, so heißt sie *orientierungsumkehrend*.

Bemerkung. Sei W eine zusammenhängende Menge, so ist jede Reparametrisierung orientierungserhaltend oder orientierungsumkehrend, denn in diesem Falle ist $\det DT(x) > 0$ oder $\det DT(x) < 0$ für alle $x \in W$.

Lemma 19.3.4. Seien $V \subset \mathbb{R}^k, U \subset \mathbb{R}^n$ offen, $\omega \in \Omega_0^k(U)$ und $\varphi : V \rightarrow U$ eine C^1 -Abbildung. Ist $\psi : W \rightarrow U$ eine orientierungserhaltende (orientierungsumkehrende) Reparametrisierung von φ , so gilt:

$$\int_{\varphi} \omega = \pm \int_{\psi} \omega.$$

Beweis. Zunächst gilt mit $\psi = \varphi \circ T$:

$$\int_{\psi} \omega = \int_W \psi^* \omega = \int_W (\varphi \circ T)^* \omega = \int_W T^* \varphi^* \omega.$$

Da $\varphi^* \omega \in \Omega_0^k(V)$ und $V \subset \mathbb{R}^k$, gilt für eine stetige Funktion $g : V \rightarrow \mathbb{R}$:

$$\varphi^* \omega = g dx_1 \wedge \dots \wedge dx_k$$

und somit folgt aus Lemma 19.2.2(b)

$$T^* \varphi^* \omega = (g \circ T) dT_1 \wedge \dots \wedge dT_k \in \Omega_0^k(W),$$

wobei $W \subset \mathbb{R}^k$ eine offene Teilmenge ist und $T = (T_1, \dots, T_k)$ die Komponentendarstellung von T . Also gilt:

$$(dT_1 \wedge \dots \wedge dT_k)(p) = a(p) dx_1 \wedge \dots \wedge dx_k$$

für eine stetige Funktion $a : W \rightarrow \mathbb{R}$. Wegen

$$(dT_1 \wedge \dots \wedge dT_k)(p)(e_1, \dots, e_k) = \det \begin{pmatrix} dT_1(p)(e_1) & \dots & dT_1(p)(e_k) \\ \vdots & & \vdots \\ dT_k(p)(e_1) & \dots & dT_k(p)(e_k) \end{pmatrix} = \det DT(p)$$

erhalten wir $a(p) = \det DT(p)$ für jedes $p \in W$ und daher

$$T^* \varphi^* \omega(p) = (g \circ T(p)) \det dT(p) dx_1 \wedge \dots \wedge dx_k.$$

Also gilt

$$\int_{\varphi} \omega = \int_W (g \circ T(p)) \det DT(p) d\lambda^k(p).$$

Auf der anderen Seite ergibt sich aus der Transformationsformel:

$$\int_W g \circ T(p) |\det DT(p)| d\lambda^k(p) = \int_{T(W)} g d\lambda^k = \int_V g d\lambda^k = \int_{\varphi} \omega.$$

Ist also $\det DT > 0$, so folgt $\int_{\varphi} \omega = \int_{\psi} \omega$. Ist $\det DT < 0$, so folgt $\int_{\varphi} \omega = - \int_{\psi} \omega$. □

Nun wollen wir Differentialformen über Untermannigfaltigkeiten integrieren. Dies geschieht mittels Parametrisierungen (siehe Def. 16.4.3).

Definition 19.3.5. Sei $M \subset \mathbb{R}^N$ eine n -dimensionale Untermannigfaltigkeit. Eine Familie $\{(\varphi_\alpha, V_\alpha)\}_{\alpha \in I}$ von offenen Mengen $V_\alpha \subset \mathbb{R}^n$ und injektiven Parametrisierungen

$$\varphi_\alpha : V_\alpha \rightarrow \varphi_\alpha(V_\alpha) =: U_\alpha \subset \mathbb{R}^N$$

mit

$$\bigcup_{\alpha \in I} U_\alpha = M.$$

heißt *Atlas* aus Parametrisierungen für M .

Zu jeder Untermannigfaltigkeit existiert ein Atlas aus lokalen (insbesondere injektiven) Parametrisierungen (vgl. Bem. (c) nach 16.4.3). Wir setzen im Folgenden immer voraus, dass die lokalen Parametrisierungen unendlich oft differenzierbar sind.

Definition 19.3.6. Sei $M \subset \mathbb{R}^N$ eine n -dimensionale Untermannigfaltigkeit. Eine *Differentialform vom Grade k* auf M ist eine Abbildung

$$\omega : M \rightarrow \bigcup_{p \in M} \Lambda^k(T_p^*M)$$

mit $\omega(p) \in \Lambda^k(T_p^*M)$ für alle $p \in M$. Die Differentialform ω heißt *m -mal stetig differenzierbar*, falls für jede lokale Parametrisierung $\varphi_\alpha : V_\alpha \rightarrow U_\alpha \subset M$ die zurückgeholte Differentialform

$$\omega_\alpha := \varphi_\alpha^* \omega$$

mit

$$\varphi_\alpha^* \omega(x)((u_1), \dots, (u_k)) := \omega(\varphi_\alpha(x))(D\varphi_\alpha(x)u_1, \dots, D\varphi_\alpha(x)u_k)$$

für $u_1, \dots, u_k \in \mathbb{R}^n$ und $x \in V_\alpha$ m -mal stetig differenzierbar ist. Die Menge aller m -mal stetig differenzierbaren k -Formen auf M bezeichnen wir mit $\Omega_m^k(M)$. Mit $\Omega^k(M)$ bezeichnen wir die unendlich oft differenzierbaren k -Formen auf M . Die auf der offenen Menge $V_\alpha \subset \mathbb{R}^n$ definierte Differentialform $\omega_\alpha = \varphi_\alpha^* \omega$ nennen wir auch *lokale Darstellung* von ω bezüglich der Parametrisierung φ_α .

Bemerkung. Seien $\varphi_\alpha : V_\alpha \rightarrow U_\alpha \subset M$ und $\varphi_\beta : V_\beta \rightarrow U_\beta \subset M$ zwei lokale Parametrisierungen mit $U_\alpha \cap U_\beta \neq \emptyset$. Definiere $V_{\alpha\beta} = \varphi_\beta^{-1}(U_\alpha \cap U_\beta)$. Dann ist $T_{\alpha\beta} : V_{\alpha\beta} \rightarrow V_{\alpha\beta}$ mit $T_{\alpha\beta} = \varphi_\beta^{-1} \circ \varphi_\alpha$ ein Diffeomorphismus und es gilt für die lokalen Darstellungen ω_α und ω_β :

$$\omega_\alpha = T_{\alpha\beta}^* \omega_\beta, \tag{19.5}$$

denn

$$T_{\alpha\beta}^* \omega_\beta = T_{\alpha\beta}^* \varphi_\beta^* \omega = (\varphi_\beta \circ T_{\alpha\beta})^* \omega = \varphi_\alpha^* \omega = \omega_\alpha.$$

Lemma 19.3.7. Sei $M \subset \mathbb{R}^N$ eine n -dimensionale Untermannigfaltigkeit und $(\varphi_\alpha, V_\alpha)_{\alpha \in I}$ ein Atlas aus Parametrisierungen für M . Seien $\omega_\alpha \in \Omega_m^k(V_\alpha)$, $\alpha \in I$, eine Familie von m -mal stetig differenzierbaren k -Formen, die die Kompatibilitätsbedingungen

$$\omega_\alpha = T_{\alpha\beta}^* \omega_\beta \text{ für alle } \alpha, \beta \in I \text{ mit } \varphi_\alpha(V_\alpha) \cap \varphi_\beta(V_\beta) \neq \emptyset \tag{19.6}$$

erfüllen. Dann existiert genau eine k -Form $\omega \in \Omega_m^k(M)$ mit $\varphi_\alpha^* \omega = \omega_\alpha$ für alle $\alpha \in I$.

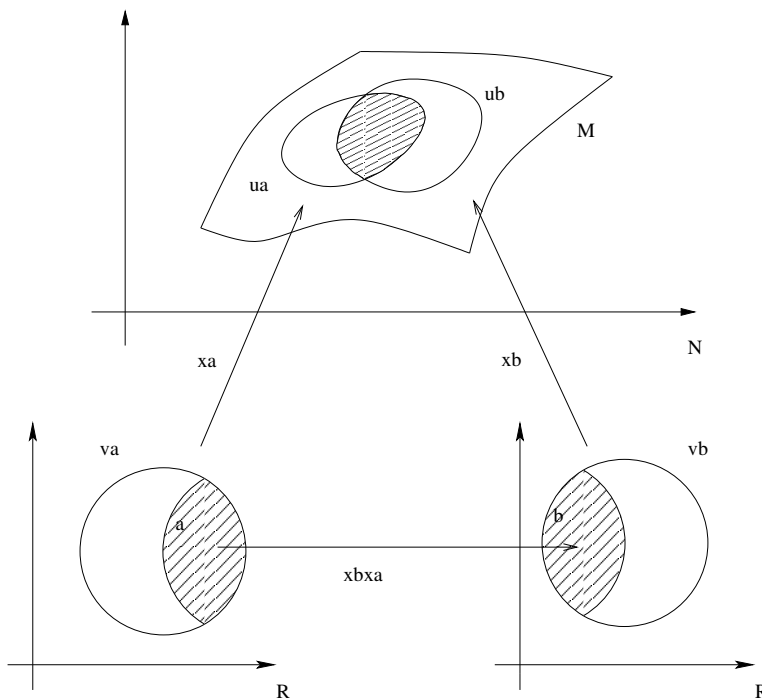


Abbildung 19.2:

Beweis. Sei $p \in M$ und $v_1, \dots, v_k \in T_p M$. Wähle ein $\alpha \in I$ mit $p \in \varphi_\alpha(V_\alpha)$. Dann existieren genau ein $x \in V_\alpha$ und eindeutige Vektoren $u_1, \dots, u_k \in \mathbb{R}^n$ mit $p = \varphi_\alpha(x)$ und $v_j = D\varphi_\alpha(x)(u_j)$. Definiere nun

$$\omega(p)(v_1, \dots, v_k) := \omega_\alpha(x)(u_1, \dots, u_k).$$

Zu zeigen ist, dass diese Definition nicht von der Wahl von α abhängt. Sei also $\beta \in I$ mit $p \in \varphi_\beta(V_\beta)$. Dann existieren wieder genau ein $x' \in V_\beta$ und eindeutige Vektoren $u'_1, \dots, u'_k \in \mathbb{R}^n$ mit $p = \varphi_\beta(x')$ und $v_j = D\varphi_\beta(x')(u'_j)$. Zu zeigen ist

$$\omega_\alpha(x)(u_1, \dots, u_k) = \omega_\beta(x')(u'_1, \dots, u'_k).$$

Es gilt

$$x' = \varphi_\beta^{-1} \circ \varphi_\alpha(x) = T_{\alpha\beta}(x)$$

und aus der Kettenregel folgt:

$$u'_j = (D\varphi_\beta(x'))^{-1} D\varphi_\alpha(x)(u_j) = D\varphi_\beta^{-1}(p) D\varphi_\alpha(x)(u_j) = D(\varphi_\beta^{-1} \circ \varphi_\alpha)(x)(u_j) = DT_{\alpha\beta}(x)(u_j).$$

Also erhalten wir zusammen mit den Kompatibilitätsbedingungen (19.6):

$$\begin{aligned} \omega_\beta(x')(u'_1, \dots, u'_k) &= \omega_\beta(T_{\alpha\beta}(x))(DT_{\alpha\beta}(x)(u_1), \dots, DT_{\alpha\beta}(x)(u_k)) = T_{\alpha\beta}^* \omega_\beta(x)(u_1, \dots, u_k) \\ &= \omega_\alpha(x)(u_1, \dots, u_k). \end{aligned}$$

□

Dieser Sachverhalt ermöglicht es uns, die äußere Ableitung einer Differentialform auf einer Untermannigfaltigkeit zu definieren.

Definition 19.3.8. Sei $M \subset \mathbb{R}^N$ eine n -dimensionale Untermannigfaltigkeit und $(\varphi_\alpha, V_\alpha)_{\alpha \in I}$ ein Atlas aus Parametrisierungen für M . Sei $m \geq 1$ und $\omega \in \Omega_m^k(M)$. Dann definiere ihr Differential

$$d\omega \in \Omega_{m-1}^{k+1}(M)$$

als die eindeutig bestimmte $(k+1)$ -Form in M mit $\varphi_\alpha^* d\omega = d\omega_\alpha$ für alle $\alpha \in I$.

Bemerkung. Die Familie $d\omega_\alpha \in \Omega_{m-1}^{k+1}(V_\alpha)$, $\alpha \in I$, erfüllt wegen

$$d\omega_\alpha = d(T_{\alpha\beta}^* \omega_\beta) = T_{\alpha\beta}^* d\omega_\beta$$

die Kompatibilitätsbedingungen (19.6). Daher ist $d\omega$ wegen Lemma 19.3.7 wohldefiniert.

Wir wollen nun das Integral von n -Formen über n -dimensionale Untermannigfaltigkeiten definieren.

Ist $V \subset \mathbb{R}^n$ offen, $\varphi : V \rightarrow U \subset M$ eine lokale Parametrisierung und $\omega \in \Omega_0^n(M)$ eine stetige n -Form, so ist es naheliegend, $\int_U \omega$ durch $\int_V \varphi^* \omega$ zu erklären. Sei $\psi : W \rightarrow U$ eine weitere lokale Parametrisierung, so ist $T := \varphi^{-1} \circ \psi : W \rightarrow V$ ein Diffeomorphismus und somit ψ eine Reparametrisierung von φ . Daher gilt wegen Lemma 19.3.3

$$\int_W \psi^* \omega = \pm \int_V \varphi^* \omega$$

je nachdem, ob T orientierungserhaltend oder umkehrend ist.

Definition 19.3.9. Eine Untermannigfaltigkeit $M \subset \mathbb{R}^N$ heißt *orientierbar*, falls ein Atlas $(\varphi_\alpha, V_\alpha)_{\alpha \in I}$ aus Parametrisierungen für M existiert mit

$$\det DT_{\alpha\beta}(x) > 0$$

für alle $x \in \varphi_\alpha^{-1}((\varphi_\alpha(V_\alpha) \cap (\varphi_\beta(V_\beta)))$. Der Atlas $(\varphi_\alpha, V_\alpha)_{\alpha \in I}$ heißt dann *orientiert*.

Definition 19.3.10. Sei $M \subset \mathbb{R}^N$ eine n -dimensionale Untermannigfaltigkeit mit orientiertem Atlas $(\varphi_\alpha, V_\alpha)_{\alpha \in I}$. Sei $\omega \in \Omega_0^n(M)$, so dass

$$\text{supp } \omega := \overline{\{x \in M \mid \omega(x) \neq 0\}} \subset \varphi_\alpha(V_\alpha)$$

für ein $\alpha \in I$. Dann definiere

$$\int_M \omega := \int_{V_\alpha} \varphi_\alpha^* \omega.$$

Die Menge $\text{supp } \omega$ heißt auch *Träger* (engl. support) der Form ω .

Bemerkung. Die Definition des Integrals hängt nicht von der Wahl der lokalen Parametrisierung φ_α ab, d.h. ist $\text{supp } \omega \subset \varphi_\alpha(V_\alpha) \cap \varphi_\beta(V_\beta)$, so folgt

$$\int_{V_\alpha} \varphi_\alpha^* \omega = \int_{V_\beta} \varphi_\beta^* \omega.$$

Falls der Träger von ω nicht im Bildbereich einer einzelnen Karte liegt, bedient man sich einer Technik, die man die "Zerlegung der Eins" nennt.

