



UNIVERSITÄT
HEIDELBERG
**SEMPER
APERTUS**

SKRIPT ZUR VORLESUNG
HÖHERE MATHEMATIK III
FÜR DAS
STUDIUM DER PHYSIK

gehalten an der
Universität Heidelberg
im
Wintersemester 2016/17
von
Johannes Walcher

INHALTSVERZEICHNIS

7	SYMMETRIEN	115
	§ 18 Multi-lineare Algebra	115
	§ 19 Gruppen und ihre Darstellungen	125
	§ 20 Euklidische und unitäre Räume	135
8	INTEGRATIONSTHEORIE	145
	§ 21 Integration im \mathbb{R}^n	145
	§ 22 Eigenschaften des Lebesgue-Integrals	162
	§ 23 Differentialformen und der Stokessche Satz	176
9	FUNKTIONALE	195
	§ 24 Hilbert-Räume	195
	§ 25 Lineare Operatoren	206
	§ 26 Sturm-Liouville-Theorie	217
	§ 27 Fourier-Transformation	217

INHALTSVERZEICHNIS

KAPITEL 7

SYMMETRIEN

Diese Vorlesung ist die Fortsetzung der **Höheren Mathematik 2** und richtet sich an Studierende der Physik im dritten Semester. Das Hauptziel ist die Vermittlung der mathematischen Grundlagen für die “moderne” (im Unterschied zur klassischen) theoretischen Physik. Dazu werden zunächst einige Ergebnisse der Linearen Algebra wiederholt, auf die wir uns bereits im letzten Semester regelmässig, und manchmal etwas unreflektiert, berufen haben, und an den Stellen weiter ausgebaut, die in der Quantenmechanik eine ausgezeichnete Rolle spielen. Gründlich motiviert kehren wir im Anschluss daran zur Analysis zurück. Wir entwickeln die Integrationstheorie im \mathbb{R}^n , besprechen Differentialformen und Integralsätze, Hilberträume und die Fourier-Transformation sowie im Rahmen der Möglichkeiten partielle Differentialgleichungen. Die Überlegungen aus dem Sommer zum Verhältnis zu den theoretischen Physik-Vorlesungen behalten natürlich weiter ihre Gültigkeit.

Für Kommentare, auch Fehlermeldungen, zum Skript, per Email an walcher@uni-heidelberg.de, bin ich sehr dankbar. Man beachte aber, dass dieses Skript keinen Anspruch auf Abgeschlossenheit erhebt: Viele Beweise werden bestenfalls angedeutet, mancher Begriff nur indirekt definiert.

· Xyper-Referenzen wie etwa hier zum Kapitel 6 (Funktionentheorie), sollten spätestens dann funktionieren, wenn Sie das **Skript des Sommersemester** ins gleiche Verzeichnis geladen haben.

· Zur Homepage der Vorlesung: www.mathi.uni-heidelberg.de/~walcher/teaching/wise1617/hoemaphys/

§ 18 Multi-lineare Algebra

Sei \mathbb{K} ein Körper (s. Def. 1.1).

Definition 18.1. Ein Vektorraum über \mathbb{K} ist eine Menge V mit einem ausgezeichneten Element $0 \in V$ und zwei binären Verknüpfungen,

$$\begin{aligned} \text{die Addition: } & V \times V \ni (v_1, v_2) \mapsto v_1 + v_2 \in V \\ \text{und die Skalarmultiplikation: } & \mathbb{K} \times V \ni (\lambda, v) \mapsto \lambda \cdot v := \lambda v \in V \end{aligned} \tag{18.1}$$

welche die folgenden Rechenregeln erfüllen:

$$\begin{aligned}
 \text{Kommutativität: } & \forall v_1, v_2 \in V : v_1 + v_2 = v_2 + v_1, \\
 \text{Assoziativität: } & \forall v_1, v_2, v_3 \in \mathbb{K} : (v_1 + v_2) + v_3 = v_1 + (v_2 + v_3) \\
 & \forall v \in V, \lambda_1, \lambda_2 \in \mathbb{K} : \lambda_2 \cdot (\lambda_1 \cdot v) = (\lambda_2 \cdot \lambda_1) \cdot v \\
 \text{Neutrale Elemente: } & \forall v \in \mathbb{K} : v + 0 = v, 1 \cdot v = v \\
 \text{Distributivgesetze: } & \forall v_1, v_2 \in V, \lambda \in \mathbb{K} : \lambda \cdot (v_1 + v_2) = \lambda \cdot v_1 + \lambda \cdot v_2 \\
 & \forall v \in V, \lambda_1, \lambda_2 \in \mathbb{K} : (\lambda_1 + \lambda_2) \cdot v = \lambda_1 \cdot v + \lambda_2 \cdot v
 \end{aligned} \tag{18.2}$$

derart, dass die Gleichung

$$v + x = 0 \tag{18.3}$$

für alle $v \in V$ eine (automatisch eindeutige) Lösung $x \in V$ besitzt.

Übungsaufgabe: $0 \cdot v = 0 \in V \forall v \in V$. Für den Rest der Wiederholung benutzen wir im wesentlichen Prosa.

Bemerkungen/Beispiele 18.2. · In der Kursvorlesung betrachtet man Vektorräume über allgemeinen Körpern. Für die Physik sind normalerweise nur \mathbb{Q} , \mathbb{R} und \mathbb{C} wichtig.

· Für jedes $n \in \mathbb{N}_0 = \{0, 1, 2, \dots\}$ ist das n -fache kartesische Produkt, d.h. die Menge der geordneten n -Tupel

$$\mathbb{K}^n = \underbrace{\mathbb{K} \times \mathbb{K} \times \dots \times \mathbb{K}}_{n \text{ mal}} = \{(x^1, \dots, x^n)^T \mid x^i \in \mathbb{K}\} \tag{18.4}$$

mit komponentenweiser Addition und Skalarmultiplikation ein Vektorraum. Dabei ist $\mathbb{K}^0 := \{0\}$, der “triviale” Vektorraum.

· Für eine beliebige Menge M ist die Menge der Abbildungen

$$\mathcal{F}(M, \mathbb{K}) = \{f : M \rightarrow \mathbb{K}\} \tag{18.5}$$

mit punktweise Addition und Skalarmultiplikation ein \mathbb{K} -Vektorraum. ($\mathcal{F}(\emptyset, \mathbb{K}) \cong \{0\}$.)

· Ein *Untervektorraum* eines Vektorraums ist eine nicht-leere Teilmenge $W \subset V$, welche unter Addition und Skalarmultiplikation abgeschlossen ist. W ist dann selbst ein \mathbb{K} -Vektorraum.

· Besitzt etwa M in (18.5) weitere Struktur, so sind häufig diejenigen Abbildungen, die mit dieser Struktur verträglich sind, Untervektorräume von $\mathcal{F}(M, \mathbb{K})$. Wichtige Beispiele aus der Höheren Mathematik 2 sind Vektorräume stetiger und differenzierbarer Funktionen.

· Ein anderes Beispiel, welches gleich noch eine wichtige Rolle spielen wird, ist der Unterraum der Abbildungen mit *endlichem Träger*

$$\begin{aligned}
 \mathcal{F}^{\text{fin}}(M, \mathbb{K}) & := \{f : M \rightarrow \mathbb{K} \mid f(m) = 0 \text{ ausser für endlich viele } m\} \\
 & = \{f : M \rightarrow \mathbb{K} \mid |\text{supp}(f)| = \#\{m \in M \mid f(m) \neq 0\} < \infty\}
 \end{aligned} \tag{18.6}$$

§ 18. MULTI-LINEARE ALGEBRA

Bei der Untersuchung der Struktur von Vektorräumen kommen die folgenden Konzepte auf.

· Für eine beliebige Teilmenge $A \subset V$ eines Vektorraums ist die lineare Hülle die Menge

$$\text{span}(A) = \left\{ \sum_{i=1}^r \lambda_i v_i \mid r \in \mathbb{N}_0, \lambda_i \in \mathbb{K}, v_i \in A \right\} \quad (18.7)$$

aller endlichen (!) Linearkombinationen. (Für $r = 0$ ordnet man der leeren Summe den Wert 0 zu.) Offensichtlich ist $\text{span}(A)$ ein Untervektorraum von V .

· Eine Teilmenge $A \subset V$ heisst *Erzeugendensystem* von V , wenn $\text{span}(A) = V$.

· Eine Teilmenge $A \subset V$ eines Vektorraums heisst *linear unabhängig*, wenn für jede endliche Teilmenge $\{v_1, \dots, v_r\} \subset A$ ($r \in \mathbb{N}$) gilt, dass aus $\lambda_i \in \mathbb{K}$ für $i = 1, \dots, r$ und

$$\sum_{i=1}^r \lambda_i v_i = 0 \Rightarrow \lambda_i = 0 \quad \forall i \quad (18.8)$$

(Die leere Menge gilt auch als linear unabhängig, der Nullvektor aber nicht.)

· Eine *Basis* B von V ist ein linear unabhängiges Erzeugendensystem. Eine Teilmenge $B \subset V$ ist genau dann eine Basis, wenn sich jedes $v \in V$ als

$$v = \sum_{i=1}^r v^i b_i \quad \text{mit } v^i \in \mathbb{K} \text{ und } b_i \in B \quad (18.9)$$

darstellen lässt *und* diese Darstellung bis auf Permutation und Addieren von 0 eindeutig ist.

· Man beachte, dass in (18.5) und in der Definition einer Basis nur endliche Linearkombinationen zugelassen sind. Eine Möglichkeit, unendlichen Linearkombinationen Sinn zu geben, ist die Einführung einer Norm auf V , vgl. Def. 4.6, unter der man dann beispielsweise wie in § 5 Reihen von Vektoren in V auf Konvergenz untersuchen kann. Dazu werden wir noch Gelegenheit erhalten.

Das wichtigste Ergebnis zur Struktur von Vektorräumen ist

Theorem 18.3. (i) *Jeder Vektorraum hat eine Basis.*

(ii) *Hat V ein endliches Erzeugendensystem (man sagt auch: V ist endlich erzeugt), so hat V eine endliche Basis und jede Basis von V hat die gleiche Kardinalität, genannt die Dimension von V , geschrieben $\dim(V)$.*

Ist V nicht endlich erzeugt, so heisst V unendlich dimensional. Verabredungsgemäss ist $\dim\{0\} = 0$. Für $\dim(V) = n$ schreiben wir auch (b_1, \dots, b_n) statt $\{b_1, \dots, b_n\}$, wenn wir uns die Basis geordnet denken, und füllen die Komponenten in (18.9) mit Nullen auf.

Bemerkungen/Beispiele 18.4. · Anfänglich scheinen die meisten Vektorräume “mit einer Basis geboren” zu werden. Beispielsweise ist

$$B = \left(e_i = (0, \dots, \underset{\substack{\uparrow \\ i\text{-te Stelle}}}{1}, \dots, 0)^T \right)_{i=1, \dots, n} \quad (18.10)$$

in “kanonischer” Weise eine Basis von \mathbb{K}^n .

· Allgemeiner ist für eine beliebige nicht-leere Menge M der *freie Vektorraum* über M die Menge der endlichen “formalen” Linearkombinationen

$$\text{Free}(M) = \{ \lambda_1 m_1 + \lambda_2 m_2 + \dots + \lambda_r m_r \mid \lambda_i \in \mathbb{K}, m_i \in M \} \quad (18.11)$$

mit Vektorraumoperationen, welche, im scheinbaren Unterschied zu allem Vorhergegangenen, erst “nachträglich” erklärt werden, ein Vektorraum mit Basis M . Tatsächlich ist $\text{Free}(M)$ nichts anderes als der oben eingeführte Unterraum $\mathcal{F}^{\text{fin}}(M, \mathbb{K})$ von $\mathcal{F}(M, \mathbb{K})$.

· In der Linearen Algebra lernt man jedoch, zwischen verschiedenen Basen zu wechseln zu schätzen, und auch für das Verständnis der Quantenmechanik ist es unerlässlich zu akzeptieren, dass Basen von Vektorräumen zusätzliche physikalische Information darstellen und keine a priori bevorzugt ist.

· Für ein lehrreiches Beispiel sei V ein \mathbb{K} -Vektorraum und $W \subset V$ ein Untervektorraum. Dann wird auf V durch

$$v_1 \sim v_2 :\Leftrightarrow v_1 - v_2 \in W \quad (18.12)$$

eine Äquivalenzrelation erklärt, und auf dem mengentheoretischen Quotienten

$$V/\sim = \{ [v] := \{ \tilde{v} \sim v \} \mid v \in V \} \quad (18.13)$$

erfüllen die Verknüpfungen

$$[v_1] + [v_2] := [v_1 + v_2], \quad \lambda[v] := [\lambda v] \quad (18.14)$$

die Vektorraumaxiome. Der resultierende Vektorraum V/W heisst *Quotientenraum* von V nach W .

Ist dann etwa V endlich-dimensional mit Basis $B = (b_1, \dots, b_n)$, so folgt zwar leicht, dass $\text{span}([b_1], \dots, [b_n]) = V/W$, d.h. V/W ist endlich erzeugt. Falls aber $\dim U \geq 1$, so ist $\{[b_1], \dots, [b_n]\} \subset V/U$ nicht mehr linear unabhängig und bildet daher keine Basis mehr. Es gibt im Allgemeinen keine weitere Information zur Auswahl einer Basis von V/W .

· Für viele praktische und auch theoretische Zwecke, insbesondere die, die mit endlich-dimensionalen Vektorräumen zu tun haben, ist die Wahl und das Rechnen in einer Basis sehr nützlich. Zur Erinnerung: Eine Abbildung $\hat{A} : V \rightarrow W$ heisst linear, falls für alle $\lambda \in \mathbb{K}$ und $v_1, v_2 \in V$, gilt, dass $\hat{A}(\lambda v_1 + v_2) = \lambda \hat{A}(v_1) + \hat{A}(v_2)$. Sind dann V und W endlich-dimensional mit Basen $B = (b_1, \dots, b_n)$ und $C = (c_1, \dots, c_m)$, dann existieren für $i = 1, \dots, n$ und $j = 1, \dots, m$ $A_i^j \in \mathbb{K}$ so, dass

$$\hat{A}(b_i) = \sum_{j=1}^m c_j A_i^j \quad \forall i = 1, \dots, n \quad (18.15)$$

und diese “Darstellung von \hat{A} bezüglich der ausgezeichneten Basen” legt die lineare Abbildung bereits vollständig fest. Man sagt auch: Der von den geordneten Basen gestiftete Isomorphismus $V \cong \mathbb{K}^n$, $W \cong \mathbb{K}^m$ induziert eine Identifikation

$$\text{Hom}_{\mathbb{K}}(V, W) \cong \text{Mat}_{m \times n}(\mathbb{K}) \quad (18.16)$$

§ 18. MULTI-LINEARE ALGEBRA

als \mathbb{K} -Vektorräume der Dimension $m \cdot n$.

· Die Komponenten w^j von $w := \hat{A}(v)$ bezüglich der Basis (c_j) erhält man aus den Komponenten v^i von $v \in V$ bezüglich der Basis (b_i) durch “Multiplikation der Matrix A auf den Spaltenvektor der Komponenten von v ”

$$w^j = \sum_{i=1}^n A_i^j v^i \quad (18.17)$$

· Ist $\tilde{B} = (\tilde{b}_1, \dots, \tilde{b}_n)$ eine andere Basis von V , so existiert eine invertierbare Matrix $R = (R_k^i)_{\substack{i=1, \dots, n \\ k=1, \dots, n}}$ so, dass

$$\tilde{b}_k = \sum_{i=1}^n b_i R_k^i \quad (18.18)$$

Die Komponenten bezüglich der neuen Basis hängen über

$$\tilde{v}^k = \sum_{i=1}^n (R^{-1})_i^k v^i \quad (18.19)$$

mit denen der alten zusammen: Aus

$$\sum_k R_k^i (R^{-1})_j^k = \delta_j^i = \begin{cases} 1 & i = j \\ 0 & \text{sonst} \end{cases} \quad (18.20)$$

folgt

$$\sum_{k=1}^n \tilde{b}_k \tilde{v}^k = \sum_{i=1}^n b_i v^i = v \quad (18.21)$$

· Ist auch noch $(\tilde{c}_1, \dots, \tilde{c}_m)$ eine andere Basis von W , mit $\tilde{c}_l = \sum_{j=1}^m c_j S_l^j$, und ist $A \in \text{Mat}_{m \times n}(\mathbb{K})$ die Darstellung einer linearen Abbildung $\hat{A} \in \text{Hom}_{\mathbb{K}}(V, W)$, so hat diese gleiche Abbildung bezüglich der neuen Basen die darstellende Matrix

$$\tilde{A} = S^{-1} \cdot A \cdot R \quad (18.22)$$

Operationen mit Vektorräumen

Wie bereits angedeutet zahlt sich der “operationelle” Standpunkt, von dem aus Vektorräume mit Basis als Räume von Funktionen aufgefasst werden können, auf lange Sicht auch gerade in der Physik aus,¹ solange man sicherstellt, dass die Betrachtungen entweder gar nicht, oder aber in “natürlicher Weise”, von der Wahl der Basis abhängen. Zur Illustration diskutieren wir die Zusammensetzung neuer Vektorräume aus gegebenen.