Definition 19.3.11. Eine offene Überdeckung $\{U_\alpha\}_{\alpha \in I}$ der Untermannigfaltigkeit M heißt *lokal endlich*, falls für jedes $\alpha \in I$ die Menge

$$\{\beta \in I \mid U_\alpha \cap U_\beta \neq \emptyset\}$$

endlich ist. Ein Atlas $(\varphi_\alpha, V_\alpha)_{\alpha \in I}$ aus Parametrisierungen für M heißt *lokal endlich*, falls die offene Überdeckung $\{U_\alpha\}_{\alpha \in I}$ der Untermannigfaltigkeit M mit $U_\alpha = \varphi_\alpha(V_\alpha)$ lokal endlich ist.

Satz 19.3.12. Sei $\{U_\alpha\}_{\alpha \in I}$ eine lokal endliche Überdeckung einer Untermannigfaltigkeit M . Dann existiert eine Zerlegung der Eins angepasst an $\{U_\alpha\}_{\alpha \in I}$, d.h. es existiert eine Familie von C^∞ -Funktionen $\rho_\alpha : M \rightarrow [0, 1]$ mit

$$\text{supp } \rho_\alpha = \overline{\{p \in M \mid \rho_\alpha(x) \neq 0\}} \subset U_\alpha \text{ und } \sum_{\alpha \in I} \rho_\alpha(p) = 1 \quad \forall p \in M.$$

Bemerkung. Da die Überdeckung $\{U_\alpha\}_{\alpha \in I}$ lokal endlich ist, sind für jedes $p \in M$ nur endlich viele Summanden in $\sum_{\alpha \in I} \rho_\alpha(p)$ ungleich null.

Beweis. Wir beweisen die Aussage nur für kompakte Mannigfaltigkeit und eine endliche Überdeckung U_1, \dots, U_n . Betrachte dann zu jedem $x \in M$ eine offene Menge $W(x)$ mit $\overline{W(x)} \subset U_i$ für ein $i \in \{1, \dots, n\}$. $\bigcup_{x \in M} W(x)$ überdeckt M und da M kompakt ist, existiert eine endliche Teilüberdeckung. Also existieren $k_i \in \mathbb{N}$ und endlich viele Punkte $x_{i,j}$, $i \in \{1, \dots, n\}$ und $j \in \{1, \dots, k_i\}$, mit $x_{i,j} \in U_i$ und

$$\bigcup_{i=1}^n \bigcup_{j=1}^{k_i} W(x_{i,j}) = M.$$

Setze

$$W_i := \bigcup_{j=1}^{k_i} W(x_{i,j}),$$

so ist $W_i \subset U_i$ offen mit $\overline{W_i} \subset U_i$ und $\bigcup_{i=1}^n W_i = M$. Konstruiere nun C^∞ -Funktionen $f_i : M \rightarrow \mathbb{R}$ mit $f_i(p) > 0$ für $p \in W_i$ und $f_i(p) = 0$ für $p \in M \setminus W_i$. Definiere dann

$$\rho_i(p) = \frac{f_i(p)}{\sum_{j=1}^n f_j(p)}.$$

□

Definition 19.3.13. Sei M eine orientierbare n -dimensionale Untermannigfaltigkeit mit orientiertem Atlas $(\varphi_\alpha, V_\alpha)_{\alpha \in I}$, so dass die Überdeckung $\varphi(V_\alpha) = U_\alpha$ lokal endlich ist.

Sei $\{\rho_\alpha\}_{\alpha \in I}$ eine Zerlegung der Eins angepasst an $\{U_\alpha\}_{\alpha \in I}$. Sei ω eine n -Form auf M , so definiere:

$$\int_M \omega := \sum_{\alpha \in I} \int_M \rho_\alpha \omega.$$

Bemerkung. Aus der Eigenschaft der Zerlegung der Eins folgt: $\omega = \sum_{\alpha \in I} \rho_\alpha \cdot \omega$.

Außerdem kann man zeigen, dass die Definition unabhängig von der Wahl der Zerlegung der Eins ist.

Eine orientierte Untermannigfaltigkeit induziert eine Orientierung der Tangentialräume. Eine Orientierung auf einem Vektorraum ist die Wahl einer Äquivalenzklasse von geordneten Basen.

Definition 19.3.14. Sei V ein n -dimensionaler reeller Vektorraum und

$$GB(V) = \{(b_1, \dots, b_n) \mid \{b_1, \dots, b_n\} \text{ Basis von } V\}$$

die Menge der *geordneten* Basen auf V (siehe auch Kapitel 8). Zwei geordnete Basen $B = (b_1, \dots, b_n)$ und $C = (c_1, \dots, c_n)$ heißen *äquivalent* (in Zeichen $B \sim C$), falls die Determinante der linearen Abbildung $T : V \rightarrow V$ mit $T(b_i) = c_i$ positiv ist. Ist $B = (b_1, \dots, b_n)$ eine Basis, so bezeichnen wir mit

$$[B] := \{C \mid C \in GB(V), C \sim B\}$$

die Äquivalenzklasse von B . Unter einer *Orientierung* auf V verstehen wir die Wahl einer Äquivalenzklasse. Ihre Elemente heißen orientierte Basen.

Die durch die kanonische Basis (e_1, \dots, e_n) des \mathbb{R}^n induzierte Orientierung $[(e_1, \dots, e_n)]$ heißt *Standardorientierung* oder *kanonische Orientierung*. Sei $\varphi : \mathbb{R}^n \rightarrow V$ ein Isomorphismus. Dann heißt $[(\varphi(e_1), \dots, \varphi(e_n))]$ die durch φ *induzierte Orientierung* auf V .

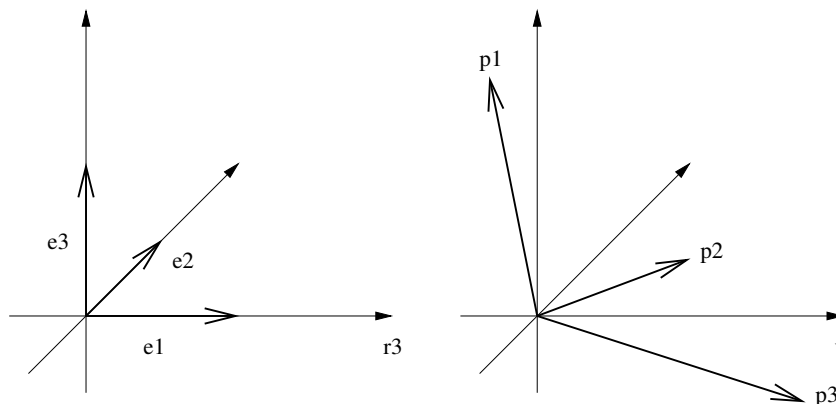


Abbildung 19.3: induzierte Orientierung

Bemerkungen.

- (a) Es existieren genau zwei Äquivalenzklassen. Haben wir eine Orientierung festgelegt, so nennen wir ihre Elemente auch manchmal *positiv* orientierte Basen. Die Elemente der anderen Klasse heißen dann *negativ* orientierte Basen.
- (b) Zwei Isomorphismen $\varphi_1, \varphi_2 : \mathbb{R}^n \rightarrow V$ induzieren genau dann die gleiche Orientierung, falls

$$\det(\varphi_2^{-1} \circ \varphi_1) > 0.$$

- (c) Ist M eine orientierbare Untermannigfaltigkeit mit orientiertem Atlas $(\varphi_\alpha, V_\alpha)_{\alpha \in I}$, so induzieren die Isomorphismen $D\varphi_\alpha(x_\alpha) : \mathbb{R}^n \rightarrow T_p M$ mit $\varphi_\alpha(x_\alpha) = p$ alle die gleiche Orientierung auf $T_p M$, denn ist $\varphi_\alpha(x_\alpha) = \varphi_\beta(x_\beta)$, so ist

$$\det DT_{\alpha\beta}(x_\alpha) = \det D(\varphi_\beta^{-1} \circ \varphi_\alpha)(x_\alpha) > 0.$$

19.4 Integralsatz von Stokes

Nun wollen wir den Integralsatz von Stokes beweisen. Dazu benötigen wir den Begriff der Untermannigfaltigkeiten mit Rand.

Untermannigfaltigkeiten sind lokal diffeomorph zu offenen Teilmengen des \mathbb{R}^n . Untermannigfaltigkeiten mit Rand sind lokal diffeomorph zu offenen Teilmengen des Halbraumes

$$H^n := \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid x_1 \leq 0\}.$$

Die Menge

$$\partial H^n = \{(0, x_2, \dots, x_n) \mid x_i \in \mathbb{R}\}$$

heißt *Rand* von H^n . Der Rand von H^n ist ein Vektorraum isomorph zu \mathbb{R}^{n-1} . Analog zur Definition der Untermannigfaltigkeit ohne Rand in 16.4.1 definieren wir:

Definition 19.4.1. Eine Teilmenge $M \subset \mathbb{R}^N$ heißt *n-dimensionale Untermannigfaltigkeit mit Rand* des \mathbb{R}^N , falls für jedes $p \in M$ eine offene Umgebung $W \subset \mathbb{R}^N$ und ein Diffeomorphismus $\psi : W \rightarrow W' \subset \mathbb{R}^N$ existiert mit

$$\psi(W \cap M) = W' \cap (H^n \times \{0\}),$$

wobei $0 \in \mathbb{R}^{N-n}$. Die Einschränkung von ψ auf $W \cap M$ heißt auch *Karte*. Die Menge der Punkte $p \in M$ mit

$$\psi(p) \in W' \cap (\partial H^n \times \{0\})$$

heißen die *Randpunkte* der Untermannigfaltigkeit. Mit ∂M bezeichnen wir die Menge aller Randpunkte von M .

Bemerkung.

- (a) Die Definition der Randpunkte hängt nicht von der Wahl der Karte ab. Falls $\partial M = \emptyset$, erhalten wir eine Untermannigfaltigkeit ohne Rand im Sinn der Definition 16.4.1.
- (b) Die Menge $W' \cap (\mathbb{R}^n \times \{0\})$ ist von der Form $V \times \{0\}$ mit $V \subset \mathbb{R}^n$ offen. Daher lässt sich ψ auch als Diffeomorphismus von $\psi^{-1}(V \times \{0\}) \subset W$ nach V auffassen. Die Inverse dieser Abbildung werde mit φ notiert. $\varphi : V \rightarrow W$ ist eine Immersion. Ihre Einschränkung $\varphi|_{V \cap H^n} : V \cap H^n \rightarrow W \cap M$ heißt *lokale Parametrisierung* von M .
- (c) Sei $M \subset \mathbb{R}^N$ eine *n-dimensionale Untermannigfaltigkeit mit Rand*. Dann existieren Familien $\{V_\alpha\}_{\alpha \in I} \subset \mathbb{R}^n$ und $\{W_\alpha\}_{\alpha \in I} \subset \mathbb{R}^N$ offener Mengen sowie eine Familie von Immersionen $\varphi_\alpha : V_\alpha \rightarrow W_\alpha$, so dass für die Parametrisierungen $\varphi_\alpha|_{V_\alpha \cap H^n} : V_\alpha \cap H^n \rightarrow W_\alpha \cap M =: U_\alpha$ gilt $\bigcup_{\alpha \in I} U_\alpha = M$. Eine solche Familie heißt *Atlas* aus Parametrisierungen für M .

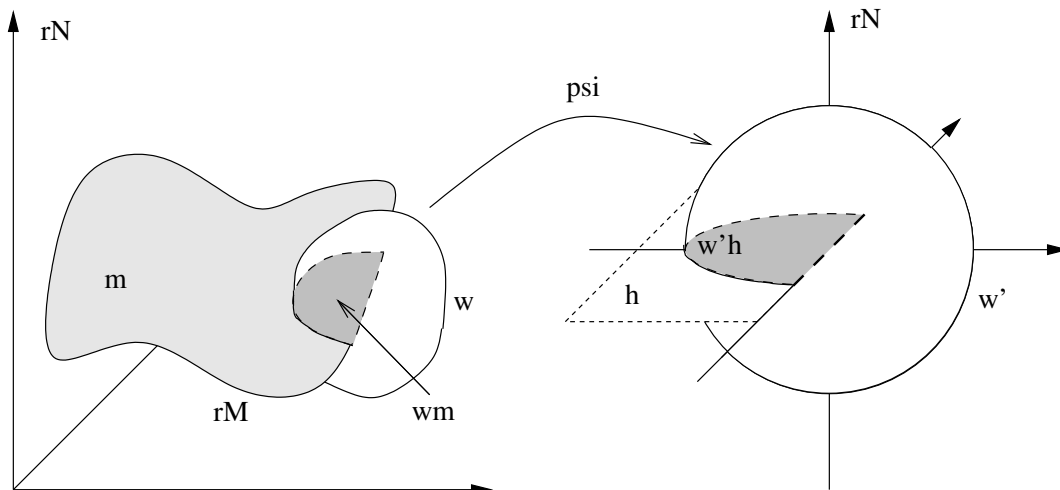


Abbildung 19.4: Untermannigfaltigkeit mit Rand

Lemma 19.4.2. Sei $M \subset \mathbb{R}^N$ eine n -dimensionale Untermannigfaltigkeit mit Rand $\partial M \neq \emptyset$ und mit Atlas $\varphi_\alpha : V_\alpha \cap H^n \rightarrow W_\alpha \cap M =: U_\alpha, \alpha \in I$. Sei $J = \{\beta \in I \mid V_\beta \cap \partial H^n \neq \emptyset\}$. Dann ist ∂M eine Untermannigfaltigkeit der Dimension $n - 1$ ohne Rand und die Abbildungen

$$\varphi_\beta : V_\beta \cap \partial H^n \rightarrow \partial M, \text{ mit } \beta \in J$$

bilden einen Atlas für ∂M . Ist $(\varphi_\alpha, V_\alpha \cap H^n)_{\alpha \in I}$ ein orientierter Atlas für M , so ist auch $(\varphi_\beta, V_\beta \cap \partial H^n)_{\beta \in J}$ ein orientierter Atlas für ∂M .

Bemerkung. Sei $p = \varphi_\alpha(x) \in M$, so ist $T_p M$ durch $(D\varphi_\alpha(x)(e_1), \dots, D\varphi_\alpha(x)(e_n))$ orientiert. Ist $p \in \partial M$, so ist $T_p \partial M$ durch $(D\varphi_\alpha(x)(e_2), \dots, D\varphi_\alpha(x)(e_n))$ orientiert. Diese Orientierung heißt die durch M induzierte Orientierung.

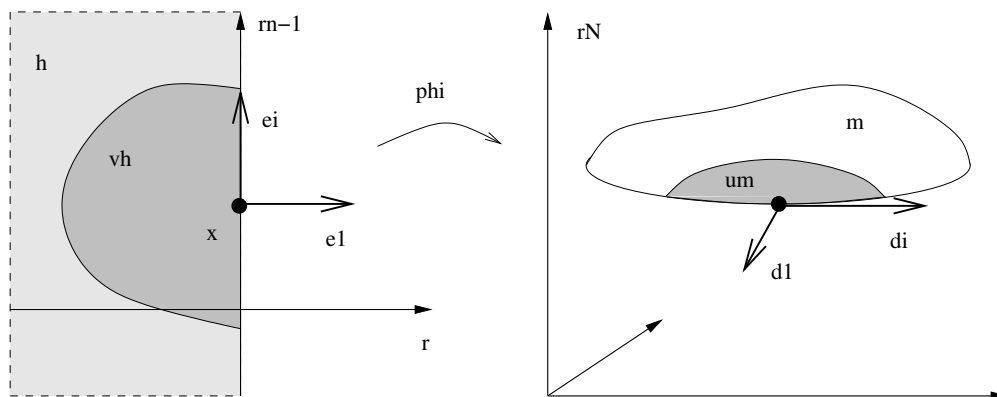


Abbildung 19.5: induzierte Orientierung

Beweis. Dass $(\varphi_\beta, V_\beta \cap \partial H^n)_{\beta \in J}$ ein Atlas für ∂M ist, folgt aus der entsprechenden

Eigenschaft des Atlas für M , denn ∂H^n ist isomorph zum Vektorraum \mathbb{R}^{n-1} . Zu zeigen bleibt, dass dieser orientiert ist.

Seien $\varphi_\alpha : V_\alpha \rightarrow W_\alpha$ und $\varphi_\beta : V_\beta \rightarrow W_\beta$ mit $x \in \partial H^n \cap V_\alpha$ und $y \in \partial H^n \cap V_\beta$, so dass $\varphi_\alpha(x) = \varphi_\beta(y) = p \in \partial M$. Betrachte das Differential $D(\varphi_\beta^{-1} \circ \varphi_\alpha)(x) : \mathbb{R}^n \rightarrow \mathbb{R}^n$. Dieses Differential ist ein Vektorraumisomorphismus mit $(n - 1)$ -dimensionalem invarianten Unterraum $\partial H^n = \{0\} \times \mathbb{R}^{n-1} = \{(0, \lambda_2, \dots, \lambda_n) \mid \lambda_i \in \mathbb{R}\}$, d.h. $D(\varphi_\beta^{-1} \circ \varphi_\alpha)(x)(\partial H^n) = \partial H^n$. Daraus erhalten wir die Darstellung

$$D(\varphi_\beta^{-1} \circ \varphi_\alpha)(x)(e_1) = \lambda_1 e_1 + \sum_{i=2}^n \lambda_i e_i,$$

mit $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ und $\lambda_1 \neq 0$. Wir zeigen:

$$\det D(\varphi_\beta^{-1} \circ \varphi_\alpha)(x)|_{\partial H^n} > 0.$$

Betrachte für $t \in (-\epsilon, 0]$ und $\epsilon > 0$ genügend klein die Kurve

$$c(t) = (c_1(t), \dots, c_n(t)) = \varphi_\beta^{-1} \circ \varphi_\alpha(x + te_1).$$

Dann ist $c_1(t) < 0$ für $t \in (-\epsilon, 0)$ und $c_1(0) = 0$. Da $\dot{c}_1(0) = \lambda_1 \neq 0$ muss $\lambda_1 > 0$ gelten.

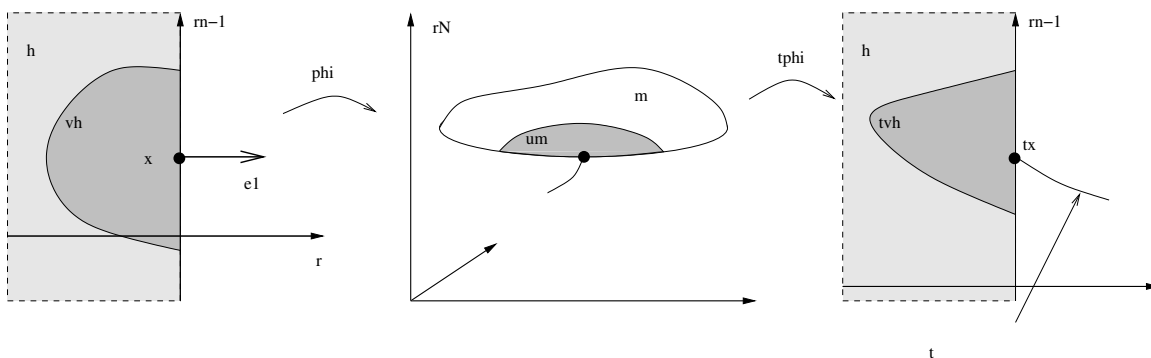


Abbildung 19.6: zum Beweis von Lemma 19.4.2

Sei für $i \geq 2$

$$D(\varphi_\beta^{-1} \circ \varphi_\alpha)(x)(e_i) = \sum_{j=2}^n a_{ij} e_j,$$

so erhalten wir, weil der Atlas für M orientiert ist:

$$0 < \det D(\varphi_\beta \circ \varphi_\alpha)(x) = \det \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ \lambda_2 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ \lambda_n & a_{n2} & \cdots & a_{nn} \end{pmatrix} = \lambda_1 \cdot \det(a_{ij}).$$

Da $\lambda_1 > 0$, folgt

$$\det(a_{ij}) = \det D(\varphi_\beta^{-1} \circ \varphi_\alpha)|_{\partial H^n}(x) > 0.$$

□

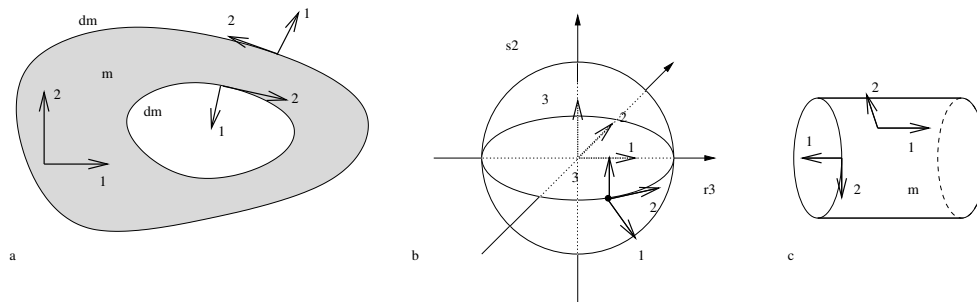


Abbildung 19.7: Beispiele

Beispiele.

Satz 19.4.3 (Satz von Stokes). Sei $M \subset \mathbb{R}^N$ eine n -dimensionale orientierbare Untermannigfaltigkeit mit Rand ∂M . Sei ω eine stetig differenzierbare $(n-1)$ -Form auf M mit kompaktem Träger. Dann gilt:

$$\int_M d\omega = \int_{\partial M} \omega.$$

Dabei benutzen wir auf ∂M die durch M induzierte Orientierung.

Beweis. Wir werden den Beweis in mehreren Schritten durchführen.

1. Sei $M = H^n \subset \mathbb{R}^n$ mit der Standard-Orientierung und ω eine stetig differenzierbare $(n-1)$ -Form auf M mit kompaktem Träger. Wir können aufgrund der Linearität der äußeren Ableitung und des Integral annehmen, dass ω ein Monom ist, d.h.

$$\omega = a_j dx_1 \wedge \dots \wedge \widehat{dx_j} \wedge \dots \wedge dx_n$$

mit $a_j \in C^1(M)$ mit kompaktem Träger. ($\widehat{dx_j}$ bedeutet, dass dx_j ausgelassen wird.) Dann folgt:

$$d\omega = \frac{\partial a_j}{\partial x_j} dx_j \wedge dx_1 \wedge \dots \wedge \widehat{dx_j} \wedge \dots \wedge dx_n = (-1)^{j-1} \frac{\partial a_j}{\partial x_j} dx_1 \wedge \dots \wedge dx_n.$$

Aus der Definition des Integrals ergibt sich:

$$\int_{H^n} d\omega = (-1)^{j-1} \int_{H^n} \frac{\partial a_j}{\partial x_j} d^n x(x),$$

wobei abkürzend $x = (x_1, \dots, x_n)$ und $d^n x = dx_1 \wedge \dots \wedge dx_n$. Da a_j kompakten Träger hat, existiert ein $K > 0$ mit $a_j(x) = 0$ für alle x mit $|x_j| > K$. Mit Fubini und partieller Integration folgt:

- (a) Für $j = 1$ erhalten wir:

$$\int_{H^n} d\omega = \int_{\mathbb{R}^{n-1}} \left(\int_{-\infty}^0 \frac{\partial a_1}{\partial x_1}(x) dx_1 \right) dx_2 \dots dx_n = \int_{\mathbb{R}^{n-1}} a_1(0, x_2 \dots x_n) dx_2 \dots dx_n.$$

(b) Für $j > 1$ erhalten wir:

$$\int_{H^n} d\omega = \int_{\mathbb{R}^{n-1}} \left(\int_{-\infty}^{+\infty} \frac{\partial a_j}{\partial x_j}(x) dx_j \right) dx_1 \dots \widehat{dx_j} \dots dx_n = 0.$$

Betrachte die Parametrisierung $\varphi : \mathbb{R}^{n-1} \rightarrow \partial H^n$ des Randes von H^n mit

$$\varphi(y) = \varphi(y_1, \dots, y_{n-1}) = (0, y_1, \dots, y_{n-1}).$$

Da φ eine lineare Abbildung ist, folgt $D\varphi(y) = \varphi$ und somit $(D\varphi(y)(e_1), \dots, D\varphi(y)(e_{n-1})) = (e_2, \dots, e_n)$. Daher ist $\varphi : \mathbb{R}^{n-1} \rightarrow \partial H^n$ orientierungserhaltend und wir bekommen:

$$\int_{\partial H^n} \omega = \int_{\mathbb{R}^{n-1}} \varphi^* \omega,$$

wobei

$$\varphi^* \omega = (a_j \circ \varphi) d\varphi_1 \wedge \dots \wedge \widehat{d\varphi_j} \wedge \dots \wedge d\varphi_n.$$

Weil $\varphi_1(y_1, \dots, y_{n-1}) = 0 \Rightarrow d\varphi_1 \equiv 0$ und $\varphi_j(y_1, \dots, y_{n-1}) = y_{j-1}$ für $j > 1$, gilt:

$$\varphi^* \omega = \begin{cases} 0 & \text{für } j \neq 1 \\ (a_1 \circ \varphi) dy_1 \wedge \dots \wedge dy_{n-1} & \text{für } j = 1. \end{cases}$$

Daraus folgt:

$$\begin{aligned} \int_{\partial H^n} \omega &= \begin{cases} 0 & \text{für } j \neq 1 \\ \int_{\mathbb{R}^{n-1}} a_1(0, y_1, \dots, y_{n-1}) d\lambda^{n-1}(y) & \text{für } j = 1 \end{cases} \\ &= \int_{H^n} d\omega. \end{aligned}$$

2. Sei nun M eine n -dimensionale Untermannigfaltigkeit mit Rand und orientiertem Atlas $(\varphi_\alpha, V_\alpha \cap H^n)_{\alpha \in I}$. Es sei ω eine stetig differenzierbare $(n-1)$ -Form mit kompaktem Träger und $\text{supp } \omega \subset \varphi_\alpha(V_\alpha \cap H^n)$ für ein $\alpha \in I$. Dann hat $\varphi_\alpha^* \omega$ – und somit auch $d(\varphi_\alpha^* \omega)$ – einen kompakten Träger in $H^n \cap V_\alpha$ und es folgt aus 1.:

$$\begin{aligned} \int_M d\omega &\stackrel{Def.}{=} \int_{V_\alpha \cap H^n} \varphi_\alpha^* d\omega \stackrel{19.2.3}{=} \int_{V_\alpha \cap H^n} d(\varphi_\alpha^* \omega) = \int_{H^n} d(\varphi_\alpha^* \omega) \\ &\stackrel{1.}{=} \int_{\partial H^n} \varphi_\alpha^* \omega = \int_{V_\alpha \cap \partial H^n} \varphi_\alpha^* \omega = \int_{\partial M} \omega. \end{aligned}$$

3. Sei nun M wie oben und ω eine stetig differenzierbare $(n-1)$ -Form mit kompaktem Träger. Dann existiert eine endliche Menge $I' \subset I$ mit $\text{supp } \omega \subset \bigcup_{\alpha \in I'} \varphi_\alpha(V_\alpha)$. Sei $(\rho_\alpha)_{\alpha \in I'}$ eine Zerlegung der Eins angepasst an $\varphi_\alpha(V_\alpha \cap H^n)$. Insbesondere sind die Träger von $\rho_\alpha \omega$ kompakt und enthalten in $\varphi_\alpha(V_\alpha \cap H^n)$. Daher folgt aus 2.:

$$\begin{aligned} \int_M d\omega &= \int_M d \sum_{\alpha \in I'} \rho_\alpha \omega = \sum_{\alpha \in I'} \int_M d(\rho_\alpha \omega) \stackrel{2.}{=} \sum_{\alpha \in I'} \int_{\partial M} \rho_\alpha \omega \\ &= \int_{\partial M} \sum_{\alpha \in I'} \rho_\alpha \omega = \int_{\partial M} \omega. \end{aligned}$$

□

Korollar 19.4.4. Sei $M \subset \mathbb{R}^N$ eine n -dimensionale Untermannigfaltigkeit mit $\partial M = \emptyset$. Ist ω eine stetig differenzierbare $(n - 1)$ -Form mit kompaktem Träger, so gilt:

$$\int_M d\omega = 0.$$

Beweis. Der Beweis folgt direkt aus dem Satz von Stokes. □

Als Anwendung des Satzes von Stokes betrachten wir folgende Beispiele:

Beispiele. Sei $\omega(x) = \sum_{i=1}^n (-1)^{i-1} x_i dx_1 \wedge \dots \wedge \widehat{dx}_i \wedge \dots \wedge dx_n$, so gilt:

$$\begin{aligned} d\omega(x) &= \sum_{i=1}^n (-1)^{i-1} dx_i \wedge dx_1 \wedge \dots \wedge \widehat{dx}_i \wedge \dots \wedge dx_n \\ &= \sum_{i=1}^n dx_1 \wedge \dots \wedge dx_n = n dx_1 \wedge \dots \wedge dx_n. \end{aligned}$$

Sei $A \subset \mathbb{R}^n$ eine kompakte n -dimensionale Untermannigfaltigkeit mit Rand ∂A . Dann folgt:

$$n\lambda^n(A) = \int_A d\omega = \int_{\partial A} \sum_{i=1}^n (-1)^{i-1} x_i dx_1 \wedge \dots \wedge \widehat{dx}_i \wedge \dots \wedge dx_n.$$

Betrachte zum Beispiel $A = B(0, r) \subset \mathbb{R}^n$ und $\partial A = S^{n-1}(r) = \{x \in \mathbb{R}^n \mid \|x\| = r\}$. Für $n = 2$ erhalten wir somit $A = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq r^2\}$ und

$$\lambda^2(A) = \frac{1}{2} \int_{\partial A} x dy - y dx.$$

Sei $\varphi : [0, 2\pi) \rightarrow \partial A$ die Parametrisierung von ∂A mit $\varphi(t) = r(\cos t, \sin t)$. Dann ist

$$\int_{\partial A} x dy - y dx = \int_0^{2\pi} r^2 \cos^2 t + r^2 \sin^2 t dt = 2\pi r^2.$$

Insbesondere folgt: $\lambda^2(A) = \pi r^2$.