¹Im allereinfachsten Fall ist ein quantenmechanischer Zustand die Zuordnung einer komplexen Zahl zu möglichem Realtext, z.B. bedeutet $|\text{kitten}\rangle = \frac{1}{\sqrt{2}} |\text{alive}\rangle + \frac{i}{\sqrt{2}} |\text{dead}\rangle$: “totes Kätzchen” hat Amplitude $\frac{1}{\sqrt{2}}$, “lebendiges Kätzchen” Amplitude $\frac{i}{\sqrt{2}}$. Die Gesamtheit der Zustände bildet einen Vektorraum.

Definition 18.5 (Dualraum). Für einen \mathbb{K} -Vektorraum heisst der \mathbb{K} -Vektorraum der linearen Abbildung von V nach \mathbb{K} (als Vektorraum) der *Dualraum* von V .

$$V^* := \text{Hom}_{\mathbb{K}}(V, \mathbb{K}) \quad (18.23)$$

Bemerkungen. · Ist V endlich-dimensional, so ist V^* ebenfalls wieder endlich-dimensional mit $\dim(V^*) = \dim(V)$. Ist $B = (b_1, \dots, b_n)$ eine Basis von V , so ist $B^* = (\beta^1, \dots, \beta^n)$ mit den linearen Abbildungen

$$\beta^j : V \rightarrow \mathbb{K} \quad \beta^j(b_i) := \delta_i^j \quad (18.24)$$

eine Basis von V^* , welche die zu B duale Basis genannt wird. Elemente von V^* heissen auch Kovektoren.

· Im unendlich-dimensionalen Fall gilt die obige Aussage im Allgemeinen nicht. So hat z.B. der Vektorraum $\mathcal{F}^{\text{fin}}(\mathbb{N}, \mathbb{R})$ der endlichen Folgen die abzählbare Basis $(e_i)_{i=1,2,\dots}$ mit

$$e_i = (0, \dots, 0, 1, 0, \dots, 0, \dots) \quad (18.25)$$

\uparrow
i-te Stelle

Allerdings sind die zu den e_i dualen Kovektoren $(\epsilon^j)_{j=1,2,\dots}$ mit

$$\epsilon^j(e_i) = \delta_i^j \quad (18.26)$$

wie vorher zwar linear unabhängig, aber kein Erzeugendensystem für den Dualraum: Die lineare Abbildung, welche auf der Basis durch $e_i \mapsto 1 \ \forall i$ erklärt ist, ist keine endliche Linearkombination der ϵ^j .

· Aber auch einen endlich-dimensionalen Vektorraum muss man sorgfältig von seinem Dualraum unterscheiden. Mittels einer Basis B hat man zwar $\mathbb{K}^n \stackrel{B}{\cong} V$ und $V^* \stackrel{B^*}{\cong} \mathbb{K}^n$, so dass die scheinbar harmlose Identifikation $\mathbb{K}^n = \mathbb{K}^n$ einen Isomorphismus $V \cong V^*$ induziert. Dieser Isomorphismus hängt aber von der Wahl der Basis ab: Ist z.B. $\tilde{b}_k = \sum b_i R_k^i$ eine zweite Basis von V , und $(\tilde{\beta}^l)$ die zu (\tilde{b}_k) duale Basis, so gilt (vgl. (18.19))

$$\tilde{\beta}^l = \sum_{j=1}^n (R^{-1})_j^l \beta^j \quad (18.27)$$

Zusammen mit (18.22) folgt daraus: Der von B, B^* induzierte Isomorphismus $V \cong V^*$, der bezüglich der Basen B, B^* durch die Einheitsmatrix dargestellt wird, hat bezüglich der Basen \tilde{B}, \tilde{B}^* die darstellende Matrix $R^T R$. Im Allgemeinen unterscheidet er sich damit von dem Isomorphismus, welcher bezüglich der Basen \tilde{B}, \tilde{B}^* natürlich erscheint (weil er bezüglich dieser Basen durch die Einheitsmatrix dargestellt wird). Man sagt zu (18.27) auch salopp: ‘‘Kovektoren transformieren sich unter Basiswechsel kontragredient zu den Vektoren’’.

Definition 18.6 (Direkte Summe). Es seien V und W zwei \mathbb{K} -Vektorräume. Dann erfüllen auf dem kartesischen Produkt $V \times W$, das ist die Menge aller Paare (v, w) mit $v \in V$ und $w \in W$, die Verknüpfungen

$$(v_1, w_1) + (v_2, w_2) := (v_1 + v_2, w_1 + w_2), \quad \lambda \cdot (v, w) := (\lambda \cdot v, \lambda \cdot w) \quad (18.28)$$

§ 18. MULTI-LINEARE ALGEBRA

(wobei die Verknüpfungen in den Klammern die von V bzw. W sind) die Vektorraumaxiome. Der resultierende Vektorraum

$$V \oplus W = (V \times W, +, \cdot, (0, 0)) \quad (18.29)$$

heisst die *direkte Summe* von V und W .

Bemerkungen. · Für (v, w) schreibt man auch $v \oplus w$, und mit der Identifikation $V \cong \{(v, 0) \mid v \in V\} \subset V \oplus W$ auch einfach nur $v + w$.

· Ist B eine Basis von V , und C eine Basis von W so ist (mit der eben gegebenen Identifikation) $B \cup C$ eine Basis von $V \oplus W$. Anders gesagt gilt

$$\mathcal{F}^{\text{fin}}(B \cup C, \mathbb{K}) = \mathcal{F}^{\text{fin}}(B, \mathbb{K}) \oplus \mathcal{F}^{\text{fin}}(C, \mathbb{K}) \quad (18.30)$$

und im endlich-dimensionalen Fall insbesondere

$$\dim(V \oplus W) = \dim(V) + \dim(W) \quad (18.31)$$

· Ist X ein \mathbb{K} -Vektorraum mit Unterräumen V und W , so heisst $V + W := \text{span}(V \cup W)$ die “innere” Summe von V und W . Dies ist der kleinste Unterraum von X , der V und W enthält, deckt allerdings auch den Fall ab, dass V und W nicht-trivialen Durchschnitt haben. Gilt ausserdem $V \cap W = \{0\}$, so heisst die innere Summe “direkt”, und man schreibt auch $V \oplus W$. Zur Unterscheidung heisst die Konstruktion in 18.6 dann “äussere” direkte Summe. Allerdings ist der Unterschied letztlich klein, da in der Situation von 18.6 V und W in “natürlicher” Weise Unterräume von $V \oplus W$ sind.

Die geometrische Vorstellung von $\mathbb{K}^n \oplus \mathbb{K}^m$ als dem kartesischen Produkt entspricht der klassischen Intuition dafür, wie man das Zusammensetzen unterschiedener physikalischer Größen und Systeme mathematisch abbilden sollte. Tatsächlich stellt man die Bewegung eines Teilchens durch Raum und Zeit als Bahn in $\mathbb{R}^3 \times \mathbb{R}$ dar, und die Konfiguration von N Teilchen beschreibt man durch Vektoren im $\mathbb{R}^{3N} = \mathbb{R}^3 \oplus \dots \oplus \mathbb{R}^3$. Denkt man hingegen über die Zustände physikalischer Systeme als “Funktionen auf dem Raum ihrer möglichen Eigenschaften”, so ist für die Zusammensetzung von Systemen eine andere Operation natürlicher, und wie die Quantenmechanik zeigt, die richtige: das Tensorprodukt.

Definition 18.7 (Multi-lineare Abbildungen). Seien V_1, \dots, V_k und X Vektorräume über \mathbb{K} . Eine Abbildung

$$T : V_1 \times \dots \times V_k \rightarrow X \quad (18.32)$$

heisst k -linear (für $k = 2$ auch bilinear), falls für alle $i = 1, \dots, k$ und alle

$$(v_1, \dots, v_{i-1}, v_{i+1}, \dots) \in V_1 \times \dots \times V_{i-1} \times V_{i+1} \times \dots \times V_k$$

die Abbildung

$$V_i \ni v \mapsto T(v_1, \dots, v_{i-1}, v, v_{i+1}, \dots, v_k) \in X \quad (18.33)$$

linear (in v) ist. Wir schreiben $L(V_1, \dots, V_k; X)$ für den Vektorraum dieser k -linearen Abbildungen.

Semidefinition 18.8 (Tensorprodukt). Seien V und W Vektorräume über \mathbb{K} . Für gegebene Basen B von V und C von W gilt

$$V = \mathcal{F}^{\text{fin}}(B, \mathbb{K}) \quad W = \mathcal{F}^{\text{fin}}(C, \mathbb{K}) \quad (18.34)$$

und wir definieren das *Tensorprodukt* von V und W als den Vektorraum

$$V \otimes W = \mathcal{F}^{\text{fin}}(B \times C, \mathbb{K}) \quad (18.35)$$

Bemerkungen. · Vor allem Anderen vergleiche man (18.35) mit (18.30).

· Dann aber beachte man, dass (18.8) eigentlich das Tensorprodukt von Vektorräumen mit Basis definiert, und zwar wieder als Vektorraum mit Basis. Um die Unabhängigkeit von der Wahl der Basis sicherzustellen, ergänzt man die Konstruktion

1. um eine *bilineare Abbildung*

$$V \times W \rightarrow V \otimes W \quad (v, w) \mapsto v \otimes w \quad (18.36)$$

welche in den gegebenen Identifikationen (18.34), (18.35) definiert ist durch

$$(f \otimes g)(b, c) := f(b) \cdot g(c) \quad (18.37)$$

(Übungsaufgabe: Man prüfe, dass diese Abbildung bilinear ist.), und mit deren Hilfe man die Basis von $V \otimes W$ auch als

$$\{b \otimes c \mid b \in B, c \in C\} \quad (18.38)$$

schreiben kann, und

2. um eine *Vorschrift*, wie ein Basiswechsel in V und W einen Basiswechsel im Tensorprodukt induziert. Seien etwa im endlich-dimensionalen Fall

$$\tilde{b}_j = \sum_{i=1}^n b_i R_j^i \quad \tilde{c}_l = \sum_{k=1}^l c_k S_l^k \quad (18.39)$$

andere Basen von V und W , und $V \tilde{\otimes} W$ das entsprechende (a priori verschiedene) Tensorprodukt. Dann gilt

$$\tilde{b}_j \otimes \tilde{c}_l = \sum_{i,k} b_i \otimes c_k R_j^i S_l^k \quad (18.40)$$

Identifiziert man dann $\tilde{b}_j \tilde{\otimes} \tilde{c}_l \in V \tilde{\otimes} W$ mit $\tilde{b}_j \otimes \tilde{c}_l \in V \otimes W$, so ist die darstellende Matrix der Abbildung (18.36) von der Wahl der Basis unabhängig.

· Das Tensorprodukt zweier Vektorräume als der Vektorraum (18.35) zusammen mit der bilinearen Abbildung (18.36) ist damit so kanonisch definiert, wie es mit Hilfe von Basen geht. Bei weiterbestehenden oder wiederkehrenden Zweifeln merke man sich (18.38), im endlich-dimensionalen Fall noch (18.40) und vor allem

$$\dim(V \otimes W) = \dim(V) \cdot \dim(W) \quad (18.41)$$

§ 18. MULTI-LINEARE ALGEBRA

im Unterschied zu (18.31).

· Es sei aber gesagt, dass man das Tensorprodukt auch basisunabhängig definieren kann als

$$\mathcal{F}^{\text{fin}}(V \times W, \mathbb{K})/L \quad (18.42)$$

für einen geeigneten Unterraum L und einer dann ganz kanonischen bilinearen Abbildung $V \times W \rightarrow V \otimes W$, siehe Lehrbücher zur Linearen Algebra unter “universeller Eigenschaft des Tensorprodukts”.

· Aus dieser universellen Eigenschaft des Tensorprodukts folgt auch, dass für gegebene $A_1 \in \text{Hom}_{\mathbb{K}}(V_1, W_1)$ und $A_2 \in \text{Hom}_{\mathbb{K}}(V_2, W_2)$ das Produkt

$$A_1 \times A_2 : V_1 \times V_2 \rightarrow W_1 \times W_2 \rightarrow W_1 \otimes W_2 \quad (18.43)$$

eine eindeutige lineare Abbildung

$$A_1 \otimes A_2 : V_1 \otimes V_2 \rightarrow W_1 \otimes W_2 \quad (18.44)$$

induziert, welche $(A_1 \otimes A_2)(v_1 \otimes v_2) = A_1(v_1) \otimes A_2(v_2)$ erfüllt. (Konkret ist das Tensorprodukt zweier linearen Abbildungen einfach die lineare Fortsetzung der Abbildung

$$(A_1 \otimes A_2)(b_1 \otimes b_2) := A_1(b_1) \otimes A_2(b_2) \quad (18.45)$$

von einer Produkt-Basis² $\{b_1 \otimes b_2 \mid b_1 \in B_1, b_2 \in B_2\}$ auf ganz $V_1 \otimes V_2$.) (18.44) ist analog zur wohl schon bekannten direkten Summe von linearen Abbildungen,

$$\begin{aligned} A_1 \oplus A_2 : V_1 \oplus V_2 &\rightarrow W_1 \oplus W_2 \\ (A_1 \oplus A_2)(v_1 \oplus v_2) &:= A_1(v_1) \oplus A_2(v_2) \end{aligned} \quad (18.46)$$

· Die Bilinearität des Tensorprodukts auf Vektoren (dass nämlich $(\lambda v_1 + v_2) \otimes w = \lambda v_1 \otimes w + v_2 \otimes w \in V \otimes W$) überträgt sich auf die Bilinearität des Tensorprodukts von Abbildungen: Für $A_1, B_1 \in \text{Hom}_{\mathbb{K}}(V_1, W_1)$ und $A_2 \in \text{Hom}_{\mathbb{K}}(V_2, W_2)$, $\lambda \in \mathbb{K}$ gilt

$$(\lambda A_1 + B_1) \otimes A_2 = \lambda A_1 \otimes A_2 + B_1 \otimes A_2 \in \text{Hom}_{\mathbb{K}}(V_1 \otimes V_2, W_1 \otimes W_2) \quad (18.47)$$

· Ausserdem erfreuen sich direkte Summe und Tensorprodukt von Vektorräumen selbst der Verträglichkeitseigenschaften, die ihren Namen und Notation rechtfertigen: Für je drei \mathbb{K} -Vektorräume V_1, V_2, V_3 gilt etwa

$$\begin{aligned} (V_1 \oplus V_2) \otimes V_3 &\cong (V_1 \otimes V_3) \oplus (V_2 \otimes V_3) \\ (V_1 \otimes V_2) \otimes V_3 &\cong V_1 \otimes (V_2 \otimes V_3) \end{aligned} \quad (18.48)$$

wobei man \cong statt $=$ schreibt, da diese Identifikationen nur “kanonisch” aber nicht “unmittelbar” sind. Die Unterscheidung scheint spitzfindig, ist aber relativ klar bei der Kommutativität: Natürlich ist $V \otimes W$ nicht *gleich* $W \otimes V$ in dem Sinne, dass ja das erste Tensorprodukt von Vektoren der Form $v \otimes w$ für $v \in V$ und $w \in W$ aufgespannt wird, während das zweite von solchen der Form $w \otimes v$. Die Abbildung

$$c_{V,W} : V \otimes W \ni v \otimes w \mapsto w \otimes v \in W \otimes V \quad (18.49)$$

²Dass nicht jede Basis von $V_1 \otimes V_2$ von dieser Form ist, lernen wir hoffentlich in den Übungen.

definiert aber einen linearen Isomorphismus, welcher kanonisch ist zumindestens in dem Sinne, dass er nicht von weiteren Wahlen abhängt. Falls allerdings von vorneherein $V = W$, so gibt es einen “noch kanonischeren” Isomorphismus, nämlich die Identität. Dadurch wird $c_{V,V}$ zu einer nicht-trivialen *Selbstabbildung* von $V \otimes V$, $v_1 \otimes v_2 \mapsto v_2 \otimes v_1$. Dies spielt unten und in der Physik insbesondere bei der Zusammensetzung “ununterscheidbarer Systeme” eine wichtige Rolle.

· In ähnlichem Sinne gelten

$$\mathbb{K} \otimes V \cong V, \quad \mathbb{K}^0 \otimes V \cong \mathbb{K}^0, \quad \mathbb{K}^0 \oplus V \cong V \quad (18.50)$$

Proposition 18.9. *Seien V und W endlich-dimensionale \mathbb{K} -Vektorräume. Dann ist die (kanonische) Abbildung*

$$\text{ev} : V^* \otimes W \rightarrow \text{Hom}_{\mathbb{K}}(V, W), \quad (18.51)$$

definiert als lineare Fortsetzung von

$$\lambda \otimes w \mapsto (v \mapsto \lambda(v) \cdot w) \quad (18.52)$$

ein Isomorphismus von Vektorräumen.