19.5 Die Volumenform und die klassischen Integralsätze

Jeder orientierbaren Untermannigfaltigkeit $M \subset \mathbb{R}^N$ der Dimension n lässt sich eine ausgezeichnete n -Form zuordnen. Sei (w_1, \dots, w_n) eine positiv orientierte Basis in $T_p M \subset \mathbb{R}^N$. Der Tangentialraum $T_p M$ ist als Untervektorraum des \mathbb{R}^N bezüglich des auf $T_p M$ eingeschränkten Standardskalarproduktes des \mathbb{R}^N ein Vektorraum mit Skalarprodukt. Nach dem Orthonormalisierungsverfahren von Gram-Schmidt existiert genau eine *ON*-Basis (b_1, \dots, b_n) mit

$$b_k = \sum_{j=1}^k a_{jk} w_j$$

und $a_{kk} \in \mathbb{R}, a_{kk} > 0$. Damit hat die lineare Abbildung $T : T_p M \rightarrow T_p M$ mit $T(w_j) = b_j$ positive Determinante (ihre Matrixdarstellung ist bezüglich der Basis (w_1, \dots, w_n) durch die obere Dreiecksmatrix (a_{jk}) gegeben) und die Basis (b_1, \dots, b_n) ist somit positiv orientiert. Wir definieren nun:

Definition 19.5.1. Sei $M \subset \mathbb{R}^N$ eine orientierbare Untermannigfaltigkeit der Dimension n und sei $p \in M$. Sei (b_1, \dots, b_n) eine positiv orientierte ON -Basis von $T_p M$, so definiere für $v_1, \dots, v_n \in T_p M$:

$$dM(p)(v_1, \dots, v_n) := \det \begin{pmatrix} \langle b_1, v_1 \rangle & \dots & \langle b_n, v_1 \rangle \\ \vdots & & \vdots \\ \langle b_1, v_n \rangle & \dots & \langle b_n, v_n \rangle \end{pmatrix}.$$

Die Differentialform dM heißt *Volumenform*.

Bemerkungen.

(a) Wegen $\langle b_i, b_j \rangle = \delta_{ij}$ folgt

$$dM(p) = b_1^* \wedge \dots \wedge b_n^*.$$

dM ist hierbei nicht als Differential einer $(n-1)$ -Form aufzufassen.

(b) Die Definition hängt nicht von der Wahl der positiv orientierten Basis (b_1, \dots, b_n) ab. Denn ist (c_1, \dots, c_n) eine weitere positiv orientierte Basis und $T : T_p M \rightarrow T_p M$ die lineare Abbildung mit $T(b_i) = c_i$, so ist T eine Isometrie und $|\det T| = 1$. Da (c_1, \dots, c_n) positiv orientiert ist, gilt $\det T > 0 \Rightarrow \det T = 1$. Nun gilt

$$\det \begin{pmatrix} \langle T(b_1), v_1 \rangle & \dots & \langle T(b_n), v_1 \rangle \\ \vdots & & \vdots \\ \langle T(b_1), v_n \rangle & \dots & \langle T(b_n), v_n \rangle \end{pmatrix} = a \cdot dM(p)(v_1, \dots, v_n)$$

für eine Konstante a , denn alternierende n -Formen können sich wegen $\dim \Lambda^n(T_p M) = \binom{n}{n} = 1$ nur um einen Faktor unterscheiden. Setzen wir die Basis (b_1, \dots, b_n) für $v_1, \dots, v_n \in T_p M$ in diese Identität ein, so erhalten wir $a = \det T = 1$. Denn $(\langle T(b_i), b_j \rangle)_{ij}$ ist gerade die Matrixdarstellung von T bezüglich der ON -Basis (b_1, \dots, b_n) . Somit ist die Unabhängigkeit der Definition von $dM(p)$ von der Wahl der positiv orientierten Basis gewährleistet.

Die Volumenform besitzt folgende lokale Darstellung.

Lemma 19.5.2. Sei $M \subset \mathbb{R}^N$ eine orientierbare Untermannigfaltigkeit, dM die Volumenform, $V \subset \mathbb{R}^n$ offen und $\varphi : V \rightarrow M$ eine lokale Parametrisierung mit $\varphi(x) = p \in M$. Dann gilt

$$\varphi^* dM(x) = \sqrt{\det(g_{ij}(x))} dx_1 \wedge \dots \wedge dx_n,$$

wobei die $n \times n$ Matrix $(g_{ij}(x))$ durch

$$g_{ij}(x) = \left\langle \frac{\partial \varphi}{\partial x_i}(x), \frac{\partial \varphi}{\partial x_j}(x) \right\rangle$$

gegeben ist.

Bemerkung. In dieser Darstellung kommt die ON -Basis nicht mehr vor. Dieser Darstellung zeigt auch die Differenzierbarkeit von dM . Für $\det(g_{ij}(x))$ schreiben wir manchmal auch kurz $g(x)$.

Beweis. Es gilt

$$\varphi^*dM(x) = f(x)dx_1 \wedge \dots \wedge dx_n$$

mit

$$f(x) = \det \begin{pmatrix} \langle \frac{\partial \varphi}{\partial x_1}(x), b_1 \rangle & \dots & \langle \frac{\partial \varphi}{\partial x_n}(x), b_1 \rangle \\ \vdots & & \vdots \\ \langle \frac{\partial \varphi}{\partial x_1}(x), b_n \rangle & \dots & \langle \frac{\partial \varphi}{\partial x_n}(x), b_n \rangle \end{pmatrix},$$

wobei (b_1, \dots, b_n) eine beliebige orientierte Basis von T_pM ist. Wegen

$$g_{ij}(x) = \sum_{k=1}^n \left\langle \frac{\partial \varphi}{\partial x_i}(x), b_k \right\rangle \left\langle b_k, \frac{\partial \varphi}{\partial x_j}(x) \right\rangle$$

folgt $\det(g_{ij}(x))_{i,j \leq n} = f^2(x)$ und somit wegen $\det AA^T = (\det A)^2$ die Aussage. □

Es gilt das folgende Kriterium für die Orientierbarkeit einer Untermannigfaltigkeit.

Satz 19.5.3. Sei $M \subset \mathbb{R}^N$ eine n -dimensionale Untermannigfaltigkeit. M ist genau dann orientierbar, falls eine n -Form $\omega \in \Omega^n(M)$ existiert mit $\omega(p) \neq 0$ für alle $p \in M$.

Beweis. Sei M orientierbar, so ist die Volumenform eine nirgends verschwindende n -Form, denn

$$dM(p)(b_1, \dots, b_n) = 1$$

für jede positiv orientierte ON -Basis (b_1, \dots, b_n) .

Sei nun umgekehrt ω eine nirgends verschwindende n -Form, so können wir die Tangentialräume T_pM wie folgt orientieren: Wir nennen eine geordnete Basis $(v_1, \dots, v_n) \in T_pM$ positiv orientiert, falls

$$\omega(p)(v_1, \dots, v_n) > 0.$$

Dies definiert eine Orientierung auf T_pM , denn ist $T : T_pM \rightarrow T_pM$ ein Isomorphismus, so ist $(T(v_1), \dots, T(v_n))$ genau dann positiv orientiert, falls $\det T > 0$, da

$$\omega(p)(T(v_1), \dots, T(v_n)) = \det T \cdot \omega(p)(v_1, \dots, v_n).$$

Weil $(v_i^*(T(v_j)))_{i,j \leq n}$ die Matrixdarstellung von T bezüglich (v_1, \dots, v_n) ist, ist nämlich

$$v_1^* \wedge \dots \wedge v_n^*(T(v_1), \dots, T(v_n)) = \det (v_i^*(T(v_j)))_{i,j \leq n} = \det T = (\det T) v_1^* \wedge \dots \wedge v_n^*(v_1, \dots, v_n)$$

und $\omega(p)$ ist wegen $\dim \Lambda^n(T_pM) = 1$ ein reelles Vielfaches von $v_1^* \wedge \dots \wedge v_n^*$.

Einen Atlas kompatibel mit dieser Orientierung erhält man dann wie folgt:

Ist $V \subset \mathbb{R}^n$ offen und $\varphi : V \rightarrow U \subset M$ eine Parametrisierung und ist

$$\omega(p)(D\varphi(x)(e_1), \dots, D\varphi(x)(e_n)) < 0$$

für ein $x \in V$ (und damit für alle $x \in V$), so betrachte die Spiegelung $S : \mathbb{R}^n \rightarrow \mathbb{R}^n$ mit

$$S(x_1, \dots, x_n) = (-x_1, x_2, \dots, x_n).$$

Ersetze φ durch die Parametrisierung $\tilde{\varphi} = \varphi \circ S : \tilde{V} \rightarrow M$ mit $\tilde{V} =: S^{-1}(V)$, so ist

$$D\tilde{\varphi}(x)(e_1) = D\varphi(S(x))(S(e_1)) = -D\varphi(S(x))(e_1)$$

und $D\tilde{\varphi}(x)(e_j) = D\varphi(S(x))(e_j)$ für $j \geq 2$. Daher ist

$$\omega(p)(D\tilde{\varphi}(x)(e_1), \dots, D\tilde{\varphi}(x)(e_n)) > 0.$$

Durch dieses Verfahren konstruiert man aus einem Atlas einen orientierten Atlas. \square

Definition 19.5.4. Sei $M \subset \mathbb{R}^N$ eine n -dimensionale orientierbare Untermannigfaltigkeit mit Volumenform dM . Sei $f : M \rightarrow \mathbb{R}$ eine messbare Funktion, d.h.

$$f \circ \varphi_\alpha : V_\alpha \rightarrow \mathbb{R}$$

sei Borel-messbar für alle $\alpha \in I$. Dann heißt $f : M \rightarrow \mathbb{R}$ (*Lebesgue-*)*integrierbar*, falls die n -Form $f dM$ integrierbar ist. In diesem Fall heißt

$$\int_M f dM$$

das (*Lebesgue-*)*Integral* von f .

Sei $M \subset \mathbb{R}^{n+1}$ eine n -dimensionale Untermannigfaltigkeit. Eine solche Untermannigfaltigkeit heißt auch *Hyperfläche*.

Satz 19.5.5. Sei $M \subset \mathbb{R}^{n+1}$ eine Hyperfläche. Dann ist M genau dann orientierbar, falls ein stetiges Einheitsnormalenfeld $N : M \rightarrow \mathbb{R}^{n+1}$ existiert, d.h. eine stetige Abbildung mit $N(p) \perp T_p M$ und $|N(p)| = 1$ für alle $p \in M$. Wenn ein solches Feld existiert, so definiert

$$\omega(p)(v_1, \dots, v_n) = \det(N(p), v_1, \dots, v_n)$$

eine nirgends verschwindende n -Form. Diese Form ist dann auch die Volumenform von M .

Beweis. Sei $N : M \rightarrow \mathbb{R}^{n+1}$ ein stetiges Einheitsnormalenfeld, so definiere

$$\omega(p)(v_1, \dots, v_n) = \det(N(p), v_1, \dots, v_n)$$

für $v_j \in T_p M$. Diese Form definiert wegen Satz 19.5.3 eine Orientierung auf M . Ist (b_1, \dots, b_n) eine positiv orientierte *ON*-Basis von $T_p M$, so ist $(N(p), b_1, \dots, b_n)$ eine *ON*-Basis von \mathbb{R}^{n+1} und es gilt

$$\omega(p)(b_1, \dots, b_n) = 1.$$

Da alternierende $(n+1)$ -Formen in \mathbb{R}^{n+1} bis auf einen Vorfaktor eindeutig bestimmt sind, ist $\omega(p)$ die zugehörige Volumenform $dM(p)$. \square

Für die umgekehrte Richtung benötigen wir folgendes Lemma:

Lemma 19.5.6. Seien $v_1, \dots, v_n \in \mathbb{R}^{n+1}$ linear unabhängig und $n \geq 2$. Sei

$$v_1 \times \dots \times v_n := \sum_{i=1}^{n+1} \det(v_1, \dots, v_n, e_i) e_i.$$

Dann folgt:

$$\langle v_1 \times \dots \times v_n, v_j \rangle = 0$$

und

$$\det(v_1, \dots, v_n, v_1 \times \dots \times v_n) > 0.$$

Beweis.

$$\langle v_1 \times \dots \times v_n, v_j \rangle = \sum_{i=1}^n \det(v_1, \dots, v_n, e_i) \langle e_i, v_j \rangle = \det(v_1, \dots, v_n, v_j) = 0.$$

Außerdem ist

$$\det(v_1, \dots, v_n, v_1 \times \dots \times v_n) = \sum_{i=1}^{n+1} \det(v_1, \dots, v_n, e_i) \det(v_1, \dots, v_n, e_i) > 0,$$

denn ist $\det(v_1, \dots, v_n, e_i) = 0$ für alle i , so wäre e_i von v_1, \dots, v_n linear abhängig für alle $i \in \{1, \dots, n+1\}$. Dies ist unmöglich, da (e_1, \dots, e_{n+1}) eine Basis von \mathbb{R}^{n+1} ist. \square

Bemerkung. Ist w ein weiterer normaler Vektor mit $\det(v_1, \dots, v_n, w) > 0$, so ist $w = \lambda v$ mit $\lambda > 0$. Sind also $\|v\| = \|w\|$, so ist $v = w$. Sei $M \subset \mathbb{R}^{n+1}$ eine Hyperfläche so definieren wir ein stetiges Einheitsnormalenfeld wie folgt: Sei $V \subset \mathbb{R}^n$ offen und $\varphi : V \cap H^n \rightarrow M$ eine positiv orientierte Parametrisierung, so definiere für $p = \varphi(x)$

$$N(p) = \frac{\frac{\partial \varphi}{\partial x_1}(x) \times \dots \times \frac{\partial \varphi}{\partial x_n}(x)}{\left\| \frac{\partial \varphi}{\partial x_1}(x) \times \dots \times \frac{\partial \varphi}{\partial x_n}(x) \right\|}.$$

Wegen Lemma 19.5.6 ist $N(p) \perp T_p M$.

Beweis von Satz 19.5.5 (Fortsetzung).

Für jede Parametrisierung $\varphi : V \cap H^n \rightarrow M$ definiert N aus der obigen Bemerkung ein stetiges Einheitsnormalenfeld auf $\varphi(V) \subset M$. Ist M orientierbar, so betrachte zwei positiv orientierte Parametrisierungen $\varphi_\alpha : V_\alpha \cap H^n \rightarrow M$, $\varphi_\beta : V_\beta \cap H^n \rightarrow M$.

Falls $p = \varphi_\alpha(x) = \varphi_\beta(y) \in M$, dann stimmen die mittels φ_α bzw. φ_β definierten Einheitsnormalenfelder $N_\alpha(p)$ bzw. $N_\beta(p)$ wegen $N_\alpha(p) \perp T_p M \perp N_\beta(p)$ und $\|N_\alpha(p)\| = 1 = \|N_\beta(p)\|$ bereits bis eventuell auf ein Vorzeichen überein. Nach Lemma 19.5.6 sind sowohl $\left(\frac{\partial \varphi_\alpha}{\partial x_1}(x), \dots, \frac{\partial \varphi_\alpha}{\partial x_n}(x), N_\alpha(p)\right)$ als auch $\left(\frac{\partial \varphi_\beta}{\partial y_1}(y), \dots, \frac{\partial \varphi_\beta}{\partial y_n}(y), N_\beta(p)\right)$ positive Basen von \mathbb{R}^{n+1} .

Weil $\left(\frac{\partial \varphi_\alpha}{\partial x_1}(x), \dots, \frac{\partial \varphi_\alpha}{\partial x_n}(x)\right)$ und $\left(\frac{\partial \varphi_\beta}{\partial y_1}(y), \dots, \frac{\partial \varphi_\beta}{\partial y_n}(y)\right)$ gleich orientiert sind, müssen auch $N_\alpha(p)$ und $N_\beta(p)$ gleich orientiert sein. Daher ist $N_\alpha(p) = N_\beta(p)$ und N hängt nicht von der Wahl der Parametrisierung ab.

Sei nun $M \subset \mathbb{R}^3$ eine orientierbare 2-dimensionale Untermannigfaltigkeit und N ein Einheitsnormalenfeld. Dann ist die Volumenform von M gegeben durch

$$dM(v_1, v_2) = \det(N, v_1, v_2) = \langle N, v_1 \times v_2 \rangle.$$

Sei $w \in \mathbb{R}^3$, so gilt:

$$\langle w, N \rangle \cdot \det(N, v_1, v_2) = \det(w, v_1, v_2)$$

für alle $v_1, v_2 \in T_p M$. Denn sind v_1, v_2 linear abhängig, so sind beide Seiten gleich 0. Andernfalls schreibe

$$w = \alpha N + \beta v_1 + \gamma v_2.$$

Dann gilt: $\langle w, N \rangle = \alpha$ und $\det(w, v_1, v_2) = \alpha \det(N, v_1, v_2)$. Für $n = 2$ ist "×" gerade das bekannte Kreuzprodukt.

Satz 19.5.7 (Klassischer Integralsatz von Stokes). Sei $M \subset U$ mit $U \subset \mathbb{R}^3$ offen eine 2-dimensionale orientierte Untermannigfaltigkeit mit Rand ∂M und $F : U \rightarrow \mathbb{R}^3$ sei ein differenzierbares Vektorfeld. Dann gilt:

$$\int_M \langle \operatorname{rot} F, N \rangle dM = \int_{\partial M} \omega_F^1 = \int_{\partial M} \langle F, T \rangle dS,$$

wobei T das orientierte Einheitstangentienfeld längs $S = \partial M$ ist.

Beweis. Aus dem allgemeinen Integralsatz von Stokes folgt mit der Bemerkung am Ende von 19.2:

$$\int_M \omega_{\operatorname{rot} F}^2 = \int_M d\omega_F^1 = \int_{\partial M} \omega_F^1.$$

Da

$$\omega_{\operatorname{rot} F}^2(v_1, v_2) = \det(\operatorname{rot} F, v_1, v_2) = \langle \operatorname{rot} F, N \rangle \cdot \det(N, v_1, v_2) = \langle \operatorname{rot} F, N \rangle \cdot dM(v_1, v_2)$$

und $\omega_F^1(v) = \langle F, T \rangle dS(v)$, folgt die Behauptung. \square

Nun wollen wir die Divergenz eines Vektorfeldes F auf einer orientierten Untermannigfaltigkeit definieren. Zunächst folgende Definition.

Definition 19.5.8. Sei $M \subset \mathbb{R}^N$ eine n -dimensionale Untermannigfaltigkeit.

- (a) Ein Vektorfeld $F : M \rightarrow \mathbb{R}^N$ auf $M \subset \mathbb{R}^N$ ist eine Abbildung mit $F(p) \in T_p M$ für alle $p \in M$. F heißt m -mal stetig differenzierbar, falls für jede Parametrisierung $\varphi : V \rightarrow M$, $V \subset \mathbb{R}^n$ offen, die Abbildung $F \circ \varphi : V \rightarrow \mathbb{R}^N$ m -mal stetig differenzierbar ist.
- (b) Sei ω eine k -Form auf M und $F : M \rightarrow \mathbb{R}^N$ ein Vektorfeld, so heißt die $(k-1)$ -Form $i_F \omega$ mit

$$i_F \omega(p)(v_1, \dots, v_{k-1}) = \omega(p)(F(p), v_1, \dots, v_{k-1})$$

das innere Produkt von F und ω .

- (c) Sei $M \subset \mathbb{R}^N$ eine orientierte n -dimensionale Untermannigfaltigkeit mit Volumenform dM und F ein stetig differenzierbares Vektorfeld. Dann heißt die eindeutig bestimmte Funktion $\operatorname{div} F : M \rightarrow \mathbb{R}$ mit

$$d(i_F dM)(p) = \operatorname{div} F dM$$

die Divergenz des Vektorfeldes F auf M .

Bemerkungen.

- (a) Sei $V \subset \mathbb{R}^n$ offen und $\varphi : V \rightarrow M$ eine lokale Parametrisierung, so ist für jedes $x \in V$ das Differential $D\varphi(x) : \mathbb{R}^n \rightarrow T_x M$ ein Vektorraumisomorphismus. Ist nun F ein Vektorfeld auf M , so heißt das Vektorfeld $\varphi^* F : V \rightarrow \mathbb{R}^N$ mit

$$\varphi^* F(x) := (D\varphi(x))^{-1}(F(\varphi(x)))$$

das mittels der Karte φ auf V zurückgeholte Vektorfeld. Sei $(F_1(x), \dots, F_n(x)) = \varphi^*F(x)$, so gilt

$$F(\varphi(x)) = D\varphi(x)(\varphi^*F(x)) = \sum_{i=1}^n F_i(x) \frac{\partial \varphi}{\partial x_i}(x).$$

Also sind die Komponenten von $\varphi^*F(x)$ die Koeffizienten des Vektors $F(\varphi(x))$ bezüglich der Basis $\left(\frac{\partial \varphi}{\partial x_1}(x), \dots, \frac{\partial \varphi}{\partial x_n}(x)\right)$ von $T_{\varphi(x)}M$.

(b) Sei ω eine k -Form auf M , so folgt

$$\varphi^*(i_F\omega) = i_{\varphi^*F}\varphi^*(\omega)$$

auf M . Zum Beweis betrachte $u_1, \dots, u_{n-1} \in \mathbb{R}^n$. Dann gilt für jedes $x \in V$:

$$\begin{aligned} \varphi^*(i_F\omega)(x)(u_1, \dots, u_{n-1}) &= i_F\omega(\varphi(x))(D\varphi(x)(u_1), \dots, D\varphi(x)(u_{n-1})) \\ &= \omega(\varphi(x))(F(\varphi(x)), D\varphi(x)(u_1), \dots, D\varphi(x)(u_{n-1})) \\ &= \omega(\varphi(x))(D\varphi(x)(\varphi^*F(x)), D\varphi(x)(u_1), \dots, D\varphi(x)(u_{n-1})) \\ &= \varphi^*\omega(x)(\varphi^*F(x), u_1, \dots, u_{n-1}) \\ &= i_{\varphi^*F}(\varphi^*\omega)(x)(u_1, \dots, u_{n-1}). \end{aligned}$$

(c) i ist linear in F , d.h. für Vektorfelder F, G auf M und $\lambda, \mu \in \mathbb{R}$ gilt

$$i_{\lambda F + \mu G} = \lambda i_F + \mu i_G.$$

(d) Sei $F = (F_1, \dots, F_n) : V \rightarrow \mathbb{R}^n$ ein Vektorfeld auf der offenen Menge $V \subset \mathbb{R}^n$ und $dx_1 \wedge \dots \wedge dx_n$ die Volumenform bezüglich der Standardorientierung auf \mathbb{R}^n . Dann gilt

$$i_F dx_1 \wedge \dots \wedge dx_n(x) = \sum_{i=1}^n (-1)^i F_i(x) dx_1 \wedge \dots \wedge \widehat{dx_i} \wedge \dots \wedge dx_n$$

und somit

$$d(i_F dx_1 \wedge \dots \wedge dx_n)(x) = \sum_{i=1}^n \frac{\partial F_i}{\partial x_i}(x) dx_1 \wedge \dots \wedge dx_n.$$

Also ist die Divergenz des Vektorfeldes $F : V \rightarrow \mathbb{R}^n$ bezüglich der Standardvolumenform des \mathbb{R}^n gegeben durch

$$\operatorname{div} F(x) = \sum_{i=1}^n \frac{\partial F_i}{\partial x_i}(x).$$

Allgemein erhält man für die Divergenz eines Vektorfeldes bezüglich einer Untermannigfaltigkeit und eines beliebigen Koordinatensystems:

Lemma 19.5.9. Sei $M \subset \mathbb{R}^N$ eine orientierte n -dimensionale Untermannigfaltigkeit mit Volumenform dM und F ein stetig differenzierbares Vektorfeld auf M . Sei $\varphi : V \rightarrow M$ eine lokale Parametrisierung, so gilt

$$\operatorname{div} F(\varphi(x)) = \frac{1}{\sqrt{g(x)}} \sum_{i=1}^n \frac{\partial(\sqrt{g}F_i)}{\partial x_i}(x),$$

wobei $(F_1(x), \dots, F_n(x)) = \varphi^*F(x)$ die lokale Darstellung des Vektorfeldes F bezüglich der Parametrisierung φ ist und $g = \det(g_{ij})$ (vgl. Bem. nach 19.5.2).

Beweis. Nach Definition der Divergenz gilt für $p \in M$:

$$d(i_F dM)(p) = \operatorname{div} F(p) dM(p).$$

Sei $\varphi : V \rightarrow M$ eine lokale Parametrisierung, so folgt aus Lemma 19.2.3 und obiger Bemerkung (a) für $x \in V$:

$$\varphi^*(d(i_F dM))(x) = d(\varphi^*(i_F dM))(x) = d(i_{\varphi^* F}(\varphi^* dM))(x).$$

Wegen Lemma 19.5.2 und obiger Bemerkung (d) erhalten wir:

$$\begin{aligned} i_{\varphi^* F} \varphi^*(dM)(x) &= \sqrt{g(x)} i_{\varphi^* F}(dx_1 \wedge \dots \wedge dx_n)(x) \\ &= \sqrt{g(x)} \sum_{j=1}^n (-1)^j F_j(x) dx_1 \wedge \dots \widehat{dx_j} \dots \wedge dx_n. \end{aligned}$$

Also folgt zusammen mit den obigen Gleichungen

$$\begin{aligned} \operatorname{div} F(\varphi(x)) \sqrt{g(x)} dx_1 \wedge \dots \wedge dx_n &= \varphi^*(\operatorname{div} F dM)(x) = \varphi^* d(i_F dM)(x) \\ &= d(i_{\varphi^* F} \varphi^* dM)(x) = \left(\sum_{i=1}^n \frac{\partial(\sqrt{g} F_i)}{\partial x_i}(x) \right) dx_1 \wedge \dots \wedge dx_n \end{aligned}$$

und daraus die Behauptung. \square

Um den Gaußschen Integralsatz formulieren und beweisen zu können, benötigen wir den Begriff des äußeren Normalenfeldes für eine positiv orientierte berandete Untermannigfaltigkeit. Sei $M \subset \mathbb{R}^N$ eine n -dimensionale Untermannigfaltigkeit mit Rand ∂M , so ist $T_p \partial M \subset T_p M$ ein Unterraum der Dimension $n-1$ von $T_p M$. Man definiert nun:

Definition 19.5.10. Sei $M \subset \mathbb{R}^N$ eine orientierte n -dimensionale Untermannigfaltigkeit mit orientiertem Rand ∂M und $p \in \partial M$. Sei $v_1, \dots, v_{n-1} \in T_p \partial M$ eine positiv orientierte Basis, so heißt der eindeutige Vektor $N(p) \perp T_p \partial M$ mit $\|N(p)\| = 1$, für den $(N(p), v_1, \dots, v_{n-1})$ eine positiv orientierte Basis des $T_p M$ bildet, *äußerer Einheitsnormalenvektor*. Das Vektorfeld $p \mapsto N(p)$ heißt *äußeres Einheitsnormalenfeld* von M .

Bemerkung. Sei $d(\partial M)$ die mit der Orientierung von ∂M kompatible Volumenform so gilt

$$i_N dM(p) = d(\partial M)(p)$$

für alle $p \in \partial M$. Denn ist (b_1, \dots, b_{n-1}) eine positiv orientierte ON -Basis von $T_p \partial M$, so folgt

$$i_N dM(p)(b_1, \dots, b_{n-1}) = dM(p)(N(p), b_1, \dots, b_{n-1}) = 1,$$

denn $(N(p), b_1, \dots, b_{n-1})$ ist eine positiv orientierte ON -Basis von $T_p M$.

Es gilt nun der folgende wichtige Integralsatz.

Satz 19.5.11 (Gaußscher Satz). Sei $M \subset \mathbb{R}^N$ eine orientierte n -dimensionale Untermannigfaltigkeit mit orientiertem Rand ∂M . Sei $N : \partial M \rightarrow \mathbb{R}^N$ das äußere Einheitsnormalenfeld. Dann gilt für jedes stetig differenzierbare Vektorfeld $F : M \rightarrow \mathbb{R}^N$ auf M :

$$\int_M \operatorname{div} F dM = \int_{\partial M} \langle F, N \rangle d(\partial M).$$

Beweis. Aus dem allgemeinen Integralsatz von Stokes und der Definition der Divergenz folgt:

$$\int_M \operatorname{div} F dM = \int_M d(i_F dM) = \int_{\partial M} i_F dM.$$

Betrachte die Zerlegung von F in Normal- und Tangentialanteil

$$F(p) = \langle F(p), N(p) \rangle N(p) + T(p)$$

mit $T(p) \in T_p \partial M$. Dann folgt aufgrund der Linearität von i :

$$i_F dM = i_{\langle F, N \rangle N} dM + i_T dM = \langle F, N \rangle i_N dM,$$

denn die $(n - 1)$ -Form $i_T dM$ verschwindet auf ∂M , weil die Vektoren $T(p), v_1, \dots, v_{n-1} \in T_p \partial M$ linear abhängig sein müssen. Aus Definition 19.5.10 und anschließender Bemerkung folgt somit die Behauptung. \square

Bemerkung. Der Gaußsche Integralsatz erlaubt auch die folgende physikalische Interpretation der Divergenz eines Vektorfeldes. Sei $M \subset \mathbb{R}^N$ eine orientierte m -dimensionale Untermannigfaltigkeit und $p \in M$. Sei $(M_m)_{m \in \mathbb{N}}$ eine Folge von n -dimensionalen Untermannigfaltigkeiten des \mathbb{R}^N mit $M_m \subset M$ und Rand ∂M_m . Wir sagen, dass diese Folge gegen p konvergiert, falls zu jeder Umgebung $U \subset M$ von p ein m_0 existiert mit $M_m \subset U$ für alle $m \geq m_0$. Sei zum Beispiel $(r_m)_{m \in \mathbb{N}} > 0$ eine gegen 0 konvergente Folge, so definiert $M_m := M \cap K(p, r_m)$ eine solche Folge. Es folgt dann aus dem Gaußschen Integralsatz für jedes stetig differenzierbare Vektorfeld $F : M \rightarrow \mathbb{R}^N$ auf M

$$\operatorname{div} F(p) = \lim_{m \rightarrow \infty} \frac{1}{\operatorname{vol} M_m} \int_{M_m} \operatorname{div} F dM_m = \lim_{m \rightarrow \infty} \frac{1}{\operatorname{vol} M_m} \int_{\partial M_m} \langle F, N \rangle d(\partial M_m).$$

Dabei wird der Quotient

$$\frac{1}{\operatorname{vol} M_m} \int_{\partial M_m} \langle F, N \rangle d(\partial M_m)$$

als der Fluss des Vektorfeldes F durch ∂M_m pro Volumeneinheit interpretiert.