Beweis. Man prüft leicht, dass ev wohldefiniert ist. (Dazu genügt es zu zeigen, dass (18.52) linear in $\lambda \in V^*$ und $w \in W$ ist.) Um zu zeigen, dass (18.51) ein Isomorphismus ist, genügt es dann, Surjektivität nachzuweisen, da linke und rechte Seite \mathbb{K} -Vektorräume der Dimension $n \cdot m = \dim(V) \cdot \dim(W)$ sind (vgl. (18.16), (18.41)). Sei dazu (b_1, \dots, b_n) eine Basis von V , $(\beta^1, \dots, \beta^n)$ die dazu duale Basis von V^* , und (c_1, \dots, c_m) eine Basis von W . Für $A \in \text{Hom}_{\mathbb{K}}(V, W)$ sei (A_i^k) die darstellende Matrix bezüglich B und C . Dann gilt

$$\text{ev}\left(\sum_{j=1}^n \sum_{k=1}^m \beta^j \otimes c_k A_j^k\right)(b_i) = \sum_{k=1}^m c_k A_i^k = A(b_i) \quad \forall i = 1, \dots, n \quad (18.53)$$

und daraus folgt $A \in \text{ev}(V^* \otimes W)$. □

Bemerkungen. · Dieser Satz gilt im Allgemeinen nicht für unendlich-dimensionale Vektorräume. (ev ist immer injektiv, aber nicht notwendigerweise surjektiv. Die Abbildung ist surjektiv, wenn wenigstens entweder V oder W endlich-dimensional ist.)

· Für endlich-dimensionale Vektorräume gilt auch

$$(V^*)^* = V \quad (18.54)$$

sowie

$$(V \otimes W)^* = V^* \otimes W^* \quad (18.55)$$

wobei die linke Seite definitionsgemäss auch gleich dem Raum der bilinearen Abbildungen $V \times W \rightarrow \mathbb{K}$ ist (vgl. 18.7).

· Dies verallgemeinert sich auf mehrere (endlich-dimensionale) Faktoren zu

$$L(V_1, \dots, V_k; \mathbb{K}) = \text{Hom}_{\mathbb{K}}(V_1 \otimes \dots \otimes V_k, \mathbb{K}) = V_1^* \otimes \dots \otimes V_k^* \quad (18.56)$$

sodass letztlich das Tensorprodukt die Theorie der multi-linearen Abbildungen auf lineare Algebra zurückgeführt hat.

§ 19. GRUPPEN UND IHRE DARSTELLUNGEN

Definition 18.10 (Komplexifizierung). Gemäss unseren Betrachtungen in § 2 ist der Körper der komplexen Zahlen $\mathbb{C} \cong \mathbb{R}^2$ ein zwei-dimensionaler reeller Vektorraum, der in einer bestimmten Basis mit einer verträglichen Multiplikation ausgerüstet wird. Ist nun V ein Vektorraum über \mathbb{R} , so erfüllt auf dem Tensorprodukt

$$V^{\mathbb{C}} = V \otimes \mathbb{C} \quad (18.57)$$

die Multiplikation mit komplexen Skalaren, welche als \mathbb{R} -lineare Fortsetzung von

$$\lambda(v \otimes \mu) := v \otimes (\lambda\mu) \quad v \in V, \lambda, \mu \in \mathbb{C} \quad (18.58)$$

definiert ist (m.A.W. ist dies $\text{id}_V \otimes \lambda$), zusammen mit der auf (18.57) bereits existierenden Addition die Axiome eines komplexen Vektorraums. Man nennt diesen Vektorraum die Komplexifizierung von V . Mit einer Basis B von V , d.h. $V \cong \mathcal{F}^{\text{fin}}(B, \mathbb{R})$, lässt sich die Komplexifizierung auch als $V^{\mathbb{C}} := \mathcal{F}^{\text{fin}}(B, \mathbb{C})$ einführen, allerdings ist wie beim Tensorprodukt der Knackpunkt, dass nicht jede Basis von $V^{\mathbb{C}}$ als komplexem Vektorraum auf diese Weise “von einer Basis von V kommt”.

Definition 18.11 (Duale Abbildung). Seien V und W Vektorräume über \mathbb{K} , und $A \in \text{Hom}_{\mathbb{K}}(V, W)$. Dann heisst die lineare Abbildung

$$\begin{aligned} A^* : W^* &\rightarrow V^* \\ A^*(\mu)(v) &:= \mu(A(v)) \end{aligned} \quad (18.59)$$

die zu A duale Abbildung. In endlichen Dimensionen wird bezüglich dualer Basen A^* durch die zu A transponierte Matrix dargestellt.

§ 19 Gruppen und ihre Darstellungen

Neben Vektorräumen ist die zweite für die theoretische Physik zentrale algebraische Struktur die von *Gruppen*. Während Vektorräume die Kenntnis von physikalischen Sachverhalten abbilden, fassen Gruppen bestimmte Wechsel der Beschreibung, wie etwa den Übergang zwischen verschiedenen Inertialsystemen in der Mechanik oder Eichtransformationen in der Elektrodynamik. Naturgemäss sind dabei diejenigen Eigenschaften, die *nicht* von der gewählten Beschreibung abhängen, von eigentlichem Interesse. Wir wiederholen die Definitionen.

Definition 19.1. (i) Eine Gruppe ist eine Menge G mit einem ausgezeichneten Element $e \in G$ und einer binären Verknüpfung

$$\cdot : G \times G \rightarrow G, \quad (g_1, g_2) \mapsto g_1 \cdot g_2 \quad (19.1)$$

mit den Rechenregeln

$$\begin{aligned} \forall g_1, g_2, g_3 \in G : (g_1 \cdot g_2) \cdot g_3 &= g_1 \cdot (g_2 \cdot g_3) \\ \forall g \in G : e \cdot g &= g \end{aligned} \quad (19.2)$$

derart, dass für all $g \in G$ die Gleichung

$$h \cdot g = e \quad (19.3)$$

eine Lösung $h \in G$ besitzt. (Die Lösung ist automatisch eindeutig und erfüllt als g^{-1} notiert auch $g \cdot g^{-1} = e$. Es gilt auch $g \cdot e = g \forall g$ und $e' \cdot g = g \forall g \Rightarrow e' = e$.)

(ii) Gilt $g_1 \cdot g_2 = g_2 \cdot g_1 \forall g_1, g_2 \in G$, so heisst G *abelsch* oder *kommutativ*. Eine abelsche Gruppe wird manchmal auch additiv notiert.

(iii) Für Gruppen G und H heisst eine Abbildung $\phi : G \rightarrow H$ ein *Gruppenhomomorphismus* falls

$$\phi(g_1 \cdot g_2) = \phi(g_1) \cdot \phi(g_2) \quad \forall g_1, g_2 \in G \tag{19.4}$$

Ein bijektiver Gruppenhomomorphismus heisst auch *Isomorphismus von Gruppen*.

(iv) Eine *Untergruppe* einer Gruppe G ist eine nicht-leere Teilmenge $H \subset G$ welche unter Multiplikation in der Gruppe und Inversenbildung abgeschlossen ist. Das ist genau dann der Fall, falls für alle $g_1, g_2 \in H$ auch $g_1 \cdot g_2^{-1} \in H$.

Beispiel 19.2. · Für eine Menge M ist die Menge der bijektiven Abbildungen $\text{Bij}(M) = \{f : M \rightarrow M \mid f \text{ bijektiv}\}$ mit der Identität (bzw. für $M = \emptyset$, der leeren Abbildung) als neutralem Element und der Verkettung von Abbildungen als Multiplikation eine Gruppe.

· Für eine endliche Menge $N \cong \{1, 2, \dots, n\}$ mit n Elementen heisst $\text{Bij}(N) = S_n$ die Permutationsgruppe oder *symmetrische Gruppe* vom Grad n . Für das Rechnen in S_n bei festem n kann man Elemente von S_n ,

$$S_n \ni \sigma : N \ni a \mapsto \sigma(a) \tag{19.5}$$

als zwei-zeilige Matrix schreiben:

$$\sigma = \begin{pmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{pmatrix} \tag{19.6}$$

In S_5 gilt etwa

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 3 & 4 & 5 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 5 & 1 & 3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 5 & 2 & 3 & 1 \end{pmatrix} \tag{19.7}$$

(Wenn man in Übung ist, so kann man auch die erste Zeile in der Notation fortlassen.)

· Für einen \mathbb{K} -Vektorraum V ist die Menge der invertierbaren linearen Abbildungen $V \rightarrow V$ in natürlicher Weise eine Gruppe, genannt die allgemein lineare Gruppe, $GL(V)$. Für einen endlich-dimensionalen Vektorraum mit $\dim(V) = n$ ist (nach Wahl einer Basis) $GL(V) \cong GL(n, \mathbb{K})$ isomorph zur Gruppe der invertierbaren $n \times n$ -Matrizen mit Einträgen in \mathbb{K} .

· Die Menge \mathbb{Z} der ganzen Zahlen mit Addition als Verknüpfung und 0 als neutralem Element ist eine abelsche Gruppe.

· Für jeden Körper \mathbb{K} ist die Menge der invertierbaren Elemente, d.h. also $\mathbb{K}^\times := \mathbb{K} \setminus \{0\}$ mit der Multiplikation als Verknüpfung und der 1 als neutralem Element eine ebenfalls abelsche Gruppe.

Bemerkungen/Beispiele 19.3. · S_n ist eine Art universelles Beispiel einer *endlichen* Gruppe (eine Gruppe heisst endlich, wenn sie endlich viele Elemente enthält). Die Anzahl Permutationen der Zahlen von 1 bis n ist $|S_n| = n!$, und jede endliche

§ 19. GRUPPEN UND IHRE DARSTELLUNGEN

Gruppe ist isomorph zu einer Untergruppe von S_n für $n = |G|$: Für $g \in G$ ist die Linksmultiplikation mit g ,

$$L_g : G \ni h \mapsto g \cdot h \in G, \quad (19.8)$$

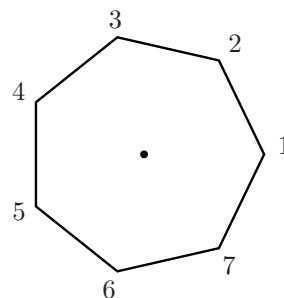
eine Permutation der G zugrundeliegenden Menge, mit Inversem $L_{g^{-1}}$. $g \mapsto L_g$ ist ein injektiver Gruppenhomomorphismus, und damit ist G eine Untergruppe von $\text{Bij}(G)$. (Diese Aussage impliziert aber sicher nicht, dass man das Studium der endlichen Gruppen auf S_n reduzieren kann.)

· Für $n \in \mathbb{Z}$ ist die Menge der Restklassen modulo n mit Addition als Verknüpfung und $[0] = \{k \cdot n \mid k \in \mathbb{Z}\}$ als neutralem Element eine endliche abelsche Gruppe, genannt die *zyklische Gruppe* $\mathbb{Z}/n\mathbb{Z}$, das sie als Untergruppe $C_n \subset S_n$ die Zahlen $\{1, \dots, n\}$ “im Kreis dreht”. Wir notieren sie dann multiplikativ

$$C_n = \{r_k \mid k \in \mathbb{Z}\} / (r_{n+k} \sim r_k) \text{ mit } r_k \cdot r_l = r_{k+l}. \quad (19.9)$$

· $GL(n, \mathbb{R})$ ist ein Beispiel einer *kontinuierlichen Gruppe*, d.h. einer Gruppe, auf deren zugrundeliegender Menge ausserdem eine metrische Struktur (im Sinne von 4.1) erklärt ist, mit der Eigenschaft, dass die Operationen in der Gruppe (Multiplikation und Inversenbildung) stetige Abbildungen sind. $GL(n, \mathbb{R})$ ist (für $n > 0$) nicht endlich, als offene Teilmenge von \mathbb{R}^{n^2} aber noch *endlich-dimensional*, und universell in dem eingeschränkten Sinn, dass alle “vernünftigen” kontinuierlichen Gruppen lokal isomorph zu Untergruppen von $GL(n, \mathbb{R})$ sind. (Die genaue Aussage und der Beweis sind wesentlich schwieriger...)

· Spezielle Gruppen treten als Symmetriegruppen algebraischer Strukturen oder geometrischer Objekte auf. Als Beispiel betrachten wir die *Diedergruppe* D_n , definiert als die Symmetriegruppe eines regelmässigen n -Ecks in der euklidischen Ebene. Um D_n zu beschreiben, betrachten wir die Wirkung von $g \in D_n$ auf die gegen den Uhrzeigersinn von 1 bis n durchnummerierten Ecken. Ist $g(1)$ der Bildpunkt der Ecke 1, so sei $k := g(1) - 1 \in \{0, \dots, n-1\}$. Damit g das Polygon invariant lässt, muss dann $g(2)$ ein Nachbar von $g(1)$ sein, d.h. es gilt entweder $g(2) = g(1) + 1 = k + 2$ oder $g(2) = g(1) - 1 = k$ (mit einer offensichtlichen periodischen Identifikation der Indizes für $k = n - 1$ bzw. $k = 0$). Im ersten Fall folgt leicht, dass $g = r_k$ eine Drehung um den Mittelpunkt mit Winkel $2\pi k/n$ ist. Im zweiten Fall ist $g = s_k$ eine Spiegelung an der Winkelhalbierenden.



Die Fortsetzungen dieser Elemente von D_n zu linearen (tatsächlich orthogonalen) Abbildungen $\mathbb{R}^2 \rightarrow \mathbb{R}^2$ haben bezüglich der Standardbasis die darstellenden Matrizen³

$$r_k = \begin{pmatrix} \cos \frac{2\pi k}{n} & -\sin \frac{2\pi k}{n} \\ \sin \frac{2\pi k}{n} & \cos \frac{2\pi k}{n} \end{pmatrix}, \quad s_k = \begin{pmatrix} \cos \frac{2\pi k}{n} & \sin \frac{2\pi k}{n} \\ \sin \frac{2\pi k}{n} & -\cos \frac{2\pi k}{n} \end{pmatrix} \quad (19.10)$$

³Für die Spiegelungen prüfe man, dass $(\cos \frac{\pi k}{n}, \sin \frac{\pi k}{n})$ und $(-\sin \frac{\pi k}{n}, \cos \frac{\pi k}{n})$ Eigenvektoren von s_k zu den Eigenwerten $+1$ und -1 sind.

Die von den r_k erzeugte Untergruppe von D_n ist isomorph zu C_n , im Übrigen gelten die Relation

$$s_k = r_k \cdot s_0, \quad s_0 r_k s_0 = r_{-k} = (r_k)^{-1}, \quad s_k \cdot s_k = r_0 = e \quad (19.11)$$

Für $n \geq 3$ ist D_n nicht-abelsch. (Mehr in den Übungen.)

· Wie diese Beispiele zeigen, hat die Kategorie der Gruppen eine reichere Struktur als die der Vektorräume. Insbesondere liegt eine endliche Gruppe nicht bereits durch die Angabe der Anzahl ihrer Elemente fest, während für einen endlich-dimensionalen Vektorraum über \mathbb{K} die Angabe der Dimension genügt, um ihn (bis auf Isomorphismus) vollständig zu bestimmen. Ausserdem ererbt im Unterschied zu (18.13) der Quotient einer Gruppe nach einer Untergruppe im Allgemeinen keine natürliche Gruppenstruktur: Zwar ist für eine Untergruppe $H \subset G$

$$g \sim g' :\Leftrightarrow g^{-1} \cdot g' \in H \Leftrightarrow \exists h \in H : g' = g \cdot h \quad (19.12)$$

eine Äquivalenzrelation auf G . Der offensichtliche Versuch, auf der Menge der Äquivalenzklassen

$$G/H := \{gH := \{g \cdot h \mid h \in H\} \mid g \in G\} \quad (19.13)$$

eine Multiplikation zu definieren,

$$(g_1 H) \cdot (g_2 H) \stackrel{?}{=} (g_1 \cdot g_2) H \quad (19.14)$$

ist aber nur dann erfolgreich, wenn diese Multiplikation wohldefiniert, d.h. von den gewählten Repräsentanten unabhängig ist. Dies ist genau dann der Fall, wenn für alle $g'_1 H = g_1 H$, d.h. $g'_1 = g_1 \cdot h$ für ein $h \in H$, gilt, dass $(g'_1 \cdot g_2) H = (g_1 \cdot g_2) H$. Mit anderen Worten, es muss ein $h' \in H$ existieren, mit dem $g'_1 g_2 = g_1 g_2 h'$, oder

$$(g_1 g_2)^{-1} (g'_1 g_2) = g_2^{-1} g_1^{-1} g_1 h g_2 = g_2^{-1} \cdot h \cdot g_2 = h' \quad (19.15)$$

Eine Untergruppe mit dieser Eigenschaft, das heisst

$$g^{-1} \cdot H \cdot g \subset H \quad \forall g \in G \quad (19.16)$$

heisst *normale Untergruppe*, häufiger auch *Normalteiler*.