Ist zum Beispiel F das Geschwindigkeitsfeld einer in M strömenden Flüssigkeit, so stellt der Quotient die pro Zeiteinheit und Volumeneinheit nach außen strömende Flüssigkeitsmenge dar. Daher heißt $\operatorname{div} F(p)$ auch die *Quellenstärke* des Vektorfeldes F im Punkte p . Das Vektorfeld F heißt *quellenfrei* oder *divergenzfrei*, falls $\operatorname{div} F(p) = 0$ für alle $p \in M$ gilt.

Aus dem Gaußschen Integralsatz lassen sich weitere wichtige Integralformeln ableiten. Sei $M \subset \mathbb{R}^N$ eine orientierte n -dimensionale Untermannigfaltigkeit ohne Rand und $f : M \rightarrow \mathbb{R}$ eine zweimal stetig differenzierbare Funktion. Sei $df(p) : T_p M \rightarrow \mathbb{R}$ das Differential von f (siehe Definition 16.6.1), so ist der Gradient von f das durch

$$df(p)v = \langle \operatorname{grad} f(p), v \rangle$$

definierte Vektorfeld $\operatorname{grad} f$ auf M . Mit

$$\Delta := \operatorname{div} \operatorname{grad} : C^2(M) \rightarrow C^0(M)$$

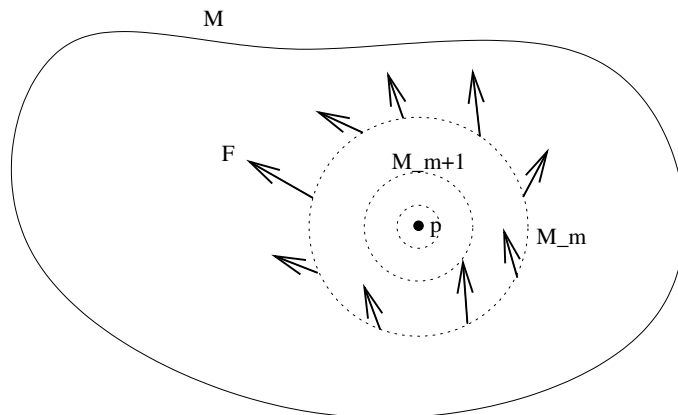


Abbildung 19.8: Anschauliche Bedeutung der Divergenz

bezeichnen wir den *Laplace-Beltrami Operator* auf M . Sei nun $A \subset M$ eine n -dimensionale Untermannigfaltigkeit von \mathbb{R}^N mit der Orientierung von M und orientiertem Rand ∂A . Ist N das äußere Einheitsnormalenfeld auf ∂A , so heißt für jedes $p \in \partial M$

$$\partial_N f := Df(p)N(p) = \langle \text{grad } f(p), N(p) \rangle$$

die *Normalenableitung* von f . Dann gelten die folgenden Greenschen Formeln.

Satz 19.5.12 (*Greensche Formeln*). *Es seien $M \subset \mathbb{R}^N$ eine orientierte Untermannigfaltigkeit, $A \subset M$ eine Untermannigfaltigkeit mit Rand und $f, h : M \rightarrow \mathbb{R}$ zweimal stetig differenzierbare Funktionen. Dann gelten folgende Formeln.*

1. *Greensche Formel 1. Art*

$$\int_A h(x)\Delta f(x) + \langle \text{grad } h(x), \text{grad } f(x) \rangle dA(x) = \int_{\partial A} h(x)\partial_N f(x)d(\partial A)(x)$$

2. *Greensche Formel 2. Art*

$$\int_A (h(x)\Delta f(x) - f(x)\Delta h(x))dA(x) = \int_{\partial A} h(x)\partial_N f(x) - f(x)\partial_N h(x)d(\partial A)(x).$$

Beweis. Betrachte das Vektorfeld $F = h \text{ grad } f$. Dann folgt:

$$\text{div } F(x) = \langle \text{grad } h(x), \text{grad } f(x) \rangle + h(x)\Delta f(x).$$

Der Beweis dieser Identität kann durch Rechnung in lokalen Koordinaten geführt werden. Wenden wir den Integralsatz von Gauß auf das Vektorfeld F an, so erhalten wir die Greensche Formel 1. Art. Zum Beweis der Greenschen Formel 2. Art betrachte die beiden Greenschen Formeln 1. Art die durch Vertauschung von f und g entstehen. Die Subtraktion beider Gleichungen ergibt dann die Greensche Formel 2. Art \square

Index

- $\mathcal{L}^p(X, \mathcal{A}, \mu)$, 474
- $\langle \cdot, \cdot \rangle$, 237
- $\| \cdot \|$, 238, 240
- \oplus , 147, 155
- \perp , 241
- $\mathcal{A} * \mathcal{B}$, 455
- $\mathcal{A} \otimes \mathcal{B}$, 455
- $\Lambda^k(V^*)$, 479
- \overline{A} , 85, 287
- $B(a, \epsilon)$, 279
- Bild f , 159
- $B(M, \mathbb{R})$, 278
- $B(\mathbb{R}, \mathbb{R})$, 289
- $\mathcal{B}(X)$, 424
- \mathbb{C} , 29
- $C^0(I, \mathbb{C})$, 169
- $C^0(M, \mathbb{R})$, 284
- $C^1(U, Y)$, 317
- $C_b^0(M, \mathbb{R})$, 284
- χ_A , 224
- χ_A , 438
- $C^k(D)$, 105
- $C^k(I, \mathbb{C})$, 169
- $C^k(\mathbb{R})$, 144
- $C^k(U, Y)$, 320
- $C^\infty(D)$, 105
- $C^\infty(I, \mathbb{C})$, 169
- $C^\infty(\mathbb{R})$, 144
- D , 158
- $d(x, y)$, 240, 278
- $d_A(x, y)$, 278
- Δf , 515
- δ -Funktional, 249
- diam, 424
- div F , 512
- div v , 303
- ∂M , 502
- $\partial_N f$, 516
- $E_\lambda(f)$, 217
- End(V), 166
- ϵ -Ball, 279
- ϵ -Umgebung, 279
- f^+ , 437
- f^- , 437
- f -invariant, 230
- GL(V), 159
- GL(n, K), 185
- grad $f(a)$, 302
- Grad p , 220
- $H_f(a)$, 322
- id $_V$, 157
- $i_F \omega$, 512
- K -lineare Abbildung, 157
- Kern f , 159
- $K(p, r)$, 287
- $\mathcal{L}_{f,b}$, 160
- $\mathcal{L}(\mathbb{R}^n)$, 421
- $L(V, W)$, 166
- $\mathcal{L}(X, Y)$, 285
- $M(m \times n, K)$, 184
- $M(n, K)$, 184
- M^\perp , 247
- $M(X, \mathcal{A})$, 438
- $M^+(X, \mathcal{A})$, 438
- \mathbb{N} , 20
- $N_p M$, 351
- $O(n, \mathbb{R})$, 345
- $\Omega^k(U)$, 483
- $\Omega_m^k(U)$, 483
- $\Omega_m^k(M)$, 497
- $O(n)$, 262
- $O(V)$, 262
- $P(F, E)$, 150
- $P(K)$, 220
- $P_n(K)$, 144
- pr $_U$, 157
- \mathbb{Q} , 21
- \mathbb{R} , 25

- rotv, 386
- $R([a, b])$, 121, 144
- $\mathbb{R}^{[a, b]}$, 144
- $R([a, b], X)$, 308
- Rang f , 189
- $\mathbb{R}^{\mathbb{N}}$, 144
- $\mathbb{R}^{\mathbb{R}}$, 144
- $SL(n, K)$, 212
- $SL(n, \mathbb{R})$, 345
- $S(M)$, 140
- S_n , 140, 205
- $SO(V)$, 267
- Span, 145
- $SU(V)$, 267
- $\text{Sym}(n, \mathbb{R})$, 345
- $T([a, b])$, 119, 144
- $T([a, b], X)$, 306
- $T_p M$, 350
- $T(X, \mathcal{A})$, 438
- $U(n)$, 262
- $U(V)$, 262
- $v_1 \times \dots \times v_n$, 510
- $V_{\mathbb{C}}$, 257
- V^M , 142
- vol, 415
- V^* , 167, 248
- \mathbb{Z} , 21
- \mathbb{Z}_- , 21
- Überdeckung
 - lokal endliche, 500
- äußere Ableitung, 484
- äußere Produktmaß, 455
- äußeres Einheitsnormalenfeld, 514

- Abbildung, 10
 - K -lineare, 157
 - adjungierte, 249
 - affin lineare, 96, 306
 - alternierende, 205
 - bijektive, 11
 - differenzierbar, 353
 - differenzierbare, 97, 169, 295
 - Grenzwert einer \sim , 86, 282
 - Identität, 157
 - injektive, 11
 - inverse, 12
 - konjugiert lineare, 238
 - kontrahierend, 333
 - lineare, 157
 - Dimensionsformel für lin. Abb., 164
 - Koordinatendarstellung, 186
 - Rang, 189
 - Lipschitz-stetige, 282
 - messbare, 427
 - multilineare, 205, 313
 - normierte, 205
 - Nullabbildung, 157
 - Projektion, 157
 - semilineare, 238
 - stetig differenzierbare, 310, 317
 - stetig fortsetzbare, 85, 283, 295
 - stetige, 97, 282
 - surjektive, 11
 - transponierte, 168
- abelsche (kommutative) Gruppe, 13, 139
- Abelscher Grenzwertsatz, 118
- abgeschlossen
 - \sim e Hülle, 85, 287
 - \sim e Kreisscheibe, 81
 - \sim e Menge, 81, 287
 - \sim er Ball, 287
 - \sim es Intervall, 38, 76
- Ableitung, 93, 169, 296
 - der Umkehrfunktion, 100
 - partielle, 300, 325
 - Richtungsableitung, 299
- Absolutbetrag, 19
- Abstand, 240
 - von Mengen, 424
- abzählbar, 43
- Addition, 97
- adjungiert
 - \sim e Abbildung, 249
 - \sim e Matrix, 252
- Äußeres Lebesguemaß, 415
- äußeres Maß, 413
- affin linear
 - \sim e Abbildung, 96, 306
 - \sim e Approximation, 96, 296
- Algebra, 419
 - Borel- σ -Algebra, 424
 - Produkt- σ -Algebra, 455
- Algorithmus
 - Euklidischer, 222

- Gaußscher, 162, 193
- allgemeine lineare Gruppe, 159
- alternierend, 205
- analytisch, 113, 393
- Antisymmetrie, 16
- Approximation
 - affin lineare, 96, 296
- Arcus Cosinus, 68
- Arcus Sinus, 68
- Assoziativgesetz, 8, 13, 15, 139
- Atlas, 497
 - für berandete Untermfk., 502
 - orientiert, 499
- aufgespannter Unterraum, 145
- äußeres Produkt, 480
- ausgeartet
 - \sim e Form, 274
 - \sim es Skalarprodukt, 248
- Aussage, 5
- Aussageform, 5
- Austauschlemma, 151
- Austauschsatz, 151
- Automorphismus, 158
- Axiome
 - Anordnungsaxiome, 17
 - Axiome der Addition, 13
 - Axiome der Multiplikation, 15
 - Gruppenaxiome, 139
 - Körperaxiome, 13
 - Vollständigkeitsaxiom, 25
- Bahnkurve, 369
- Ball, 279
 - abgeschlossener, 287
- Banachraum, 280
- Banachscher Fixpunktsatz, 333
- Basis, 149
 - duale, 167
 - geordnete, 165
 - numerierte, 165
 - Orthogonalbasis, 242
 - Orthonormalbasis, 242
- Berührungspunkt, 85, 287
- Bernoullische Ungleichung, 22
- Beschleunigung, 98
- beschränkt
 - \sim e Menge, 288
- beschränkte
 - Folge, 35, 45
 - Menge, 25, 27
- Besselsche Ungleichung, 244
- Betrag, 19, 31
- Bild, 10, 159
- Bildmaß, 428
- Bilinearform, 270
- Binomialkoeffizienten, 23
 - verallgemeinerte, 114
- Binomialreihe, 114
- Binomische Formel, 24
- Borel- σ -Algebra, 424
- Borel-Maß, 427
- Cauchy-Kriterium, 49
- Cauchy-Riemannsche Differentialgleichungen, 392
- Cauchy-Schwarzsche Ungleichung, 239, 475
- Cauchyfolge, 40, 280
- Cauchykernel, 398
- Cauchyprodukt, 59
- Cauchysche Integralformel, 398
- Cauchyscher Integralsatz, 395
- Cayley-Hamilton, 227
- charakteristische Funktion, 438
- charakteristisches Polynom, 224
- Cosinus, 61
- Cosinussatz, 241
- Cotangens, 68
- Cramersche Regel, 213
- Dachprodukt, 480
- de Morgansche Regeln, 8
- de Rham'sche Kohomologiegruppe, 490
- Definitionsbereich, 10
- dehnungsbeschränkt
 - \sim e Funktion, 71
- Determinante, 205, 207
- Determinantenmultiplikation, 210
- diagonalisierbarer Endomorphismus, 217
- Differentialform
 - Zurückholen (pull-back), 487
- Diffeomorphismus, 337
 - lokaler, 337
- Differential, 296, 484
- Differentialform

- k -ten Grades, 483
- m -mal stetig differenzierbare, 483
- äußere Ableitung von, 484
- 1. Grades, 379
- auf Untermannigfaltigkeiten, 497
- exakte, 380, 489
- geschlossene, 385, 489
- Pfaffsche, 379
- Windungsform, 379
- Differentialformen
 - Gravitationsform, 380
- Differentialgleichung
 - k -ter Ordnung, 363
 - gewöhnliche, 357
 - Hamiltonsche, 358
 - homogene Schwingungsgleichung, 177
 - lineare, 161, 358
 - homogene, 144, 358
 - lineare $\sim n$ -ter Ordnung, 169
 - Schwingungsgleichung, 161
- Differentialgleichungen
 - Cauchy-Riemannsche, 392
- Differentialoperator, 158
- Differentiation unter dem Integralzeichen, 454
- Differenzenquotient, 93, 97
- Differenzialoperator
 - linearer, n -ter Ordnung, 169
- differenzierbar
 - \sim e Abbildung, 97, 169, 295
 - \sim e Abbildung auf Untermannigf., 353
 - \sim e Funktion, 93
 - \sim e Untermannigfaltigkeit, 342
 - k -mal, 319
 - komplex, 391
 - zweimal, 319
- Dimension, 152
- Dimensionsformel, 153
 - für lineare Abbildungen, 164
- Dirac Funktional, 249
- Diracmaß, 414
- direkte Summe, 147, 155
- Distributivgesetz, 8, 15
- divergente
 - Folge, 34, 45
 - Reihe, 47
- Divergenz, 512
 - eines Vektorfeldes, 303
- Doppelreihen, 56
- Drehachse, 269
- Drehebene, 269
- Drehstreckung, 261
- Drehung, 261
- Drehwinkel, 269
- Dreiecksungleichung, 19, 31, 239, 277, 278
- dual
 - \sim e Basis, 167
- Dualraum, 167, 248
- Durchmesser, 424
- dynamisches System, 370
- Eigenraum, 218
- Eigenvektor, 218
- Eigenwert, 218
 - Multiplizität, 219
- Einbettung, 349
 - isometrische, 253
- Eindeutigkeitssatz, 170
- einfach zusammenhängend, 390
- Einheitsmatrix, 184
- elementare Funktionen, 438
- σ -endlich, 460
- Endomorphismus, 166
 - diagonalisierbarer, 217
 - hermitescher, 254
 - Komplexifizierung, 257
 - nilpotenter, 233
 - normaler, 254
 - orthogonaler, 254
 - schiefadjungierter, 254
 - schiefhermitescher, 254
 - schiefsymmetrischer, 254
 - selbstadjungierter, 254
 - symmetrischer, 254
 - unitärer, 254
- Entwicklungspunkt, 68, 105, 109
- Erzeugendensystem, 145
- Erzeuger, 424
- erzeugter Unterraum, 145
- euklidisch
 - \sim e Norm, 277
 - \sim er Vektorraum, 238
- Euklidischer Algorithmus, 222
- Euler-Lagrange-Gleichung, 373
- Eulersche Formel, 62