· Beispielsweise ist in D_n die zyklische Untergruppe C_n der Drehungen eine normale Untergruppe (vgl. insbes. die zweite Relation in (19.11)). Andererseits ist die von s_0 erzeugte Untergruppe $\{e, s_0\} \cong \mathbb{Z}/2\mathbb{Z}$ kein Normalteiler:

$$r_1^{-1} s_0 r_1 = r_1^{-1} r_1^{-1} s_0 = r_2^{-1} s_0 \neq s_0 \quad (19.17)$$

(für $n > 2$).

Zur Vorbereitung der nächsten Definition bemerken wir, dass man die erste Klasse von Beispielen in 19.2 in die zweite einbetten kann, indem man die Menge M im Sinne von (18.6) (bzw. (18.11)) zu einer Basis eines Vektorraums erklärt. Konkret ist dies für $S_n = \text{Bij}(\{1, 2, \dots, n\})$ der injektive Gruppenhomomorphismus

$$\begin{aligned} \phi : S_n &\rightarrow GL(n, \mathbb{R}) \\ \phi(\sigma)(e_i) &:= e_{\sigma(i)} \end{aligned} \quad (19.18)$$

§ 19. GRUPPEN UND IHRE DARSTELLUNGEN

Vorsicht: Für die Komponenten $(x^i)_{i=1,\dots,n}$ von $x \in \mathbb{R}^n$ bezüglich der Basis $(e_i)_{i=1,\dots,n}$ gilt $\phi(\sigma)(x)^i = x^{\sigma^{-1}(i)}$. Man beachte auch, dass diese Konstruktion verschieden ist von der linearen Fortsetzung (19.10).

Jede endliche Gruppe wird also durch Verkettung von (19.8) und (19.18) zu einer Untergruppe von $GL(n, \mathbb{R})$. Will man sich von der Wahl der Basis aber wieder unabhängig machen, so macht es Sinn, die folgende allgemeine Verbindung zwischen Gruppentheorie und linearer Algebra zu erklären.

Definition 19.4. Eine (*lineare*) *Darstellung* einer Gruppe G ist ein Vektorraum V zusammen mit einem Gruppenhomomorphismus

$$\rho : G \rightarrow GL(V) \quad (19.19)$$

· In dieser Definition wird nicht angenommen, dass G endlich ist oder ρ injektiv. Im Allgemeinen muss auch V nicht endlich-dimensional sein, und der zugrundeliegende Körper kann beliebig sein. Wir beschränken uns im Folgenden aber auf endlich-dimensionale Darstellungen über $\mathbb{K} = \mathbb{C}$ (oder allenfalls \mathbb{R}).

· Man interessiert sich dann bei festgehaltener Gruppe G für die verschiedenen Weisen, auf die “ G als Symmetriegruppe eines physikalischen Systems” realisiert werden kann. Dabei tauchen eine Reihe von assoziierten Begriffen auf.

· Eine G -lineare Abbildung (oder G -Homomorphismus) zwischen zwei Darstellungen (V_1, ρ_1) und (V_2, ρ_2) ist eine \mathbb{K} -lineare Abbildung $\phi : V_1 \rightarrow V_2$ von Vektorräumen derart, dass für alle $g \in G$

$$\phi \circ \rho_1(g) = \rho_2(g) \circ \phi \in \text{Hom}_{\mathbb{K}}(V_1, V_2) \quad (19.20)$$

Solche G -linearen Abbildungen bilden einen Untervektorraum aller \mathbb{K} -linearen Abbildungen, den wir mit $\text{Hom}_G(V_1, V_2)$ bezeichnen.

· (V_1, ρ_1) und (V_2, ρ_2) heißen isomorph (oder äquivalent) falls eine invertierbare G -lineare Abbildung $V_1 \xrightarrow{\cong} V_2$ existiert.

· Eine *Unterdarstellung* einer Darstellung (V, ρ) ist ein Untervektorraum $W \subset V$ mit der Eigenschaft, dass $\rho(g)(W) \subset W \forall g \in G$. Ein solches W heißt auch *invarianter Unterraum*. Durch Einschränken von $\rho(g)$ auf W ist dann $(W, \rho|_W)$ ebenfalls eine Darstellung von G .

· Man beachte, dass die Aussage “ W ist ein invarianter Unterraum” nicht bedeutet, dass notwendig $\rho(g)(w) = w \forall w \in W$. Man schreibt die Menge der invarianten Vektoren auch als

$$V^G := \{v \in V \mid \rho(g)(v) = v \forall g \in G\} \quad (19.21)$$

· Die direkte Summe von zwei Darstellungen ist der Vektorraum $V_1 \oplus V_2$ mit dem Homomorphismus

$$g \mapsto (\rho_1 \oplus \rho_2)(g) := \rho_1(g) \oplus \rho_2(g) \quad (19.22)$$

· Das Tensorprodukt von Darstellungen ist der Vektorraum $V_1 \otimes V_2$ mit dem Homomorphismus

$$g \mapsto (\rho_1 \otimes \rho_2)(g) := \rho_1(g) \otimes \rho_2(g) \quad (19.23)$$

· Zwei Trivialitäten:

* Für jeden Vektorraum V und beliebiges G wird durch $\rho(g) := \text{id}_V \ \forall g \in G$ eine Darstellung erklärt.

* Für jede Darstellung (V, ρ) von G sind der triviale Vektorraum $\{0\}$ und das gesamte V invariante Unterräume.

Definition 19.5. Eine Darstellung (V, ρ) von G mit $V \neq \{0\}$ heisst *irreduzibel* falls sie keine nicht-triviale Unterdarstellung enthält, d.h. es gibt genau zwei invariante Unterräume, nämlich $\{0\}$ und ganz V . Andernfalls heisst die Darstellung *reduzibel*.

· Eine Darstellung heisst *zerlegbar* falls sie äquivalent zu einer nicht-trivialen direkten Summe zweier Darstellungen ist, andernfalls unzerlegbar.

· Eine *Irrep* ist eine Isomorphismenklasse irreduzibler Darstellungen.

Beispiel: ϕ aus (19.18) ist eine n -dimensionale Darstellung von S_n , die sogenannte Permutationsdarstellung. Sie ist reduzibel (für $n > 1$), denn der von $\bar{e} := e_1 + \dots + e_n$ aufgespannte Unterraum ist invariant unter S_n .

Theorem 19.6 (Lemma von Schur). *Seien (V_1, ρ_1) und (V_2, ρ_2) Darstellungen einer Gruppe G über \mathbb{K} , und $\phi \in \text{Hom}_G(V_1, V_2)$ ein G -Homomorphismus.*

(i) $\text{Ker}(\phi) := \{v \in V_1 \mid \phi(v) = 0\} \subset V_1$ ist eine Unterdarstellung von (V_1, ρ_1) und $\text{Im}(\phi) := \{\phi(v) \mid v \in V_1\} \subset V_2$ ist eine Unterdarstellung von (V_2, ρ_2) . Ausserdem induziert ρ_2 eine Darstellung auf $\text{Coker}(\phi) = V_2 / \text{Im}(\phi)$.

(ii) Falls (V_1, ρ_1) und (V_2, ρ_2) irreduzibel sind, so ist ϕ invertierbar oder 0.

(iii) Angenommen, $\mathbb{K} = \mathbb{C}$ und $\phi' \in \text{Hom}_G(V_1, V_2)$ ist eine weitere solche G -lineare Abbildung. Falls dann (V_1, ρ_1) und (V_2, ρ_2) irreduzibel sind, und V_1 endlich-dimensional, sowie $\phi \neq 0$, so existiert ein $\alpha \in \mathbb{C}$ so, dass $\phi' = \alpha\phi$.

Beweis. (i) ist trivial. Z.B. impliziert $\phi(v) = 0$ sofort $\phi(\rho_1(g)v) = \rho_2(g)(\phi(v)) = \rho_2(g)(0) = 0$, d.h. $v \in \text{Ker}(\phi) \Rightarrow \rho_1(g)v \in \text{Ker}(\phi) \ \forall g \in G$. Dies ist genau die Definition eines invarianten Unterraums. Die beiden übrigen Aussagen werden analog gezeigt.

(ii) Falls $\phi \neq 0$, so ist $\text{Ker}(\phi) \neq V_1$. Da $\text{Ker}(\phi)$ G -invariant ist und V_1 irreduzibel, muss $\text{Ker}(\phi) = 0$ sein. Also ist ϕ injektiv. In ähnlicher Weise ist im Falle $\phi \neq 0$ $\text{Im}(\phi) \neq 0$, und da V_2 irreduzibel ist, folgt $\text{Im}(\phi) = V_2$. Damit ist ϕ surjektiv, mithin invertierbar.

(iii) Wegen (ii) und $\phi \neq 0$ ist ϕ invertierbar, es existiert also $\phi^{-1} \circ \phi' \in \text{Hom}_{\mathbb{C}}(V_1, V_1)$. Man prüft leicht, dass $\phi^{-1} \in \text{Hom}_G(V_2, V_1)$, und daher auch $\phi^{-1} \circ \phi' \in \text{Hom}_G(V_1, V_1)$. Nun hat über dem algebraisch abgeschlossenen Körper \mathbb{C} $\phi^{-1} \circ \phi'$ wie jeder Endomorphismus eines nicht-trivialen endlich-dimensionalen Vektorraums einen Eigenwert $\alpha \in \mathbb{C}$. Dies folgt aus der Tatsache, dass das charakteristische Polynom $P(\alpha) = \det(\phi^{-1} \circ \phi' - \alpha \text{id}_{V_1})$ gemäss 7.6 mindestens eine komplexe Nullstelle besitzt, und ist gleichbedeutend damit, dass $\phi^{-1} \circ \phi' - \alpha \text{id}_{V_1}$ nicht invertierbar ist. Da $\phi^{-1} \circ \phi' - \alpha \text{id}_{V_1}$ aber auch G -linear ist, muss sie wegen (ii) identisch verschwinden. Daraus folgt sofort $\phi' = \alpha\phi$. \square

Bemerkungen. Für endlich-dimensionale irreduzible Darstellungen schreibt man Aussage (iii) auch als

$$\text{Hom}_G(V_1, V_2) = (V_1^* \otimes V_2)^G = \begin{cases} \mathbb{C} & \text{falls } (V_1, \rho_1) \text{ isomorph zu } (V_2, \rho_2) \\ \{0\} & \text{sonst} \end{cases} \quad (19.24)$$

§ 19. GRUPPEN UND IHRE DARSTELLUNGEN

Hier haben wir in der ersten Gleichung zunächst 18.9 benutzt, um $\text{Hom}_{\mathbb{C}}(V_1, V_2) = V_1^* \otimes V_2$ zu schreiben. Dann haben wir mittels $\rho_1^*(g) := (\rho_1(g^{-1}))^*$ (vgl. (18.11)) V_1^* zu einer Darstellung aufgerüstet. Man prüft dann, dass unter dieser Identifikation $\text{Hom}_G(V_1, V_2)$ (vgl. (19.20)) mit dem Raum der invarianten Vektoren (vgl. (19.21)) im Tensorprodukt übereinstimmt.

Theorem 19.7. *Sei G abelsch und (V, ρ) eine endlich-dimensionale irreduzible Darstellung von G über \mathbb{C} . Dann ist V ein-dimensional, d.h. $\rho \in \text{Hom}(G, \mathbb{C}^\times)$.*

Beweis. Wir behaupten, dass für jedes $g_0 \in G$, $\rho(g_0) = \alpha(g_0) \text{id}_V$ für $\alpha(g_0) \in \mathbb{C}^\times$. Tatsächlich gilt ja für jedes $g \in G$

$$\rho(g_0)\rho(g) = \rho(g_0g) = (G \text{ abelsch}) = \rho(gg_0) = \rho(g)\rho(g_0) \quad (19.25)$$

d.h. $\rho(g_0) \in \text{Hom}_G(V, V)$. Mit dem Lemma von Schur folgt $\rho(g_0) = \alpha(g_0) \text{id}_V$, und $\alpha(g_0) \neq 0$, da $\rho(g_0)$ invertierbar ist. Dann aber ist jeder Unterraum von V invariant. Da (V, ρ) irreduzibel ist, muss V ein-dimensional sein. \square

Übungsaufgabe: Jede irreduzible Darstellung einer endlichen Gruppe ist endlich-dimensional.

Beispiel 19.8 (Die irreduziblen Darstellungen der zyklischen Gruppe). Da C_n von r_1 erzeugt wird (vgl. (19.9)), ist eine endlich-dimensionale irreduzible Darstellung über \mathbb{C} , welche nach 19.7 ein-dimensional ist, durch $\rho(r_1) \in \mathbb{C}^\times$ bestimmt. $(\rho(r_1))^n = \rho(e) = 1$ impliziert dann, dass $\rho(r_1)$ eine n -te Einheitswurzel ist, d.h. $\rho(r_1) = \exp\left(\frac{2\pi i l}{n}\right)$ für ein $l \in \{0, \dots, n-1\}$ (vgl. 2.2). Die Darstellungen (\mathbb{C}, ρ_l) mit

$$\rho_l(r_k) = (\rho_l(r_1))^k = \exp\left(\frac{2\pi i l \cdot k}{n}\right) \quad (19.26)$$

sind paarweise inäquivalent und geben damit die vollständige Liste der Irreps von C_n .

Beispiel 19.9 (Die irreduziblen Darstellungen der Diedergruppe). Wir verweisen auf die Präsentation der Gruppe in (19.11).

1. Ist (V, ρ) eine irreduzible Darstellung von D_n über \mathbb{C} , dann erhalten wir durch Einschränkung von ρ (als Homomorphismus $D_n \rightarrow GL(V)$) auf C_n eine Darstellung der zyklischen Gruppe.
2. Da V gemäss der Übung endlich-dimensional ist, enthält $(V, \rho|_{C_n})$ eine irreduzible Unterdarstellung (Übungsaufgabe). Wegen 19.8 existiert also ein $l \in \{0, \dots, n-1\}$ und ein ein-dimensionaler Unterraum $V_l \subset V$ mit

$$\rho(r_k)|_{V_l} = \exp\left(\frac{2\pi i l \cdot k}{n}\right) \text{id}_{V_l} . \quad (19.27)$$

$\rho(s_0)(V_l)$ ist wieder ein ein-dimensionaler Unterraum, der wegen (19.11), speziell wegen $r_k s_0 = s_0 r_{-k}$, invariant ist mit

$$\rho(r_k)|_{\rho(s_0)(V_l)} = \rho(s_0)\rho(r_{-k})|_{V_l} = \exp\left(\frac{-2\pi i l \cdot k}{n}\right) \text{id}_{\rho(s_0)(V_l)} \quad (19.28)$$

3. Wegen $s_0^2 = 1$ und $s_k = r_k \cdot s_0$ ist dann $V_l + \rho(s_0)(V_l)$ invariant unter ganz D_n , also wegen der Irreduzibilität von V gleich V . Damit ist V entweder ein- oder zwei-dimensional.

4. Gilt $V_l = \rho(s_0)(V_l) = V$, d.h. ist V ein-dimensional, so folgt aus $s_0^2 = 1$ dass $\rho(s_0) = \pm 1$, und aus $s_0 r_1 s_0 = r_{-1} = r_1^{-1}$ dass $r_1 = \pm 1$. Für n gerade sind alle vier Kombinationen möglich und es gibt vier paarweise inäquivalente ein-dimensionale Darstellungen. Für n ungerade muss $r_1 = 1$ sein, und es gibt nur zwei ein-dimensionale Irreps.

5. Gilt $V_l \cap \rho(s_0)(V_l) = \{0\}$, d.h. $V = V_l \oplus \rho(s_0)(V_l)$ ist zwei-dimensional, so haben wir bezüglich einer angepassten Basis die darstellenden Matrizen

$$\rho(r_k) = \begin{pmatrix} \exp\left(\frac{2\pi i k \cdot l}{n}\right) & 0 \\ 0 & \exp\left(\frac{-2\pi i k \cdot l}{n}\right) \end{pmatrix}, \quad \rho(s_0) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad (19.29)$$

Man prüft dann leicht, dass diese Matrizen für jedes l tatsächlich eine zwei-dimensionale Darstellung definieren, welche für $l \neq 0, \frac{n}{2}$ irreduzibel ist, für $l = 0$ oder $\frac{n}{2}$ aber in zwei ein-dimensionale Darstellungen zerlegbar ist. Ausserdem sieht man ein, dass die zu l und $n - l$ gehörigen Darstellungen äquivalent sind.⁴

Fazit: D_n besitzt für n ungerade genau zwei ein-dimensionale Irreps und $\frac{n-1}{2}$ zwei-dimensionale Irreps, und für n gerade genau vier ein-dimensionale Irreps und $\frac{n}{2} - 1$ zwei-dimensionale Irreps.

Alternierende Multilinearformen und das Dachprodukt

Als Anwendung der in den letzten beiden §§ entwickelten Konzepte bringen wir noch eine nützliche Verallgemeinerung des bereits vertrauten Determinantenkalküls, die in der Integrationstheorie (siehe § 23) sowie bei fermionischen Mehrteilchensystemen eine wichtige Rolle spielt.