- exakt, 489
- Exponentialfunktion, 53
- Extremum
 - globales, 101
 - lokales, 101, 354
 - für norm. Vektorr., 330
- Fatou, Lemma v., 450
- Flusslinie, 369
- Folge, 33, 44, 279
 - beschränkte, 35, 45
 - Cauchyfolge, 40, 280
 - der Partialsummen, 46
 - divergente, 34, 45
 - Grenzwert einer \sim , 33, 45
 - konvergente, 33, 45, 279
 - monoton fallende, 37
 - monoton wachsende, 37
 - Nullfolge, 34
 - Teilfolge, 35, 279
- folgenkompakt
 - \sim e Menge, 288
- Form
 - alternierende k -Form, 479
 - hermitesche, 269
 - Hauptachse, 271
 - indefinite, 330
 - negativ definite
 - auf norm.Vektorr., 330
 - negativ semidefinite
 - auf norm.Vektorr., 330
 - nicht ausgeartete, 274
 - Nullraum, 274
 - positiv definite, 269
 - auf norm.Vektorr., 330
 - positiv semidefinite, 269
 - auf norm.Vektorr., 330
 - quadratische, 330
 - Sesquilinearform, 270
 - Signatur, 274
 - symmetrische, 269
 - Fundamentalmatrix, 271
 - Hauptachse, 271
 - symmetrische Bilinearform, 270
- k -Form (Differentialform), 483
- n -Form, siehe Differentialform, 379
- Formel
 - Binomische Formel, 24
 - Eulersche, 62
 - Polarisationsformeln, 240
 - Taylorsche, 107
- Fourierintegral, 408
- Fourierkoeffizienten, 245
- Fouriertransformierte, 408
- freie Homotopie, 389
- Fubini, 462
- Fundamentalmatrix, 366
 - einer symmetrischen Form, 271
- Fundamentalsatz der Algebra, 83
- Fundamentalsystem, 171, 366
- Funktion
 - analytische, 113
 - dehnungsbeschränkte, 71
 - differenzierbare, 93
 - ganze, 400
 - gleichmäßig stetige, 88
 - Grenzwert einer \sim , 86
 - Heaviside-Funktion, 89
 - linksseitig stetige, 89
 - Lipschitz-stetige, 71
 - Potenzfunktion, 113
 - rechtsseitig stetige, 89
 - Regelfunktion, 121
 - Stammfunktion, 124, 303
 - stetig fortsetzbare, 85, 283, 295
 - stetige, 73
 - Treppenfunktion, 119, 306
- Funktional
 - Dirac Funktional, 249
- Funktionenfolge, 131
 - gleichmäßig konvergente, 131
 - punktweise konvergente, 131
- Gammafunktion, 128
- ganze Funktionen, 400
- Gaußklammer, 28
- Gaußscher Satz, 514
- Gaußscher Algorithmus, 162, 193
- geometrische Reihe, 47
- geordnet
 - \sim e Basis, 165
- gerade
 - Permutation, 207
- geschlossene Differentialform, 385

- Geschwindigkeit, 98, 297
- Geschwindigkeitsfeld, 98
- Geschwindigkeitsvektor, 98
- gewöhnliche Differentialgleichung, 357
- gleichmäßig
 - ~ konvergente Funktionenfolge, 131
- Gleichung
 - lineare Differentialgleichung, 161
 - Euler-Lagrange-, 373
 - homogene, 160
 - homogene Schwingungsgleichung, 177
 - inhomogene, 160
 - lineare, 160
 - Schwingungsgleichung, 161
- Gleichungssystem
 - lineares skalares, 161
- global
 - ~es Extremum, 101
 - ~es Maximum, 101
 - ~es Minimum, 101
- Grad
 - des Polynoms, 69
 - Nilpotenzgrad, 233
- Gradient, 302
- Gradientenvektorfeld, 302
- Graph, 12
- Gravitationsform, 380
- Greensche Formeln, 516
- Grenzwert
 - einer Abbildung, 282
 - einer Folge, 33, 45, 90
 - einer Funktion, 86
 - gleichmäßiger, 131
 - in ∞ und $-\infty$, 90
 - linksseitiger, 89
 - punktweiser, 131
 - rechtsseitiger, 89
 - uneigentlicher, 91
- Gronwall, Lemma v., 362
- Grundbereich, 5
- Gruppe, 139
 - abelsche, 139
 - allgemeine lineare \sim , 159
 - der Permutationen, 140
 - der symmetrischen Matrizen, 345
 - orthogonale, 262, 345
 - Permutationsgruppe, 205
 - spezielle Lineare, 345
 - spezielle lineare, 212
 - spezielle orthogonale, 267
 - spezielle unitäre, 267
 - symmetrische, 140
 - unitäre, 262
- Gruppenaxiome, 139
- Häufungspunkt, 40, 46, 93, 279
- Häufungspunkt (HP), 433
- Höldersche Ungleichung, 474
- Halbnorm, 475
- halboffen
 - ~es Intervall, 76
- Hamilton Prinzip, 375
- Hamiltonsche Differentialgleichungen, 358
- Hamiltonsche Vektorfeld, 358
- harmonische Reihe, 50
- Hauptachse
 - einer hermiteschen/symmetrischen Form, 271
- Hauptachsensystem, 271
- Hauptachsentransformation, 272
- Hauptsatz über Maße, 422
- Hauptteil, 401, 403
- Heaviside-Funktion, 89
- hermitesch
 - ~e Form, 269
 - ~e Matrix, 255, 261
 - ~er Endomorphismus, 254
- Hessematrix, 322
- Hilbertraum, 280
- holomorph, 391
- Homöomorphismus, 349
- homogen
 - ~e Gleichung, 160
 - ~e Schwingungsgleichung, 177
 - ~e lineare Differentialgleichung, 144, 358
 - mit konstanten Koeffizienten, 366
- Homotopie, 387
 - freie, 389
- Horizontalschnitt, 457
- Identität, 157
- Identitätssatz, 71
- imaginäre Einheit, 30
- Imaginärteil, 30

- Immersion, 346
- implizite Funktionen
 - Satz über \sim , 339
- indefinite Form, 330
- Index
 - Spaltenindex, 184
 - Zeilenindex, 184
- Induktion, 20
 - Induktionsanfang, 21
 - Induktionsbehauptung, 21
 - Induktionsschritt, 21
 - Induktionsvoraussetzung, 21
 - verallgemeinerte Induktion, 22
- induziert
 - \sim e Metrik, 278
- induzierte Orientierung, 501
- Infimum, 27
- inhomogen
 - \sim e Gleichung, 160
- inner
 - \sim er Punkt, 286
- innere Produkt, 512
- innerer Punkt, 433
- Integrabilitätsbedingungen, 385
- Integral
 - \sim einer Regelfunktion, 122
 - \sim einer Treppenfunktion, 120
 - auf Untermannigfaltigkeiten, 510
 - von Treppenfunktionen, 441
 - einer Regelfunktion, 308
 - einer Treppenfunktion, 307
 - uneigentliches, 126
 - von k -Formen, 494
- Integralkurve, 357
- Integration
 - partielle, 125
 - Substitutionsregel, 125
- integrierbar, 447, 473
- Intervall
 - abgeschlossenes, 38, 76
 - Charakterisierung, 76
 - halboffenes, 76
 - kompaktes, 76
 - maximales, 360
 - offenes, 33, 76
 - uneigentliches, 76
- Intervallschachtelung, 38
- invariant
 - \sim er Unterraum, 230
- inverse
 - Matrix, 185
- Inverse Funktionen - Theorem, 336
- inverses Element, 13, 15, 139
- isolierte Singularität, 405
- Isometrie, 253, 254
 - Matrix, 261
- isometrisch
 - \sim e Einbettung, 253
 - \sim e Vektorräume, 253
- isomorph, 158
- Isomorphie, 165
- Isomorphismus, 158
- Jacobimatrix, 301
- k -mal differenzierbar, 319
- kanonisch
 - \sim es Skalarprodukt, 238
- kanonische Paarung, 168
- Kardinalität, 23
- Karte, 342
 - für berandete Untermfk., 502
- Kartenwechsel, 342
- kartesisches Produkt, 9
- Kern, 159
- Kettenregel, 99, 305, 393
- Körper, 16
 - algebraisch abgeschlossener, 224
 - archimedisch angeordneter, 28
- kommutativ
 - \sim er Ring, 221
- Kommutativgesetz, 8, 13, 15, 139
- kompakt
 - \sim e Menge, 82, 288
 - \sim es Intervall, 76
- Komplement, 418
 - lineares, 147, 153
- komplex differenzierbar, 391
- komplex konjugierte Zahl, 30
- Komplexifizierung
 - eines Endomorphismus, 257
 - eines Vektorraumes, 257
- Komponentenfunktion, 97
- Komposition, 11

- konjugiert
 - ~ lineare Abbildung, 238
 - ~ lineares Skalarprodukt, 238
 - ~e Flüsse, 371
 - ~e Matrizen, 218
- kontrahierende Abb., 333
- Kontrapositionsgesetz, 11
- konvergent
 - ~e Folge, 279
- konvergente
 - Folge, 33, 45
 - gleichmäßig ~ Funktionenfolge, 131
 - punktweise ~ Funktionenfolge, 131
 - Reihe, 46
- Konvergenzkriterium
 - Majorantenkriterium, 51
 - Quotientenkriterium, 52
 - Vergleichskriterien, 51
 - von Cauchy, 49
 - von Leibniz, 49
 - Wurzelkriterium, 51
- Konvergenzradius, 69
- konvexe Menge, 312
- Koordinaten, 166
- Koordinatendarstellung
 - einer linearen Abbildung, 186
- Koordinatensystem, 166
- Koordinatentransformation, 166, 371
- Kreisring, 401
- Kriterium von Caratheodory, 424
- kritisch
 - ~er Punkt, 102
 - für norm. Vektorr., 330
- kritischer Punkt, 354, 372
- Kurve, 97, 297
 - stetige, 312
- Kurvenintegral, 381
 - komplexes, 394
- Länge, 377
 - von einem Vektor, 238
- Lagrangefunktion, 372
- Lagrangesche Multiplikatoren, 354
- Lagrangesches Funktional, 372
- Laplace-Beltrami Operator, 516
- Laplacescher Entwicklungssatz, 215
- Laurententwicklung, 404
- Laurentreihe, 401
- Lebesgue, Satz v., 451
- Lebesgue-Integral, 447
- Lebesgue-messbar, 421
- Lebesgue-Zahl, 290
- Leibnizkriterium, 49
- Lemma
 - Austauschlemma, 151
 - v. Poincaré, 387
 - von Gronwall, 362
 - von Hahn-Banach, 309
- Lemma von Fatou, 450
- Levi, Satz von B., 447
- Liegruppen, 345
- Limes inferior, 41
- Limes superior, 41
- linear
 - ~e Differentialgleichung, 161
 - ~e Abbildung, 157
 - Dimensionsformel für lin. Abb., 164
 - Rang, 189
 - ~e Differentialgleichung, 358
 - ~e Differentialgleichung n -ter Ordnung, 169
 - ~e Gleichung, 160
 - ~e homogene Differentialgleichung, 144, 358
 - mit konstanten Koeffizienten, 366
 - ~er Differenzialoperator n -ter Ordnung, 169
 - ~er Operator, 169
 - ~es Komplement, 147, 153
 - abhängig, 147
 - beschränktes Vektorfeld, 363
 - unabhängig, 147
- Linearform, 167
- Linearität des Integrals, 449
- Liouville, Satz v., 366
- Lipschitz-stetig, 359
 - ~e Abbildung, 282
 - ~e Funktion, 71
- Lipschitzkonstante, 359
- Lösung
 - Fundamentalsystem, 171
 - partikuläre, 161
- Lösungsmenge der Gleichung, 160
- Logarithmus, 61

- lokal
 - \sim e Parametrisierung, 342
 - \sim er Diffeomorphismus, 337
 - \sim es Extremum, 101, 354
 - \sim es Maximum, 101, 354
 - \sim es Minimum, 101, 354
- lokal endlich, 500
- lokal Lipschitz-stetig, 359
- Mächtigkeit, 23
- Maß
 - äußeres, 413
 - äußeres Lebesguemaß, 415
 - äußeres Produktmaß, 455
 - Borel-, 427
 - Diracmaß, 414
 - Wahrscheinlichkeitsmaß, 414
 - Zählmaß, 414
- Majorante, 51
- Majorantenkriterium, 51
- majorisierte Konvergenz, Satz über, 451
- Matrix, 184
 - adjungierte, 252
 - Determinantenmultiplikation, 210
 - Einheitsmatrix, 184
 - hermitesche, 255, 261
 - Hessematrix, 322
 - inverse, 185
 - Isometrie, 261
 - Jacobimatrix, 301
 - konjugierte Matrizen, 218
 - normale, 261
 - orthogonale, 255, 261
 - Produkt, 185
 - schiefadjungierte, 261
 - schiefhermitesche, 255, 261
 - schiefsymmetrische, 255, 261
 - selbstadjungierte, 261
 - Spaltenrang, 190
 - symmetrische, 255, 261
 - transponierte, 192
 - unitäre, 255, 261
 - Zeilenrang, 190
 - Zeilenstufenform, 194
- Matrixdarstellung
 - einer linearen Abbildung, 186
- maximale Integalkurve, 360
- maximale Intervall, 360
- Maximum, 28
 - globales, 101
 - lokales, 101, 354
 - für norm.Vektorr., 330
- Maximumsnorm, 277
- Maß, 421
- Maßraum, 422
- Menge, 5
 - abgeschlossene, 81, 287
 - Abstand v. \sim en, 424
 - abzählbar unendliche, 43
 - abzählbare, 43
 - beschränkte, 288
 - endliche, 23
 - folgenkompakte, 288
 - Indexmenge, 9
 - induktive Teilmenge des \mathbb{R} , 20
 - kompakte, 82, 288
 - konvexe, 312
 - leere, 6
 - messbare, 419
 - nach oben beschränkte, 25
 - nach unten beschränkte, 27
 - offene, 286
 - Potenzmenge, 8
 - Teilmenge, 7
 - überabzählbare, 43
 - wegzusammenhängende, 312
- Mengengesetze, 8
- Mengenoperationen
 - Differenz, 6
 - Durchschnitt, 6
 - Vereinigung, 6
- messbar, 422, 473
 - \sim e Abbildung, 427
 - \sim e Menge, 417
- messbare Mengen, 419
- Messraum, 419
- Metrik, 240, 278
 - durch eine \sim induz. Ball, 279
 - induzierte, 278
- metrisch
 - \sim er Raum, 240, 278
 - vollständiger, 280
 - zusammenhängender \sim er Raum, 312
- Minimalpolynom, 229