Wir erinnern dazu zunächst an die Def. 18.7 des Vektorraums $L(V_1, \dots, V_k; W)$ der multi-linearen Abbildungen $V_1 \times \dots \times V_k \rightarrow W$, beschränken uns aber jetzt auf den Fall, dass $V_1 = V_2 = \dots = V_k = V$ ein fester endlich-dimensionaler Vektorraum mit $n := \dim(V)$ und $W = \mathbb{K}$ (wie immer, = \mathbb{C} oder \mathbb{R}). Es gilt dann die Identifikation (vgl. (18.56))

$$L(V, \dots, V; \mathbb{K}) = \underbrace{(V \otimes \dots \otimes V)}_{k \text{ Mal}}^* = \underbrace{V^* \otimes \dots \otimes V^*}_{k \text{ Mal}} =: (V^*)^{\otimes k} \quad (19.30)$$

was sich mit (18.50) für $k = 0$ noch in natürlicher Weise zu $V^{\otimes 0} := \mathbb{K}$ fortsetzt.

· Dieser Vektorraum trägt eine natürliche Darstellung der symmetrischen Gruppe S_k , definiert als lineare Fortsetzung von

$$S_k \ni \sigma \mapsto c_\sigma(\lambda_1 \otimes \lambda_2 \otimes \dots \otimes \lambda_k) := \lambda_{\sigma(1)} \otimes \lambda_{\sigma(2)} \dots \otimes \lambda_{\sigma(k)} \quad (19.31)$$

für $\lambda_1, \dots, \lambda_k \in V^*$. Als multilineare Abbildungen geschrieben ist dies

$$(c_\sigma(T))(v_1, \dots, v_k) := T(v_{\sigma(1)}, \dots, v_{\sigma(k)}) \quad (19.32)$$

⁴Bezüglich der obigen Basen stellt $\phi := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ den Isomorphismus zwischen V_l und V_{n-l} dar.

§ 19. GRUPPEN UND IHRE DARSTELLUNGEN

Beispielsweise wird für $k = 2$ das nicht-triviale Element von $S_2 = \mathbb{Z}/2\mathbb{Z}$ durch die Abbildung c_{V^*, V^*} aus (18.49) dargestellt.

· Die Darstellung (19.31) ist im Allgemeinen nicht irreduzibel. Sie enthält als Unterdarstellung den Raum der invarianten oder symmetrischen Tensoren, welchen wir mit Hilfe von (19.30) schreiben können als

$$\text{Sym}^k V^* := \{T \in L(V, \dots, V; \mathbb{K}) \mid T(v_1, \dots, v_k) = T(v_{\sigma(1)}, \dots, v_{\sigma(k)}) \forall \sigma \in S_k\} \quad (19.33)$$

und der im nächsten § eine Rolle spielt, sowie ausserdem den Raum der “alternierende” oder (total) anti-symmetrischen Multilinearformen

$$\wedge^k V^* := \{T \in L(V, \dots, V; \mathbb{K}) \mid T = \text{sgn}(\sigma)c_\sigma(T) \forall \sigma \in S_k\} \quad (19.34)$$

wobei $\text{sgn} : S_k \rightarrow \{\pm 1\}$ den Signum- oder Paritätshomomorphismus in die multiplikative Gruppe $\{\pm 1\}$ bezeichnet. $((V^*)^{\otimes k}, c)$ enthält im Allgemeinen noch eine Reihe weiterer Unterdarstellungen, auf die wir aber hier nicht eingehen wollen. Vgl. für $V = \mathbb{K}$, $k = n$ auch mit der in (19.18) definierten Darstellung von S_n .

· Elemente $\omega \in \wedge^k V^*$ heissen meistens auch einfach k -Formen auf V .

· Wir definieren dann für $\lambda_1, \dots, \lambda_k \in V^*$ das *Dachprodukt* $\lambda_1 \wedge \dots \wedge \lambda_k \in \wedge^k V^*$ durch

$$\lambda_1 \wedge \dots \wedge \lambda_k := \sum_{\sigma \in S_k} \text{sgn}(\sigma)c_\sigma(\lambda_1 \otimes \dots \otimes \lambda_k) \quad (19.35)$$

(Mit anderen Worten ist für $v_1, \dots, v_k \in V$:

$$\begin{aligned} (\lambda_1 \wedge \dots \wedge \lambda_k)(v_1, \dots, v_k) &:= \sum_{\sigma \in S_k} \text{sgn}(\sigma)\lambda_1(v_{\sigma(1)}) \cdot \lambda_2(v_{\sigma(2)}) \cdot \dots \cdot \lambda_k(v_{\sigma(k)}) \\ &= \det(\lambda_i(v_j))_{\substack{i=1, \dots, k \\ j=1, \dots, k}} \end{aligned} \quad (19.36)$$

gemäss der Leibnizschen Formel für die Entwicklung der Determinanten einer Matrix.)

· Es gilt dann: Ist (e_1, \dots, e_n) eine Basis von V mit dualer Basis $(\epsilon^1, \dots, \epsilon^n)$ von V^* , so ist

$$\{\epsilon^{i_1} \wedge \dots \wedge \epsilon^{i_k} \mid 1 \leq i_1 < \dots < i_k \leq n\} \quad (19.37)$$

eine Basis von $\wedge^k V^*$. Daraus folgt

$$\dim \wedge^k V^* = \binom{n}{k} \quad (19.38)$$

Insbesondere ist $\wedge^n V^*$ ein-dimensional und $\wedge^k V^* = \{0\}$ für $k > n$.

· Wegen (19.37) können wir dann für beliebige $k, l \geq 1$ das Dachprodukt

$$\begin{aligned} \cdot \wedge \cdot : \wedge^k V^* \times \wedge^l V^* &\rightarrow \wedge^{k+l} V^* \\ (\omega, \alpha) &\mapsto \omega \wedge \alpha \end{aligned} \quad (19.39)$$

definieren durch die Formel

$$(\lambda_1 \wedge \dots \wedge \lambda_k) \wedge (\mu_1 \wedge \dots \wedge \mu_l) := \lambda_1 \wedge \dots \wedge \lambda_k \wedge \mu_1 \wedge \dots \wedge \mu_l \quad (19.40)$$

für faktorisierbare k - und l -Formen. (Für k oder $l = 0$ setzt man noch $\omega \wedge a = a \wedge \omega := a \cdot \omega$.) Etwas expliziter gilt (Übungsaufgabe):

$$\omega \wedge \alpha(v_1, \dots, v_{k+l}) = \frac{1}{k!l!} \sum_{\sigma \in S_{k+l}} \operatorname{sgn}(\sigma) \omega(v_{\sigma(1)}, \dots, v_{\sigma(k)}) \cdot \alpha(v_{\sigma(k+1)}, \dots, v_{\sigma(k+l)}) \quad (19.41)$$

· Für $\omega_i \in \wedge^{k_i} V^*$ gelten dann die Rechenregeln:

$$\begin{aligned} (\omega_1 \wedge \omega_2) \wedge \omega_3 &= \omega_1 \wedge (\omega_2 \wedge \omega_3) \\ \omega_1 \wedge \omega_2 &= (-1)^{k_1 k_2} \omega_2 \wedge \omega_1 \end{aligned} \quad (19.42)$$

Man sagt: Der Raum

$$\wedge V^* = \bigoplus_{k=0}^n \wedge^k V^*, \quad (19.43)$$

ausgerüstet mit dem Dachprodukt ist eine $\mathbb{Z}/2$ -graduieret-kommutative Algebra, genannt die äussere Algebra (oder auch Grassmann Algebra) über V^* .

$$\dim \wedge V^* = \sum_{k=0}^n \binom{n}{k} = 2^n \quad (19.44)$$

Geometrische Interpretation: Für einen n -dimensionalen reellen Vektorraum V ist wegen (19.38) $\wedge^n V^*$ ein-dimensional. Man nennt ein nicht-verschwindendes Element $\omega \in \wedge^n V^*$, $\omega \neq 0$ eine (orientierte) Volumenform. Sie ordnet nämlich einem linear unabhängigen Satz von Vektoren (b_1, \dots, b_n) (m.a.W., einer Basis von V) eine Zahl $\omega(b_1, \dots, b_n) \in \mathbb{R}$ zu, welche die "anschaulichen" Erwartungen an das gezeichnete Volumen des von (b_1, \dots, b_n) aufgespannten Parallelotops erfüllt.

Ist W ein weiterer n -dimensionaler Vektorraum mit einer Volumenform $\alpha \in \wedge^n W^*$, und $A : W \rightarrow V$ eine lineare Abbildung, so wird durch

$$A^*(\omega)(c_1, \dots, c_n) := \omega(Ac_1, \dots, Ac_n) \quad (19.45)$$

eine alternierende n -Form $A^*(\omega)$ definiert. (Übungsaufgabe). Wegen $\dim \wedge^n W^* = 1$ gilt $A^*(\omega) = d \cdot \alpha$ für ein $d \in \mathbb{R}$. Dabei misst die Proportionalitätskonstante d die Streckung von Parallelotopvolumina. (In Standardsituationen, z.B. $(W, \alpha) = (V, \omega)$ gilt $d = \det A$.)

Als Verallgemeinerung davon definiert für $k \leq n$ eine alternierende k -Form einen orientierten Flächeninhalt für k -dimensionale Parallelootope in einem n -dimensionalen umgebenden Raum. Im Unterschied zu der Situation für $k = n$ sind k -Formen nicht mehr bis auf Skalierung eindeutig, und können auf verschiedene Arten entarten.

Betrachten wir beispielsweise für $V = \mathbb{R}^3$ mit Standardbasis die 2-Form

$$\omega = \epsilon^1 \wedge \epsilon^2 + 2\epsilon^2 \wedge \epsilon^3 \quad (19.46)$$

Dann ordnet ω dem Standardquadrat in der 1-2-Ebene den Flächeninhalt 1 zu, dem Standardquadrat in der 2-3-Ebene den Flächeninhalt 2, und dem Standardquadrat in

§ 20. EUKLIDISCHE UND UNITÄRE RÄUME

der 1-3-Ebene den Inhalt 0. Das von $(e_1 + e_2 + e_3, 2e_3)$ aufgespannte Parallelogramm erhält den Flächeninhalt

$$\begin{aligned}\omega(e_1 + e_2 + e_3, 2e_3) &= (\epsilon^1 \wedge \epsilon^2)(e_1 + e_2 + e_3, 2e_3) + 2(\epsilon^2 \wedge \epsilon^3)(e_1 + e_2 + e_3, 2e_3) \\ &= (\epsilon^1 \otimes \epsilon^2 - \epsilon^2 \otimes \epsilon^1)(e_1 + e_2 + e_3, 2e_3) \\ &\quad + 2(\epsilon^2 \otimes \epsilon^3 - \epsilon^3 \otimes \epsilon^2)(e_1 + e_2 + e_3, 2e_3) \\ &= 0 + 2\epsilon^2(e_2) \cdot \epsilon^3(2e_3) = 4\end{aligned}\tag{19.47}$$

§ 20 Euklidische und unitäre Räume

In den Beispielen 19.8 und 19.9 leuchtet durch, dass die Eigenwerte der darstellenden linearen Abbildungen $\rho(g)$ für $g \in G$ zur Unterscheidung und Charakterisierung von endlich-dimensionalen komplexen Darstellungen einer Gruppe G dienen können. Dies ist eine Verallgemeinerung der Aussage, dass Eigenwerte einzelner linearer Abbildungen $A \in \text{Hom}_{\mathbb{K}}(V, V)$, ebenso wie etwa Determinante und Spur, “basisunabhängige” Information über A enthalten. Es stellt sich heraus, dass diese Information in einem gewissen Sinn “vollständig” ist, wenn unser Vektorraum mit einem mit A bzw. ganz ρ verträglichen inneren Produkt ausgerüstet ist. Solche inneren Produkte und die daraus abgeleiteten Normen (s. § 4 und (10.29)) sind ganz allgemein das Werkzeug, um Objekte aus Vektorräumen mit physikalischen Messgrößen in Verbindung zu setzen. Wir wiederholen auch hier erst wieder die hervorstechenden Aussagen aus der linearen Algebra.

Definition 20.1. Ein *euklidisches inneres Produkt* auf einem reellen Vektorraum V ist eine positiv definite symmetrische Bilinearform $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}$. Ein *euklidischer Raum* ist ein reeller Vektorraum mit einem euklidischen inneren Produkt.

Eine *Sesquilinearform* auf einem komplexen Vektorraum V ist eine Abbildung

$$\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{C}\tag{20.1}$$

welche im zweiten Argument linear, und im ersten konjugiert-linear (auch “anti-” oder “halb-linear” genannt) ist, d.h. $\forall v_1, v_2, w_1, w_2 \in V, \lambda, \mu \in \mathbb{C}$ gilt:

$$\langle \lambda v_1 + v_2, \mu w_1 + w_2 \rangle = \bar{\lambda}\mu \langle v_1, w_1 \rangle + \bar{\lambda} \langle v_1, w_2 \rangle + \mu \langle v_2, w_1 \rangle + \langle v_2, w_2 \rangle\tag{20.2}$$

($\bar{\lambda}$ = komplexe Konjugation aus (2.6)). Eine Sesquilinearform heisst (konjugiert) symmetrisch, falls für alle $v, w \in V$

$$\langle w, v \rangle = \overline{\langle v, w \rangle}\tag{20.3}$$

Ein *hermitesches inneres Produkt* auf einem komplexen Vektorraum ist eine positiv definite symmetrische Sesquilinearform. Ein *unitärer* (oder Hermitescher) *Raum* ist ein komplexer Vektorraum V mit einem hermiteschen inneren Produkt.

Bemerkungen. Aus der positiv Definitheit eines euklidischen oder hermiteschen inneren Produkts, das heisst

$$\langle v, v \rangle > 0 \quad \text{für } v \neq 0\tag{20.4}$$

folgt insbesondere, dass $\langle \cdot, \cdot \rangle$ nicht entartet ist, d.h. $\langle v, w \rangle = 0 \ \forall w \in V \Rightarrow v = 0$. Daraus wiederum folgt, dass die Abbildung

$$V \mapsto V^* \quad v \mapsto (w \mapsto \langle v, w \rangle) \tag{20.5}$$

injektiv ist. Für $\mathbb{K} = \mathbb{R}$ ist diese Abbildung linear, für $\mathbb{K} = \mathbb{C}$ konjugiert linear. Im Falle $\dim(V) < \infty$ folgt aus der Injektivität, dass (20.5) ein (konjugiert) linearer Isomorphismus ist.

· Für $V = \mathbb{R}^n$ ist

$$\langle x, y \rangle := \sum_{i=1}^n x^i y^i \tag{20.6}$$

ein euklidisches, für $V = \mathbb{C}^n$

$$\langle z, w \rangle := \sum_{i=1}^n \bar{z}^i w^i \tag{20.7}$$

ein hermitesches inneres Produkt. Durch Argumente, welche denen zu 20.2 sehr ähneln, kann man zeigen, dass jedes euklidische bzw. hermitesche innere Produkt auf einem endlich-dimensionalen Vektorraum zu den Standardformen äquivalent ist, das heisst es existiert jeweils eine Orthonormalbasis (ONB) $B : \mathbb{K}^n \cong V$, bezüglich der $\langle \cdot, \cdot \rangle$ die Form (20.6) bzw. (20.7) hat.

Fakten 20.2 (Selbstadjungierte Endomorphismen). Sei $(V, \langle \cdot, \cdot \rangle)$ ein euklidischer oder unitärer Vektorraum. Eine lineare Abbildung $A \in \text{Hom}_{\mathbb{K}}(V, V)$ heisst *symmetrisch* oder *selbst-adjungiert*⁵, falls

$$\langle v, Aw \rangle = \langle Av, w \rangle \quad \forall v, w \in V \tag{20.9}$$

Es sei nun V ein endlich-dimensionaler euklidischer oder unitärer Vektorraum, und $A \in \text{Hom}_{\mathbb{K}}(V, V)$ selbst-adjungiert. Dann existiert eine ONB (b_1, \dots, b_n) von V aus Eigenvektoren von A : Mit reellen Zahlen $\lambda_i \in \mathbb{R} \subset \mathbb{K}$ gilt

$$A(b_i) = \lambda_i \cdot b_i \quad \forall i \quad \text{und} \quad \langle b_i, b_j \rangle = \begin{cases} 1 & \text{für } i = j \\ 0 & \text{sonst} \end{cases} \tag{20.10}$$

Beweisideen. Für $\dim V = 0$ ist die Aussage leer. In der Absicht, für $\dim V \geq 1$ einen induktiven Beweis nach der Dimension von V zu führen, zeigt man zunächst die Existenz eines Eigenwertes $\lambda \in \mathbb{R} \subset \mathbb{K}$ von A : Für $\mathbb{K} = \mathbb{C}$ folgt wie im Beweis von 19.6 aus dem Fundamentalsatz der Algebra 7.6 zunächst die Existenz eines

⁵Für eine beliebige lineare Abbildung $A \in \text{Hom}_{\mathbb{K}}(V, V)$ und $v \in V$ ist die Abbildung $V \ni w \mapsto \langle v, Aw \rangle$ eine Linearform auf V . Wegen der Nichtentartetheit des inneren Produkts folgt daraus im endlich-dimensionalen Fall mit Hilfe des Isomorphismus (20.5), dass ein eindeutiger Vektor $A^{\text{ad}}v \in V$ existiert, mit dem

$$\langle v, Aw \rangle = \langle A^{\text{ad}}v, w \rangle \quad \forall w \in V \tag{20.8}$$

Die Zuordnung $v \mapsto A^{\text{ad}}v$ ist linear (auch für $\mathbb{K} = \mathbb{C}$!) und eindeutig durch die Eigenschaft (20.8) bestimmt. Sie heisst die zu A adjungierte Abbildung. Die Bedingung (20.9) ist dann die Gleichung $A = A^{\text{ad}}$. Für $\mathbb{K} = \mathbb{R}$ schreibt man (in Anlehnung an Matrizen) auch $A^{\text{ad}} = A^T$, für $\mathbb{K} = \mathbb{C}$ auch $A^{\text{ad}} = A^\dagger$.

§ 20. EUKLIDISCHE UND UNITÄRE RÄUME

komplexen Eigenwertes $\lambda \in \mathbb{C}$. Mit einem Eigenvektor $v \neq 0$ zum Eigenwert λ folgt aber aus

$$\bar{\lambda}\langle v, v \rangle = \langle Av, v \rangle = \langle v, Av \rangle = \lambda\langle v, v \rangle \quad (20.11)$$

und $\langle v, v \rangle \neq 0$, dass tatsächlich $\bar{\lambda} = \lambda$, d.h. $\lambda \in \mathbb{R} \subset \mathbb{C}$. Für $\mathbb{K} = \mathbb{R}$ betrachtet man A auf der Komplexifizierung $V^{\mathbb{C}}$ von V (s. 18.10) und zieht die eben gewonnene Aussage wieder auf V zurück.

- Ist dann $v \in V \setminus \{0\}$ ein Eigenvektor von A zu $\lambda \in \mathbb{R}$, so ist $\hat{v} := \frac{v}{\underbrace{\sqrt{\langle v, v \rangle}}_{>0}}$ ein Eigenvektor mit $\langle \hat{v}, \hat{v} \rangle = 1$. Für $\dim V = 1$ ist man hier fertig.
- Für $\dim V > 1$ ist das orthogonale Komplement

$$\hat{v}^{\perp} := \{w \in V \mid \langle \hat{v}, w \rangle = 0\} \quad (20.12)$$

ein Vektorraum der Dimension $\dim(\hat{v}^{\perp}) = \dim V - 1$. Wegen

$$w \in \hat{v}^{\perp} \Rightarrow \langle \hat{v}, w \rangle = 0 \Rightarrow \langle \hat{v}, Aw \rangle = \langle A\hat{v}, w \rangle = \lambda\langle \hat{v}, w \rangle = 0 \Rightarrow Aw \in \hat{v}^{\perp} \quad (20.13)$$

ist \hat{v}^{\perp} invariant unter A . Die Anwendung der Induktionsannahme auf die Einschränkung von A auf \hat{v}^{\perp} liefert dann eine ONB aus Eigenvektoren von A für \hat{v}^{\perp} , welche durch \hat{v} zu einer ONB von V vervollständigt wird. \square

Die gleichen Ideen in leicht abgewandelter Form stecken hinter den Aussagen zu den Normalformen von orthogonalen, unitären⁶, und allgemein von normalen Endomorphismen von V . Sie verallgemeinern sich unter geeigneten Voraussetzung auch zu Darstellungen von Gruppen:

Definition 20.3. Eine komplexe Darstellung (V, ρ) von G heisst *unitär*, falls auf V ein hermitesches inneres Produkt $\langle \cdot, \cdot \rangle$ existiert, welches *G-invariant* ist, d.h.

$$\langle \rho(g)v, \rho(g)w \rangle = \langle v, w \rangle \quad \forall v, w \in V, \forall g \in G \quad (20.14)$$

M.a.W., für alle $g \in G$ ist $\rho(g)$ eine unitäre Abbildung von V .

Proposition 20.4. Sei (V, ρ) eine unitäre Darstellung von G , und $U \subset V$ eine Unterdarstellung. Dann existiert eine weitere Unterdarstellung $W \subset V$ so, dass als Darstellungen von G

$$V = U \oplus W \quad (20.15)$$

Beweis. Sei $\langle \cdot, \cdot \rangle$ ein G -invariantes hermitesches inneres Produkt auf V . Setze $W := U^{\perp} = \{w \in V \mid \langle w, u \rangle = 0 \forall u \in U\}$ (das orthogonale Komplement). Falls $w \in U^{\perp}$, so ist $\forall u \in U$

$$\langle \rho(g)w, u \rangle = \langle w, \rho(g)^{-1}u \rangle = 0 \quad (20.16)$$

(da ja wegen der Invarianz von U $\rho(g)^{-1}u \in U$), also $\rho(g)w \in U^{\perp}$. Dies bedeutet, dass U^{\perp} auch G -invariant ist. \square

Die Existenz eines invarianten inneren Produktes mag wie eine starke Voraussetzung aussehen. Für viele in der Praxis vorkommende Gruppen ist sie aber automatisch erfüllt.

⁶Eine lineare Abbildung $U \in \text{Hom}_{\mathbb{K}}(V, V)$ heisst orthogonal/unitär, falls $U^{\text{ad}} = U^{-1}$, d.h. $\langle Uv, Uw \rangle = \langle v, w \rangle \forall v, w \in V$

Proposition 20.5. *Sei jetzt G eine endliche Gruppe und (V, ρ) eine komplexe Darstellung von G . Dann ist (V, ρ) unitär.*

Beweis. Sei (\cdot, \cdot) ein beliebiges (nicht notwendigerweise G -invariantes) hermitesches inneres Produkt (so etwas existiert immer). Dann ist

$$\langle v, w \rangle := \sum_{h \in G} (\rho(h)v, \rho(h)w) \quad (20.17)$$

ein G -invariantes hermitesches inneres Produkt:

$$\langle \rho(g)v, \rho(g)w \rangle = \sum_{h \in G} (\rho(hg)v, \rho(hg)w) = \sum_{hg=h' \in G} (\rho(h')v, \rho(h')w) = \langle v, w \rangle \quad (20.18)$$

und die anderen Eigenschaften sind klar. (Allerdings ist es wichtig, als Startpunkt ein positiv definites inneres Produkt zu verwenden. Ein nur nicht-entartetes inneres Produkt gemischter Signatur kann beim Mitteln über die Gruppe entarten.) \square

Korollar 20.6 (Vollständige Zerlegbarkeit). *Sei V eine endlich-dimensionale komplexe Darstellung einer endlichen Gruppe G . Dann existieren (paarweise nicht-isomorphe) irreduzible Darstellungen V_1, \dots, V_k und Multiplizitäten $a_1, \dots, a_k \in \mathbb{N}$ so, dass*

$$V \cong \bigoplus_{i=1}^k V_i \otimes \mathbb{C}^{a_i} = \bigoplus_{i=1}^k V_i^{\oplus a_i} \quad (20.19)$$

und diese Zerlegung ist bis auf Reihenfolge und Isomorphismus eindeutig.

Beweis. Die Existenz folgt durch wiederholtes Anwenden aus Prop. 20.4. Für die Eindeutigkeit sei

$$V = \bigoplus_{j=1}^l W_j \otimes \mathbb{C}^{b_j} \quad (20.20)$$

eine weitere solche Zerlegung. Da $\text{id}_V|_{V_i} \rightarrow V$ ein G -Morphismus ist, muss nach dem Lemma von Schur 19.6 das Bild eine irreduzible Darstellung isomorph zu V_i sein. Wiederhole. \square

Beachte, dass falls G nicht endlich ist (und auch nicht kompakt), vollständige Zerlegbarkeit verloren gehen kann. Betrachte z.B.

$$(\mathbb{R}, +) \ni a \mapsto \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \in GL(2, \mathbb{R}) \quad (20.21)$$

Wir verlassen an dieser Stelle die allgemeinen Betrachtungen zur Darstellungstheorie endlicher Gruppen und geben statt dessen noch zwei Querverbindungen zwischen der linearen Algebra der letzten beiden §§ und der Differentialrechnung der Kap. 4 und 5.

Erzeugende von Matrix-Lie-Gruppen

Die erste Verbindung entsteht als Anwendung der Exponentialfunktion auf die Untersuchung der Struktur kontinuierlicher Gruppen (s. Def. S. 127)

§ 20. EUKLIDISCHE UND UNITÄRE RÄUME

Zur Erinnerung: Wir hatten im § 5 die Exponentialfunktion \exp durch die Potenzreihe (5.31) für Argumente in ganz allgemeinen Banach-Algebren eingeführt.

· Das Paradebeispiel für solche Banach-Algebren waren dabei aber bereits die Endomorphismen $(\text{Hom}_{\mathbb{R}}(V, V), \|\cdot\|)$ eines endlich-dimensionalen normierten reellen Vektorraums, $(V, \|\cdot\|_V)$, siehe 4.10. Hierbei ist die Operatornorm eines solchen Endomorphismus definiert als (4.27)

$$\|A\| := \sup\{\|Av\| \mid \|v\|_V = 1\} \quad (20.22)$$

· Seinerseits war das motivierende Beispiel für einen normierten Vektorraum der \mathbb{R}^n mit seiner vom euklidischen Standardprodukt (20.6) induzierten Norm, $\|x\|_2 := \sqrt{\langle x, x \rangle}$.

· Die wichtigste analytische Eigenschaft der normierten Räume $(\mathbb{R}^n, \|\cdot\|_2)$ und $(\text{Hom}_{\mathbb{R}}(V, V), \|\cdot\|)$ war dabei ihre metrische Vollständigkeit.

· In ebendieser Weise induziert das hermitesche Standardprodukt (20.7) eine Norm auf \mathbb{C}^n , und folglich auf $\text{Hom}_{\mathbb{C}}(\mathbb{C}^n, \mathbb{C}^n)$: Die Argumente auf S. 25 bzw. 28, insbesondere die für die Gültigkeit der Cauchy-Schwarz-Ungleichung (C.S.U.),

$$|\langle v, w \rangle| \leq \|v\| \cdot \|w\| \quad \forall v, w \in V \quad (20.23)$$

lassen sich leicht für einen allgemeinen unitären Vektorraum anpassen.

· Die metrische Vollständigkeit des \mathbb{C}^n unter dieser Norm folgt vielleicht am einfachsten aus der Tatsache, dass unter geeigneter Identifikation $\mathbb{C}^n \cong \mathbb{R}^{2n}$ (vgl. 18.10) (20.7) identisch zu (20.6) ist.

· Wegen der Existenz von Orthonormalbasen sind allgemeine endlich-dimensionale euklidische oder unitäre Räume äquivalent zu den Standardformen. Im Folgenden steht also \mathbb{K}^n für \mathbb{R}^n , ausgerüstet mit (20.6), bzw. für \mathbb{C}^n mit (20.7).

· Gemäss den Betrachtungen im § 5 konvergiert die Exponentialreihe also absolut auf ganz $\text{Hom}_{\mathbb{K}}(\mathbb{K}^n, \mathbb{K}^n) = \text{Mat}(n, \mathbb{K})$, und gleichmässig auf der Kugel $B_R(0) \subset \text{Mat}(n, \mathbb{K})$ für jedes $R > 0$. Nach den Betrachtungen des § 10 ist \exp daher differenzierbar im Sinne der Differentialrechnung mehrerer *reeller* Veränderlicher.⁷ Gemäss (11.18) ist das Differential bei $0 \in \text{Mat}(n, \mathbb{K})$ die Identität, d.h. ausgedrückt durch die Richtungsableitungen (vgl. 10.4) gilt

$$D \exp(0)(A) = \left. \frac{d}{dt} \right|_{t=0} \exp(tA) = A \quad \forall A \in \text{Mat}(n, \mathbb{K}) \quad (20.24)$$

· Die hierfür und auch sonst wichtigste charakteristische Eigenschaft der Exponentialfunktion ist die “Funktionalgleichung” genannte Identität (s. 5.13)

$$\exp(A) \cdot \exp(B) = \exp(A + B) \quad \forall A, B \in \text{Mat}(n, \mathbb{K}) \mid AB = BA \quad (20.25)$$

· Aus dieser Gleichung folgt nämlich (vgl. (5.41)), dass für alle $A \in \text{Mat}(n, \mathbb{K})$

(i) $\exp(A)$ invertierbar ist, und

(ii) für alle $t, s \in \mathbb{R}$

$$\exp(tA) \exp(sA) = \exp((t + s)A) \quad (20.26)$$

⁷Im Falle $\mathbb{K} = \mathbb{C}$ ist \exp sogar komplex differenzierbar in mehreren Variablen. Wir hatten diesen Begriff aber nicht eingeführt und werden ihn im Folgenden auch nicht benutzen können, da die komplexe Konjugation in der Definition des hermiteschen inneren Produkts nicht mit der Holomorphie verträglich ist.

Man sagt in der Sprache der Gruppentheorie: Die Abbildung

$$\mathbb{R} \ni t \mapsto \exp(tA) \in GL(n, \mathbb{K}) \quad (20.27)$$

ist für alle A eine *ein-Parameter Untergruppe* von $GL(n, \mathbb{K})$.

· Ein-Parameter Untergruppen spielen für die Untersuchung der kontinuierlichen Gruppen eine ähnliche Rolle wie die von einzelnen Elementen endlicher Gruppen erzeugten zyklischen Untergruppen (vgl. $C_n := \langle r_1 \rangle \subset D_n$ in (19.11)).

· Denn wie wir schon im Zusammenhang des Umkehrsatzes im §12 (s. um (12.16) herum) festgehalten hatten, folgt aus der Tatsache, dass das Differential $D \exp(0)$ vollen Rang hat, dass \exp eine offene Umgebung von $0 \in \text{Mat}(n, \mathbb{K})$ diffeomorph (bijektiv und in beiden Richtungen stetig differenzierbar) auf eine offene Umgebung des neutralen Elements $\text{id} \in GL(n, \mathbb{K})$ abbildet. Im Sinne dieses Zusammenhangs nennt man die Elemente A des Vektorraums $\text{Mat}(n, \mathbb{K})$ die *infinitesimal Erzeugenden* der Gruppe $GL(n, \mathbb{K})$.⁸

· In der Physik und Geometrie erhalten die infinitesimal Erzeugenden derjenigen ein-Parameter Untergruppen eine besondere Bedeutung, die mit dem jeweils gegebenen inneren Produkt verträglich sind.

Definition 20.7. Für $\mathbb{K} = \mathbb{R}$ heisst die Gruppe

$$\begin{aligned} O(n) &:= \{R \in GL(n, \mathbb{R}) \mid \langle Rv, Rw \rangle = \langle v, w \rangle \forall v, w \in \mathbb{R}^n\} \\ &= \{R \in GL(n, \mathbb{R}) \mid R^T = R^{-1}\} \end{aligned} \quad (20.28)$$

die *orthogonale Gruppe*. Wir schreiben für die Menge ihrer infinitesimal Erzeugenden

$$\mathfrak{so}(n) := \{A \in \text{Mat}(n, \mathbb{R}) \mid \exp(tA) \in O(n) \forall t \in \mathbb{R}\} \quad (20.29)$$

Für $\mathbb{K} = \mathbb{C}$ heisst die Gruppe

$$\begin{aligned} U(n) &:= \{U \in GL(n, \mathbb{C}) \mid \langle Uv, Uw \rangle = \langle v, w \rangle \forall v, w \in \mathbb{C}^n\} \\ &= \{U \in GL(n, \mathbb{C}) \mid U^\dagger = U^{-1}\} \end{aligned} \quad (20.30)$$

die *unitäre Gruppe*. (Zur Notation siehe Fussnote 5.) Wir schreiben für die Menge ihrer infinitesimal Erzeugenden

$$\mathfrak{u}(n) := \{A \in \text{Mat}(n, \mathbb{C}) \mid \exp(tA) \in U(n) \forall t \in \mathbb{R}\} \quad (20.31)$$

Proposition 20.8. (i) Für $\mathbb{K} = \mathbb{R}$ gilt:

$$\mathfrak{so}(n) = \{A \in \text{Mat}(n, \mathbb{R}) \mid A^T = -A\} \quad (20.32)$$

Für $\mathbb{K} = \mathbb{C}$ gilt:

$$\mathfrak{u}(n) = \{A \in \text{Mat}(n, \mathbb{C}) \mid A^\dagger = -A\} \quad (20.33)$$

(ii) $\mathfrak{so}(n)$ bzw. $\mathfrak{u}(n)$ sind reelle Vektorräume der Dimension $\frac{n(n-1)}{2}$ bzw. n^2 , und abgeschlossen unter der Kommutator-Bildung, d.h. für $A_1, A_2 \in \mathfrak{so}(n)$ (bzw. $\mathfrak{u}(n)$) gilt

$$[A_1, A_2] := A_1 \cdot A_2 - A_2 \cdot A_1 \in \mathfrak{so}(n) \text{ (bzw. } \mathfrak{u}(n)\text{)} \quad (20.34)$$

⁸Tatsächlich erzeugt dann eine offene Umgebung von id (evtl. zusammen mit einer endlichen Menge) $GL(n, \mathbb{K})$ algebraisch.