- Minimum, 28
 - globales, 101
 - lokales, 101, 354
 - für norm.Vektorr., 330
- Minkowskische Ungleichung, 475
- Minorante, 51
- Mittelwertsatz
 - der Differentialrechnung, 102
 - der Integralrechnung, 310
 - der Integralrechnung für Kurven, 310
 - für reellwertige Funktionen, 306
 - verallgemeinerter, 104
- Monom, 149
- monotone Konvergenz, Satz v. d., 447
- Monotonie
 - monoton fallende Folge, 37
 - monoton fallende Funktion, 18
 - monoton wachsende Folge, 37
 - monoton wachsende Funktion, 18
 - Monotonie der Addition, 17
 - Monotonie der Multiplikation, 17
- Monotonie des Integrals, 449
- multilinear, 205
 - \sim e Abbildung, 313
- Multiplikation
 - skalare, 97, 141
- Multiplizität, 219, 223
- Nebenteil, 401, 403
- negativ definit
 - \sim e Form
 - auf norm.Vektorr., 330
 - \sim er Unterraum, 274
- negativ orientiert, 394
- negativ semidefinit
 - \sim e Form
 - auf norm.Vektorr., 330
- neutrales Element, 13, 15, 139
- nilpotent
 - \sim er Endomorphismus, 233
- Nilpotenzgrad, 233
- Norm, 240, 277
 - euklidische, 277
 - induz. durch ein Skalarprodukt, 277
 - Maximumsnorm, 277
 - Operatornorm, 285
 - Supremumsnorm, 121, 278
- normal
 - \sim e Matrix, 261
 - \sim er Endomorphismus, 254
- Normalenableitung, 516
- Normalraum, 351
- normiert, 205
 - \sim er Vektorraum, 240, 277
- Nullabbildung, 157
- nullhomotop, 390
- Nullraum
 - einer Form, 274
- Nullstelle, 405
 - Vielfachheit der \sim , 84
- Nullstellenordnung, 405
- numeriert
 - \sim e Basis, 165
- obere Einhüllende, 435
- offen
 - \sim e Menge, 286
 - \sim es Intervall, 33, 76
- OG-Basis, 242
- OG-System, 242
- ON-Basis, 242
- ON-System, 242
- Operator, 157
 - Differentialoperator, 158
 - linearer, 169
- Operatornorm, 285
- Ordnung
 - Übereinstimmung in der n -ten \sim , 108
 - Übereinstimmung in der ersten \sim , 96
 - einer Singularität, 405
- Ordnungsrelation, 16
 - totale, 16
- orientierbare Untermannigfaltigkeit, 499
- Orientierung
 - auf Vektorräumen, 501
 - einer Kurve in \mathbb{C} , 394
 - induzierte, 501
- orientierungserhaltend, 495
- orientierungsumkehrend, 495
- orthogonal
 - \sim e Gruppe, 262, 345
 - \sim e Matrix, 255, 261
 - \sim e Unterräume, 241
 - \sim e Vektoren, 241

- ~e unitäre Gruppe, 267
- ~er Endomorphismus, 254
- Orthogonalbasis, 242
- Orthogonalraum, 247
- Orthogonalsystem, 242
- Orthonormalbasis, 242
- Orthonormalisierungsverfahren von Gram-Schmidt, 246
- Orthonormalsystem, 242
- Paarung
 - kanonische, 168
- parameterabhängige Integrale, 453
- Parametrisierung, 346
- Parametrisierung, lokale, 342
- partiell
 - ~e Ableitung, 300, 325
 - ~e Integration, 125
 - auf Banachräumen, 328
- partikuläre Lösung, 161
 - Variation der Konstanten, 181
- Pascalsches Dreieck, 23
- Permutation, 140
 - gerade, 207
 - ungerade, 207
- Permutationsgruppe, 205
 - Transposition, 206
- Pfaffsche Form, 379
- Picard-Lindelöf, Satz v., 359
- Poincaré, Lemma v., 387
- Poincaré, Lemma v., 492
- Polarisationsformeln, 240
- Polarkoordinaten, 81
- Polstelle, 405
- Polstellenordnung, 405
- Polynom, 69, 144
 - charakteristisches, 224
 - Grad des \sim s, 69, 220
 - komplexes, 83
 - Minimalpolynom, 229
 - reelles, 83
 - Teiler, 222
 - trigonometrisches, 245
 - Zerlegung in Linearfaktoren, 84
- Polynomdivision, 221
- positiv definit
 - ~e Form, 269
 - auf norm.Vektorr., 330
 - ~er Unterraum, 274
- positiv orientiert, 394
- positiv semidefinit
 - ~e Form, 269
 - auf norm.Vektorr., 330
- Potential, 380
 - eines Vektorfeldes, 303
- Potenzfunktion, 113
- Potenzreihe, 68
- Prinzip der kleinsten Wirkung, 375
- Prinzip der vollständigen Induktion, 20
- Prinzip von Cavalieri, 462
- Produkt
 - äußeres \sim von 1-Formen, 480
 - der Matrizen, 185
- Produkt- σ -Algebra, 455
- Produktregel, 99, 304
- Projektion, 157
- Punkt
 - innerer, 286
 - kritischer, 102
- punktweise
 - \sim konvergente Funktionenfolge, 131
 - \sim r Grenzwert, 131
- quadratische Form, 330
- quasiintegrierbar, 448
- Quotientenkriterium, 52
- Quotientenregel, 99
- Rand
 - von Untermannigfaltigkeiten, 502
- Rang
 - einer linearen Abbildung, 189
 - Spaltenrang einer Matrix, 190
 - Zeilenrang einer Matrix, 190
- Raum
 - Banachraum, 280
 - Dualraum, 167, 248
 - Eigenraum, 218
 - Hilbertraum, 280
 - metrischer, 240, 278
 - vollständiger, 280
 - Unterraum
 - invarianter, 230
- Realteil, 30

- Rechenregeln
 - für differenzierbare Abbildungen, 304
 - für differenzierbare Funktionen, 98
 - für konvergente Folgen, 36, 37, 45
 - für konvergente Reihen, 48
- Rechteckmengen, 455
- reelle Einheit, 30
- Reflexivität, 16
- Regel
 - Cramersche, 213
- Regelfunktion, 121, 308
- regulärer Wert, 344
- Reihe, 46
 - absolut konvergente, 50
 - alternierende, 49
 - Binomialreihe, 114
 - divergente, 47
 - Doppelreihe, 56
 - geometrische, 47
 - harmonische, 50
 - konvergente, 46
 - Laurentreihe, 401
 - Potenzreihe, 68
 - Taylorreihe, 109
- Reparametrisierung, 382, 495
- Residuensatz, 406
- Residuum, 405
- Restglied, 106
 - von Cauchy, 107
 - von Lagrange, 107
 - von Schlömilch, 107
- Richtungsableitung, 299
- Riemannscher Hebbarkeitssatz, 401
- Riesz-Fischer, 477
- Rieszscher Darstellungssatz, 248
- Ring
 - kommutativer, 221
- Rotation, 386
- Rotationsfläche, 347
- Sattelpunkt, 332
- Satz
 - über Maße, 422
 - über implizite Funktionen, 339
 - Abelscher Grenzwertsatz, 118
 - Austauschsatz, 151
 - Banachscher Fixpunktsatz, 333
 - Cauchyscher Integralsatz, 395
 - Cosinussatz, 241
 - Differentiation unter dem Integralzeichen, 454
 - Eindeutigkeitssatz, 170
 - Fundamentalsatz der Algebra, 83
 - Hauptsatz der Differential- und Integralrechnung, 123
 - verallgemeinerter, 309
 - Identitätssatz, 71
 - Inverse Funktionen - Theorem, 336
 - Laplacescher Entwicklungssatz, 215
 - Laurententwicklung, 404
 - Lemma von Poincaré, 492
 - Linearität des Integrals, 449
 - Mittelwertsätze, *siehe* Mittelwertsatz
 - Monotonie des Integrals, 449
 - Prinzip v. Cavalieri, 462
 - Residuensatz, 406
 - Riemannscher Hebbarkeitssatz, 401
 - Riesz-Fischer, 477
 - Rieszscher Darstellungssatz, 248
 - Stetige Abhängigkeit des Integrals von einem Parameter, 453
 - Trägheitssatz von Sylvester, 274
 - v. d. majorisierten Konvergenz, 451
 - v. d. monotonen Konvergenz, 447
 - v. Fubini, 462
 - v. Gauß, 514
 - v. Lebesgue, 451
 - v. Levi, 447
 - v. Stokes, 505
 - v. Stokes (klassisch), 512
 - Variation der Konstanten, 181, 367
 - von Bolzano-Weierstraß, 39, 46
 - von Cauchy-Schwarz, 239
 - von Cayley-Hamilton, 227
 - von Hahn-Banach (Lemma), 309
 - von Heine-Borel, 291
 - von Liouville, 366, 400
 - von Picard-Lindelöf, 359
 - von Pythagoras, 241
 - von Rolle, 102
 - von Schwarz, 323
 - von Taylor, 107
 - Weierstraßscher Approximationssatz, 134
 - Zwischenwertsatz, 65, 77

- schiefadjungiert
 - \sim e Matrix, 261
 - \sim er Endomorphismus, 254
- schiefhermitesch
 - \sim e Matrix, 255, 261
 - \sim er Endomorphismus, 254
- schiefsymmetrisch
 - \sim e Matrix, 255, 261
 - \sim er Endomorphismus, 254
- Schranke
 - obere, 25
 - kleinste, 25
 - untere, 27
 - größte, 27
- Schwingungsgleichung, 161
- selbstadjungiert
 - \sim e Matrix, 261
 - \sim er Endomorphismus, 254
- semilinear
 - \sim e Abbildung, 238
 - \sim es Skalarprodukt, 238
- Sesquilinearform, 270
- σ -Algebra, 419
- Signatur
 - einer hermiteschen/symmetrischen Form, 274
- Singularität, 405
- Sinus, 61
- Skalar, 141
- skalare Multiplikation, 141
- Skalarprodukt, 237
 - kanonisches, 238
 - konjugiert lineares, 237
 - nicht ausgeartetes, 248
 - Norm, induz. durch ein \sim , 277
 - semilineares, 237
 - Standardskalarprodukt, 238
- Spaltenindex, 184
- Spaltenvektoren, 184
- Span, 145
- speziell
 - \sim e unitäre Gruppe, 267
- spezielle Lösung
 - Variation der Konstanten, 181
- spezielle lineare Gruppe, 212, 345
- Spur, 224, 269
- Stammform, 489
- Stammfunktion, 124, 303, 309, 380
- Standardorientierung, 501
- Standardskalarprodukt, 238
- stereographische Projektionen, 343
- sternförmige Menge, 386
- stetig
 - \sim e Abbildung, 97, 282
 - \sim e Fortsetzung, 85, 283, 295
 - \sim e Funktion, 73
 - \sim e Kurve, 312
 - \sim er Weg, 312
 - differenzierbar, 310, 317
 - Lipschitz-, 359
 - lokal, 359
- Stokes, 512
- Stokes, Satz v., 505
- Summationsindex, 47
- Summe
 - der Unterräume, 146
 - direkte \sim von Unterräumen, 147, 155
- Superpositionsprinzip, 144
- Supremum, 25
- Supremumsnorm, 121, 131, 278
- symmetrisch
 - \sim e Form, 269
 - \sim e Matrix, 255, 261
 - \sim er Endomorphismus, 254
- symmetrische Gruppe, 140
- System
 - Orthogonalsystem, 242
 - Orthonormalsystem, 242
- Tangens, 68
- Tangente, 96
- Tangentialraum, 350
- Tangentialvektor, 98, 297
- Tangentialvektoren, 350
- Taylorpolynom, 105
- Taylorreihe, 109
- Taylorsche Formel, 107
 - für Kurven, 328
 - für norm. Vektorräume, 327
- Teiler, 222
- Teilfolge, 35, 279
- Teilraum, 143
- Träger, 499
- Trägheitssatz von Sylvester, 274

- Trajektorie, 369
- Transformationsformel, 468
- Transitivität, 16
- Translation, 428
- transponiert
 - \sim e Abbildung, 168
 - \sim e Matrix, 192
- Transposition, 206
- Treppenfunktion, 119, 306, 438
- trigonometrisch
 - \sim es Polynom, 245
- trivial
 - \sim e Untergruppe, 140
 - \sim er Unterraum, 143
- überabzählbar, 43
- Überdeckung, 288
 - Lebesgue-Zahl der \sim , 290
- Umgebung, 33, 44, 279
- Umkehrfunktion, 79
- Umlaufzahl, 396
- Umordnung, 55
- uneigentlich
 - \sim er Grenzwert, 90
 - \sim es Integral, 126
 - \sim es Intervall, 76
- unendlich dimensional, 152
- ungerade
 - Permutation, 207
- Ungleichung
 - Besselsche, 244
 - Cauchy-Schwarzsche, 239
 - Dreiecksungleichung, 239
- unitär
 - \sim e Gruppe, 262
 - \sim e Matrix, 255, 261
 - \sim er Endomorphismus, 254
 - orthogonale \sim e Gruppe, 267
 - spezielle \sim e Gruppe, 267
- untere Einhüllende, 435
- Untergruppe, 140
 - triviale, 140
- Untermannigfaltigkeit, 342
 - orientierte, 499
- Untermannigfaltigkeit mit Rand, 502
- Unterraum, 143
 - aufgespannter, 145
 - erzeugter, 145
 - invarianter, 230
 - negativ definit, 274
 - orthogonale Unterräume, 241
 - positiv definit, 274
 - Summe der Unterräume, 146
 - trivialer, 143
- Urbild, 10
- Variation, 372
- Variation der Konstanten, 181, 367
- Variationsvektorfeld, 372
- Vektor, 97
 - Eigenvektor, 218
 - Länge, 238
 - orthogonale Vektoren, 241
 - Spaltenvektor, 184
 - Tangentialvektor, 297
 - Winkel zwischen Vektoren, 241
 - Zeilenvektor, 184
- Vektorraum
 - Banachraum, 280
- Vektorfeld, 303
 - C^k -Vektorfeld, 357
 - auf Untermfk., 512
 - Divergenz, 303
 - Gradientenvektorfeld, 302
 - Hamiltonsches, 358
 - Integralkurve, 357
 - linear beschränktes, 363
 - Potential, 303
- Vektorraum, 141
 - Dualraum, 167, 248
 - euklidischer, 238
 - Hilbertraum, 280
 - isometrisch, 253
 - isomorpher, 158
 - Komplexifizierung, 257
 - normierter, 240, 277
 - Orthogonalraum, 247
 - Skalarprodukt, 237
 - unitärer, 238
 - Unterraum
 - orthogonale Unterräume, 241
- Vergleichskriterien, 51
- Vertikalschnitt, 457
- vollständig

~er metrischer Raum, 280
Vollständigkeit, 25
Volumenform, 495, 508
Wahrscheinlichkeitsmaß, 414
Weg
 stetiger, 312
Weierstraßscher Approximationssatz, 134
Wertebereich, 10
wesentliche Singularität, 405
Windungsform, 379
Windungszahl, 396
Winkel
 zwischen Vektoren, 241
Wurzel
 n -te, 27
 Quadratwurzel, 19, 26
Wurzelkriterium, 51
Zählmaß, 414
Zeilenindex, 184
Zeilenstufenform
 der Matrix, 194
Zeilenvektoren, 184
Zerlegung, 377
Zerlegung der Eins, 500
zusammenhängend, 312
 einfach, 390
zweimal differenzierbar, 319
Zwischenwertsatz, 65, 77