§ 20. EUKLIDISCHE UND UNITÄRE RÄUME

Beweis. Wir zeigen (i) für $\mathbb{K} = \mathbb{C}$, der Rest sei Übung. Zunächst halten wir fest, dass $\text{Mat}(n, \mathbb{C}) \ni A \mapsto A^\dagger \in \text{Mat}(n, \mathbb{C})$ eine stetige Abbildung ist. Daher gilt nach dem Folgenkriterium für Stetigkeit 6.5 $\forall A \in \text{Mat}(n, \mathbb{C})$:

$$\begin{aligned} (\exp(A))^\dagger &= \left(\lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{A^k}{k!} \right)^\dagger = \lim_{n \rightarrow \infty} \left(\sum_{k=0}^n \frac{A^k}{k!} \right)^\dagger = \lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{(A^\dagger)^k}{k!} \\ &= \exp(A^\dagger) \end{aligned} \quad (20.35)$$

Beh.: $A \in \mathfrak{u}(n) \Leftrightarrow A^\dagger = -A$

· “ \Rightarrow ”: Gilt

$$\exp(tA^\dagger) = \exp(tA)^\dagger = \exp(tA)^{-1} = \exp(-tA) \quad \forall t \in \mathbb{R}, \quad (20.36)$$

so folgt durch Ableiten

$$\left. \frac{d}{dt} \right|_{t=0} \exp(tA^\dagger) = \left. \frac{d}{dt} \right|_{t=0} \exp(-tA) \Rightarrow A^\dagger = -A \quad (20.37)$$

· “ \Leftarrow ”: Gilt umgekehrt $A^\dagger = -A$, so folgt mit Hilfe von (20.35)

$$(\exp(tA))^\dagger = \exp(tA^\dagger) = \exp(-tA) = (\exp(tA))^{-1} \quad \forall t \in \mathbb{R} \quad (20.38)$$

(Im ersten Schritt haben wir noch benutzt, dass zwar $A \mapsto A^\dagger$ konjugiert linear ist, aber $t \in \mathbb{R}$.) \square

Bemerkungen. · Es sei noch zweimal betont, dass auch im Falle $\mathbb{K} = \mathbb{C}$ $\mathfrak{u}(n)$ kein komplexer, sondern nur ein reeller Vektorraum ist. Dies soll heissen: Die Addition ist zwar die von $\text{Mat}(n, \mathbb{C})$ geerbte, $\mathfrak{u}(n)$ ist aber nur unter Multiplikation mit $\mathbb{R} \subset \mathbb{C}$ abgeschlossen: Ist $A^\dagger = -A$ (anti-selbst-adjungiert), so ist $(iA)^\dagger = \bar{i}A^\dagger = (-i) \cdot (-A) = iA$, d.h. iA ist selbst-adjungiert.

· Für $\mathbb{K} = \mathbb{C}$ lässt sich aber dann der Vektorraum der infinitesimal Erzeugenden der unitären Gruppe durch einfaches Multiplizieren mit i mit den symmetrischen Abbildungen identifizieren, welche sich der Eigenschaften 20.2 erfreuen. Für $\mathbb{K} = \mathbb{R}$ gilt dies hingegen nicht (schon aus Dimensionsgründen!!!) Die Möglichkeit eines solchen fundamentalen Zusammenhangs zwischen Observablen und kontinuierlichen Symmetrien (Stichwort: Noether-Theoreme) ist letztlich der Grund für die Bedeutung von \mathbb{C} in der Quantenmechanik.

· Die Klammer (20.34) ist anti-symmetrisch

$$[A_1, A_2] = -[A_2, A_1] \quad (20.39)$$

und erfüllt die Jacobi-Identität

$$[A_1, [A_2, A_3]] + [A_2, [A_3, A_1]] + [A_3, [A_1, A_2]] = 0 \quad (20.40)$$

Eine solche Struktur heisst Lie-Algebra.

Rayleigh-Ritz-Verfahren

Unsere zweite Anwendung der Analysis auf die lineare Algebra liefert ein häufig benutztes Verfahren zur Approximation von Eigenwerten symmetrischer Abbildungen in hoher (sowie mit geeigneten weiteren Annahmen auch unendlicher) Dimension. Wir formulieren sie für $\mathbb{K} = \mathbb{R}$, $V = \mathbb{R}^n$ ($n \geq 1$) mit dem euklidischen Standardprodukt (20.6).

Theorem 20.9. Für eine symmetrische Matrix $A \in \text{Mat}(n, \mathbb{R})$, $A^T = A$ ist

$$\lambda := \min\{\langle x, Ax \rangle \mid \langle x, x \rangle = 1\} \quad (20.41)$$

der kleinste Eigenwert von A , und jedes x_* mit $\langle x_*, x_* \rangle = 1$ und $\langle x_*, Ax_* \rangle = \lambda$ ist ein Eigenvektor zum Eigenwert λ .

Bemerkungen. · Die $(n - 1)$ -dimensionale Sphäre

$$S^{n-1} := \left\{ x \in \mathbb{R}^n \mid \langle x, x \rangle = \sum_{i=1}^n (x^i)^2 = 1 \right\} \quad (20.42)$$

ist beschränkt und abgeschlossen, und daher nach dem Satz von Bolzano-Weierstrass 4.17 folgenkompakt im Sinne von 7.5. Daher nimmt die stetige Funktion $x \mapsto \langle x, Ax \rangle$ auf S^{n-1} ein Minimum an, so dass die Behauptung Sinn macht.

· Der Rest des Beweises ist aber eigentlich eine Ausrede, nochmals einige Begriffe aus dem § 12 zu wiederholen. Die folgenden Betrachtungen hängen zunächst nicht von der euklidischen Metrik auf \mathbb{R}^n ab.

Fakten 20.10. Es sei $M \subset \mathbb{R}^n$ nicht leer und $p \in M$. Für eine offene Teilmenge $U \subset \mathbb{R}^n$ mit $p \in U$ heisst eine Abbildung

$$\varphi : U \cap M \rightarrow \mathbb{R}^k \quad (20.43)$$

eine k -dimensionale (glatte) Karte von M um p falls

(i) φ injektiv und stetig, und $V := \varphi(U \cap M) \subset \mathbb{R}^k$ offen ist sowie
 (ii) die Umkehrabbildung $\psi := \varphi^{-1} : V \rightarrow \mathbb{R}^n$ (unendlich oft⁹) differenzierbar ist, mit $\text{rang}(D\psi(y)) = k \forall y \in V$.

· Eine k -dimensionale (glatte) Untermannigfaltigkeit des \mathbb{R}^n ist eine nicht-leere Teilmenge $M \subset \mathbb{R}^n$ mit der Eigenschaft, dass für jeden Punkt $p \in M$ eine (glatte) k -dimensionale Karte von M um p existiert.

· Die Umkehrabbildung $\psi = \varphi^{-1} : V \rightarrow U \cap M$ heisst auch lokale (reguläre) Parametrisierung von M um p . Sie ist anfangs eingängiger, üblich ist aber tatsächlich die Bezeichnung “Karte” für das Paar $(U \cap M, \varphi)$.

· Beispielsweise ist für $V \subset \mathbb{R}^k$ offen und eine glatte Abbildung $f : V \rightarrow \mathbb{R}^{n-k}$ der Graph $\text{Graph}(f) := \{(x, f(x)) \mid x \in V\}$ eine k -dimensionale Untermannigfaltigkeit des $\mathbb{R}^n = \mathbb{R}^k \times \mathbb{R}^{n-k}$. (Es existiert eine globale Karte mit $U = \mathbb{R}^n$, $\varphi(x, f(x)) = x$.)

⁹Es macht auch Sinn, Untermannigfaltigkeiten zu betrachten, bei denen φ^{-1} nur stetig, oder ein-, zwei- etc. Mal stetig differenzierbar ist. Häufig ist wie hier bei uns “glatt” synonym mit “unendlich oft differenzierbar”.

§ 20. EUKLIDISCHE UND UNITÄRE RÄUME

· Der Inhalt des Satzes 12.3 über implizite Funktionen ist die Aussage, dass für eine glatte Abbildung $F : U \rightarrow \mathbb{R}^{n-k}$ die Niveaumengen $F^{-1}(\text{const.})$ in der Umgebung von Punkten p mit $\text{rang}(DF(p)) = n - k$ lokal Graphen glatter Funktionen und daher k -dimensionale Untermannigfaltigkeiten sind. Tatsächlich lässt sich auch jede k -dimensionale Untermannigfaltigkeit so darstellen.

· Für einen Punkt $p \in M \subset \mathbb{R}^n$ definieren wir den Tangentialraum von M in p als Unterraum von \mathbb{R}^n (wie in (10.27) als in p angeheftet gedacht), indem wir eine lokale Parametrisierung $\psi : V \rightarrow \mathbb{R}^n$ von M um p wählen (insbesondere existiert ein $q \in V$ mit $p = \psi(q)$), und setzen:

$$T_p M := D\psi(q)(\mathbb{R}^k) = \{D\psi(q)(w) \mid w \in \mathbb{R}^k\} \quad (20.44)$$

Anschaulich ist $T_p M$ die Menge der Richtungsvektoren $\dot{\gamma}(0)$ aller glatten Kurven $\gamma : (-\epsilon, \epsilon) \rightarrow \mathbb{R}^n$ mit $\gamma((-\epsilon, \epsilon)) \subset M$ und $\gamma(0) = p$.

· Ist $F : U \rightarrow \mathbb{R}^{n-k}$ mit $U \cap M = U \cap F^{-1}(0) = \{\psi(y) \mid y \in V\}$ wie im Satz über implizite Funktionen, so folgt aus $F(\psi(y)) = 0 \forall y \in V$ durch Ableiten aus der Kettenregel

$$v = D\psi(q)(w) \Rightarrow DF(p)(v) = DF \circ D\psi(q)(w) = 0 \quad (20.45)$$

und aus Dimensionsgründen dann

$$T_p M = \text{Ker}(DF(p)) \subset \mathbb{R}^n \quad (20.46)$$

Insbesondere ist $T_p M$ als Untervektorraum von \mathbb{R}^n nicht von der Wahl einer Parametrisierung (von denen es im Allgemeinen viele gibt) abhängig.

Beispiel: Die $(n-1)$ -dimensionale Sphäre ist eine $(n-1)$ -dimensionale glatte Untermannigfaltigkeit des \mathbb{R}^n : Das Differential der definierenden Gleichung $\langle x, x \rangle - 1 = 0$ am Punkte $x_0 \in S^{n-1}$ ist

$$\mathbb{R}^n \ni v \mapsto D(\langle x, x \rangle - 1)(x_0)(v) = 2\langle x_0, v \rangle \quad (20.47)$$

und hat Rang $n - (n-1) = 1$ (Übungsaufgabe). Der Tangentialraum an die $(n-1)$ -dimensionale Sphäre im Punkte $x_0 \in S^{n-1}$ ist also

$$\begin{aligned} T_{x_0} S^{n-1} &= \text{Ker } D(\langle x, x \rangle - 1)(x_0) = \{v \in \mathbb{R}^n \mid \langle x_0, v \rangle = 0\} \\ &= x_0^\perp \end{aligned} \quad (20.48)$$

(vgl. (20.12)).

Proposition 20.11. *Es sei $M \subset \mathbb{R}^n$ eine k -dimensionale Untermannigfaltigkeit, und $f \in \mathcal{C}^\infty(\mathbb{R}^n, \mathbb{R})$. Angenommen, f hat in $p_* \in M$ ein lokales Minimum auf M , d.h. für eine Umgebung $U \subset \mathbb{R}^n$ von p_* gilt¹⁰*

$$f(p) \geq f(p_*) \quad \forall p \in U \cap M \quad (20.49)$$

Dann gilt:

$$T_{p_*} M \subset \text{Ker}(Df(p_*)) \quad (20.50)$$

¹⁰ p_* ist i.A. kein lokales Minimum von f in \mathbb{R}^n , d.h. es könnte Punkte $q \notin M$ aber in der Nähe von p_* geben, mit $f(q) < f(p_*)$.

In Worten: f ist in den Tangentialrichtungen in erster Ordnung stationär.

Beweis. Für $v \in T_{p_*}M$ sei $\gamma : (-\epsilon, \epsilon) \rightarrow \mathbb{R}^n$ eine glatte Kurve mit $\gamma((-\epsilon, \epsilon)) \subset M$, $\gamma(0) = p_*$, und $\dot{\gamma}(0) = v$. OBdA nehmen wir an, dass $\gamma((-\epsilon, \epsilon)) \subset U \cap M$, mit U wie in (20.49). Dann hat $f \circ \gamma : (-\epsilon, \epsilon) \rightarrow \mathbb{R}$ in 0 ein lokales Minimum und es folgt daher aus 9.6

$$0 = \left. \frac{d}{dt} \right|_{t=0} f(\gamma(t)) = Df(p_*)(\dot{\gamma}(0)) = Df(p_*)(v) \quad (20.51)$$

Dies impliziert (20.50). □

Beweis von 20.9. Wir wissen bereits, dass die auf \mathbb{R}^n glatte Funktion

$$x \mapsto \langle x, Ax \rangle \quad (20.52)$$

auf S^{n-1} ein globales also auch lokales Minimum in $x_* \in S^{n-1}$ besitzt. Das Differential dieser Funktion ist

$$\mathbb{R}^n \ni v \mapsto D\langle x, Ax \rangle(x_*)(v) = 2\langle x_*, Av \rangle \quad (20.53)$$

Mit (20.48) bedeutet dann (20.50), dass

$$v \in x_*^\perp \Rightarrow \langle x_*, Av \rangle = 0 \quad (20.54)$$

Wegen der Symmetrie von A ist also $\langle Ax_*, v \rangle = 0 \forall v \in x_*^\perp$. In der Zerlegung

$$\mathbb{R}^n = \mathbb{R} \cdot x_* \oplus x_*^\perp \quad (20.55)$$

ist aber die Einschränkung von $\langle \cdot, \cdot \rangle$ auf x_*^\perp ebenfalls positiv definit, also insbesondere nicht entartet. Daher ist $Ax_* \in \mathbb{R} \cdot x_*$. Es existiert also ein λ mit $Ax_* = \lambda x_*$. Es folgt der Rest der Behauptungen. □

KAPITEL 8

INTEGRATIONSTHEORIE

Die physikalisch interessantesten Aussagen aus der linearen Algebra, die wir im letzten Kapitel für endliche Dimension bewiesen haben, übertragen sich auf den (physikalisch noch interessanteren) Fall unendlich-dimensionaler Vektorräume, aber *nicht ohne weitere Anstrengungen* und Annahmen. Die Schwierigkeiten entstehen dadurch, dass einerseits die metrische Vollständigkeit normierter Räume (im Sinne von 4.2) in die Beweise einging, insbesondere bei der Eigenwerttheorie und der infinitesimalen Erzeugung kontinuierlicher Gruppen, andererseits Normen auf unendlich-dimensionalen Vektorräumen aber nicht mehr alle äquivalent sind (vgl. § 24 im Kontrast zu 4.18). Es existieren daher auch verschiedene Konvergenzbegriffe und wir müssen uns darüber Rechenschaft ablegen, welcher davon für die jeweilige Anwendung der “richtige” ist.¹¹

Ein verwandtes Problem tritt bei der Lösung von linearen und nicht-linearen Differentialgleichungen auf. Während wir für gewöhnliche Differentialgleichungen im § 14 Existenz und Eindeutigkeit von Lösungen auf der Basis der stetigen Funktionen und der gleichmässigen Konvergenz kontrollieren konnten, so reicht dieser Rahmen für die meisten partiellen Differentialgleichungen nicht mehr aus. Unstetige und unbeschränkte Funktionen treten zumindestens in Zwischenschritten unweigerlich auf.

Um all diese Fragen in Angriff nehmen zu können, ist es jetzt unabdingbar, dass wir uns zunächst mit dem allgemeinen und flexiblen Integralbegriff auseinandersetzen.

§ 21 Integration im \mathbb{R}^n

Zur Erinnerung: Wir hatten im § 13 (s. Def. 13.3) das Integral einer stetigen Funktion $f : I \rightarrow V$ ($I \subset \mathbb{R}$ ein Intervall, V ein vollständiger normierter Vektorraum) von a nach b in I definiert als Grenzwert einer speziellen Folge von Riemannschen Summen:

$$\int_a^b f(s)ds = \lim_{n \rightarrow \infty} \sum_{k=0}^{n-1} \left(\frac{b-a}{n}\right) \cdot f\left(a + \frac{k}{n}(b-a)\right) \quad (21.1)$$

¹¹Es ist intuitiv nicht ganz einfach zu sagen, welche physikalischen Prinzipien bei der Auswahl helfen sollen. Ein wichtiges Rationalitätskriterium ist sicherlich die Möglichkeit der “beliebig genauen Approximation durch endlich viele Schritte”.

Unsere Motivation dazu war die Umkehrung der ein-dimensionalen Ableitung, aufgefasst als lineare Abbildung

$$\mathcal{C}^1(I, V) \rightarrow \mathcal{C}^0(I, V) \tag{21.2}$$

im Sinne des Hauptsatzes der Differential- und Integralrechnung: Für jede Funktion $f \in \mathcal{C}^0(I, V)$ ist

$$F(t) := \int_{t_*}^t f(s) ds \tag{21.3}$$

eine Stammfunktion von f ist, d.h. $\dot{F}(t) = f(t)$.

Die Konzeption des Integrals als Umkehrung der Ableitung lässt sich nun nicht ohne Weiteres auf beliebige Funktionen von $n > 1$ reellen Veränderlichen verallgemeinern. Schon im Falle $n = 2$ hatten wir im § 15 gesehen, dass wir Stammfunktionen von Funktionen $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ nur unter Zuhilfenahme der komplexen Multiplikation im Wertebereich $\mathbb{R}^2 \cong \mathbb{C}$ gewinnen konnten, und auch nur unter der zusätzlichen Annahme, dass die Ableitung bereits holomorph ist, d.h. auch im Definitionsbereich mussten wir $\mathbb{R}^2 \cong \mathbb{C}$ identifizieren. Für $n > 2$ steht solch eine algebraische Struktur einfach nicht zur Verfügung.

Eine flexiblere Interpretation, die sich unmittelbar und viel weitreichender verallgemeinern lässt, ist die des bestimmten Integrals als ‘‘Mittelwert’’ einer Funktion über ihren Definitionsbereich bzw. als ‘‘gewichtete Summe’’ der Funktionswerte. Dass dies tatsächlich die ‘‘richtige’’ physikalische Deutung des Integrals ist, wird in der statistischen Physik und dann insbesondere im Pfadintegralzugang zur Quantenmechanik nachdrücklich klar. Wie wir im § 23 sehen werden, lassen sich aber auch in n Dimensionen (manchmal) Stammfunktionen durch Integration erhalten, und letztlich ist der Zusammenhang zwischen Mittelung und infinitesimalem Zuwachs die eigentliche mathematische Aussage des Hauptsatzes der Differential- und Integralrechnung.

Ein Vorteil der Mittelwertinterpretation der Integrals ist, dass sie im Prinzip auch für beliebige Funktionen auf Mengen eingeführt werden kann, die nicht mit allen Strukturen von \mathbb{R} bzw. \mathbb{R}^n ausgerüstet sind. (Beispielsweise ist schon die Abbildung

$$\begin{aligned} \mathbb{R}^n &= \{x : N \rightarrow \mathbb{R}\} \rightarrow \mathbb{R} \\ x &= (x^1, \dots, x^n)^T \mapsto \sum_{i=1}^n x^i \end{aligned} \tag{21.4}$$

(vgl. S. 130) der Permutationsdarstellung von S_n auf ihren invarianten Unterraum eine interessante Sorte Integral in diesem Sinne.) Als Zugeständnis an das Fehlen einer natürlichen Anordnung für $n > 1$ schreiben wir jetzt auch im ein-dimensionalen Fall¹²

$$\int_{[a,b]} f(x) dx := \int_a^b f(s) ds \tag{21.6}$$

¹²Beachte, dass wir dadurch Information verlieren. Für $a > b$ gilt:

$$\int_a^b f(s) ds = - \int_{[b,a]} f(x) dx !!! \tag{21.5}$$

§ 21. INTEGRATION IM \mathbb{R}^n

Auf der anderen Seite ist die Einführung der Mittelung zunächst am einfachsten, wenn der Wertebereich unserer Funktionen angeordnet ist. Beispielsweise gilt:

Lemma 21.1 (Mittelwertsatz der Integralrechnung). *Es seien $a, b \in \mathbb{R}$ mit $a \leq b$ und $f : [a, b] \rightarrow \mathbb{R}$ eine stetige Funktion¹³. Dann existiert ein $\xi \in [a, b]$ so, dass*

$$\int_{[a,b]} f(x)dx = \int_a^b f(s)ds = f(\xi) \cdot (b - a) = f(\xi) \cdot |b - a| \quad (21.7)$$

Beweis. Für $a = b$ steht mit $\xi = a$ auf beiden Seiten 0. Es sei also im Folgenden $a < b$. Da $[a, b]$ nach dem Satz von Bolzano-Weierstrass 3.11 folgenkompakt ist, existieren gemäss Thm. 7.5 Stellen x_m und x_M in $[a, b]$ so, dass

$$\begin{aligned} f(x_m) &= \inf f([a, b]) = \inf\{f(x) \mid x \in [a, b]\} \\ f(x_M) &= \sup f([a, b]) = \sup\{f(x) \mid x \in [a, b]\} \end{aligned} \quad (21.8)$$

Es gilt dann $f(x) - f(x_m) \geq 0 \forall x$ und daher (Übungsaufgabe!)

$$\int_{[a,b]} f(x)dx - f(x_m) \cdot (b - a) = \int_{[a,b]} (f(x) - f(x_m))dx \geq 0 \quad (21.9)$$

Ebenso gilt

$$\int_{[a,b]} f(x)dx \leq f(x_M) \cdot (b - a) \quad (21.10)$$

Zusammen

$$f(x_m) \leq \frac{1}{b - a} \cdot \int_{[a,b]} f(x)dx \leq f(x_M) \quad (21.11)$$

Mit einer leichten Verallgemeinerung des Zwischenwertsatzes 7.1 existiert dann ein $\xi \in [a, b]$ mit

$$f(\xi) = \frac{1}{b - a} \cdot \int_{[a,b]} f(x)dx \quad (21.12)$$

wie gefordert. □

Um die Sache überschaubar zu halten, geben wir also nun das Ziel dieses § aus als die Definition des Integrals als gewichtete Summe für möglichst viele Funktionen $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Wir benötigen dazu einige topologische Grundbegriffe, bezüglich einer beliebigen Norm $\|\cdot\|$ auf \mathbb{R}^n .

• Eine Teilmenge $U \subset \mathbb{R}^n$ heisst *offen*, falls für jedes $x \in U$ ein $r > 0$ existiert so, dass

$$B_r(x) := \{y \in \mathbb{R}^n \mid \|x - y\| < r\} \subset U \quad (21.13)$$

• Eine Menge $A \subset \mathbb{R}^n$ heisst *abgeschlossen*, falls einer jeden Folge $(x_k) \subset A$, welche als Folge in \mathbb{R}^n konvergiert, Grenzwert in A liegt. Dies ist genau dann der Fall, wenn $\mathbb{R}^n \setminus A$ offen ist.

• Abgeschlossene und beschränkte Teilmengen des \mathbb{R}^n sind folgenkompakt. Wir sagen im Folgenden auch einfach “kompakt”.

¹³Wir lassen als Verallgemeinerung von 1.13 jetzt auch die Notation $[a, a] := \{a\} \subset \mathbb{R}$ zu.

· Der *Abschluss* einer beliebigen Teilmenge $X \subset \mathbb{R}^n$ ist die “kleinste” abgeschlossene Menge \overline{X} , die X noch ganz enthält. Eine nützliche Charakterisierung von \overline{X} ist als Vereinigung von X mit den Grenzwerten aller (in \mathbb{R}) konvergenten Folgen mit Werten in X .

$$\overline{X} = \bigcap_{\substack{A \supset X \\ A \text{ abgeschlossen}}} A = X \cup \left\{ \lim_{n \rightarrow \infty} x_n \mid (x_n)_{n \in \mathbb{N}} \subset X, x_n \text{ ist Cauchy-Folge} \right\} \quad (21.14)$$

· Der *offene Kern* oder das Innere einer beliebigen Teilmenge $X \subset \mathbb{R}^n$ ist die “grösste” offene Menge $\overset{\circ}{X}$, die noch ganz in X enthalten ist.

· Der *Rand* von X ist

$$\partial X = \overline{X} \setminus \overset{\circ}{X} \quad (21.15)$$

· Für eine Funktion

$$f : \mathbb{R}^n \rightarrow \mathbb{R} \quad (21.16)$$

heisst

$$\text{supp}(f) := \overline{\{x \in \mathbb{R}^n \mid f(x) \neq 0\}} \quad (21.17)$$

der *Träger* (Support) von f

· Für eine beliebige Menge $X \subset \mathbb{R}^n$ heisst $\chi_X : \mathbb{R}^n \rightarrow \mathbb{R}$ mit

$$\chi_X(x) := \begin{cases} 1 & x \in X \\ 0 & x \notin X \end{cases} \quad (21.18)$$

die *charakteristische Funktion* von X . Übungsaufgabe: $\text{supp}(\chi_X) = \overline{X}$.

Wir führen nun sukzessive eine Reihe von Unterräumen des Vektorraums aller reellwertigen Funktion auf \mathbb{R}^n ein, und definieren ein zugehöriges Integral mit charakteristischen Eigenschaften. Der Einfachheit halber lassen wir den Wertebereich aus der Notation weg (Bsp. $\mathcal{F}(\mathbb{R}^n) := \mathcal{F}(\mathbb{R}^n, \mathbb{R})$.)

Treppenfunktionen

· Ein (abgeschlossener) *Quader* ist eine Teilmenge $Q \subset \mathbb{R}^n$ der Form

$$Q = [a_1, b_1] \times \cdots \times [a_n, b_n] \quad (21.19)$$

für reelle Zahlen $a_1 \leq b_1, a_2 \leq b_2, \dots, a_n \leq b_n$. Ein Quader heisst *entartet*, falls $a_i = b_i$ für wenigstens ein $i \in \{1, \dots, n\}$, sonst *nicht-entartet*. Wir schreiben für das (euklidische!) *Volumen eines Quaders*

$$\text{vol}(Q) := \prod_{i=1}^n |b_i - a_i| \quad (21.20)$$

Offensichtlich ist $\text{vol}(Q) = 0 \Leftrightarrow Q$ entartet.

· Eine Funktion $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ heisst *Treppenfunktion*, falls eine endliche Anzahl Quader Q_1, \dots, Q_k mit reellen Zahlen c_1, \dots, c_k existiert mit der Eigenschaft, dass $\overset{\circ}{Q}_i \cap \overset{\circ}{Q}_j = \emptyset$ falls $i \neq j$ und

$$\varphi = \sum_{i=1}^k c_i \chi_{Q_i} \quad (21.21)$$

§ 21. INTEGRATION IM \mathbb{R}^n

Die Menge $\mathcal{T}(\mathbb{R}^n)$ der Treppenfunktionen ist abgeschlossen unter punktweise Addition und Skalarmultiplikation, bildet also einen Untervektorraum von $\mathcal{F}(\mathbb{R}^n)$. Wir werden den Beweis dieser Aussage nicht formal aufschreiben, illustrieren aber die notwendigen Grundideen durch die folgenden Beobachtungen.

· Die charakteristischen Funktionen von offenen Quadern sind Treppenfunktionen: Betrachte einen nicht-entarteten Quader $Q = [a_1, b_1] \times \cdots \times [a_n, b_n]$, und setze für $i = 1, \dots, n$

$$\begin{aligned} \chi_i &:= \chi_{\mathbb{R}^{i-1} \times [a_i, b_i] \times \mathbb{R}^{n-i}}, & \overset{\circ}{\chi}_i &:= \chi_{\mathbb{R}^{i-1} \times (a_i, b_i) \times \mathbb{R}^{n-i}}, \\ \varphi_i &:= \chi_{\mathbb{R}^{i-1} \times \{a_i\} \times \mathbb{R}^{n-i}}, & \psi_i &:= \chi_{\mathbb{R}^{i-1} \times \{b_i\} \times \mathbb{R}^{n-i}} \end{aligned} \quad (21.22)$$

Dann gilt

$$\begin{aligned} \overset{\circ}{\chi}_i &= \chi_i - \varphi_i - \psi_i \\ \chi_Q &= \prod_{i=1}^n \chi_i & \chi_{\overset{\circ}{Q}} &= \prod_{i=1}^n \overset{\circ}{\chi}_i \end{aligned} \quad (21.23)$$

Ausmultiplizieren ergibt die explizite Darstellung von $\chi_{\overset{\circ}{Q}}$ als Linearkombination von charakteristischen Funktionen von abgeschlossenen Quadern, welche ausser Q alle entartet sind und daher leeres Innere haben.

· Ebenso sind charakteristische Funktionen von kartesischen Produkten offener, halb-offener, und abgeschlossener Intervalle Treppenfunktionen. Wir wollen solche Mengen *teilweise offene* Quader nennen.

· Für zwei Quader Q_1 und Q_2 lässt sich die Summe wieder in die Form (21.21) zerlegen (Übungsaufgabe).

· Man beachte insbesondere, dass die Darstellung einer Treppenfunktion wie in (21.21) nicht eindeutig ist. Für $n = 1$ gilt beispielsweise

$$\chi_{[a,b]} = \chi_{[a, \frac{a+b}{2}]} + \chi_{[\frac{a+b}{2}, b]} - \chi_{\{\frac{a+b}{2}\}} \quad (21.24)$$

· Aus diesen Betrachtungen folgt etwa auch, dass

$$\mathcal{T}(\mathbb{R}^n) = \text{span}\{\chi_Q \mid Q \subset \mathbb{R}^n \text{ ein Quader}\} \subset \mathcal{F}(\mathbb{R}^n) \quad (21.25)$$

(Charakteristische Funktionen abgeschlossener Quader bilden nur eben keine Basis von $\mathcal{T}(\mathbb{R}^n)$, siehe (21.24).)

· Schliesslich überlegt man sich, dass für eine Treppenfunktion $\varphi \in \mathcal{T}(\mathbb{R}^n)$ die Zahl

$$I_{\mathcal{T}}(\varphi) := \sum_{i=1}^k c_i \text{vol}(Q_i) \quad (21.26)$$

tatsächlich *nicht* von der gewählten Darstellung (21.21) abhängt. (Beweis wird nachgereicht). Wir nennen $I_{\mathcal{T}}(\varphi)$ das *Integral* von φ über \mathbb{R}^n , später auch geschrieben als

$$I_{\mathcal{T}}(\varphi) = \int_{\mathbb{R}^n} \varphi(x) dx \quad (21.27)$$

Lemma 21.2. (i) Für alle $\varphi_1, \varphi_2 \in \mathcal{T}(\mathbb{R}^n)$, $\lambda \in \mathbb{R}$ gilt

$$I_{\mathcal{T}}(\lambda\varphi_1 + \varphi_2) = \lambda I_{\mathcal{T}}(\varphi_1) + I_{\mathcal{T}}(\varphi_2) \quad (21.28)$$

(ii) Für alle $\varphi \in \mathcal{T}(\mathbb{R}^n)$ mit $\varphi(x) \geq 0 \forall x \in \mathbb{R}^n$ ist $I_{\mathcal{T}}(\varphi) \geq 0$. (Monotonie)

(iii) Für alle $\varphi \in \mathcal{T}(\mathbb{R}^n)$ ist auch $|\varphi| \in \mathcal{T}(\mathbb{R}^n)$ und $|I_{\mathcal{T}}(\varphi)| \leq I_{\mathcal{T}}(|\varphi|)$.¹⁴

Beweis. wird nachgeliefert □

Stetige Funktionen mit kompaktem Träger

· Es sei

$$\mathcal{C}_c(\mathbb{R}^n) := \{f : \mathbb{R}^n \rightarrow \mathbb{R} \mid \text{supp}(f) \text{ kompakt}\} \quad (21.29)$$

Da $\text{supp}(f)$ per Definition abgeschlossen ist, gilt offensichtlich: $f \in \mathcal{C}_c(\mathbb{R}^n) \Leftrightarrow f$ ist stetig und $f(x) = 0$ für $\|x\| > R$ für ein $R > 0 \Leftrightarrow$ Es existiert ein Quader $Q \subset \mathbb{R}^n$ mit $\text{supp}(f) \subset Q$.

· Zur Übertragung des Integrals von $\mathcal{T}(\mathbb{R}^n)$ auf $\mathcal{C}_c(\mathbb{R}^n)$ benutzen wir, dass stetige Funktionen auf folgenkompakten Mengen gleichmässig stetig sind (s. Thm. 8.4). Es folgt daraus, dass sich stetige Funktionen mit kompaktem Träger gleichmässig durch Treppenfunktionen approximieren lassen. Genauer gilt:

Lemma 21.3. Es sei $f \in \mathcal{C}_c(\mathbb{R}^n)$ und Q ein Quader mit $Q \supset \text{supp}(f)$. Dann existiert für jedes $\epsilon > 0$ ein $\varphi \in \mathcal{T}(\mathbb{R}^n)$ mit $\text{supp}(\varphi) \subset Q$ und

$$\|f - \varphi\|_{\infty} \left(:= \sup\{|f(x) - \varphi(x)| \mid x \in \mathbb{R}^n\} \right) < \epsilon \quad (21.30)$$

Beweis. Da Q kompakt ist und f stetig, existiert ein $\delta > 0$ so, dass

$$|f(x) - f(y)| < \epsilon \quad \text{für alle } x, y \in Q \text{ mit } \|x - y\| < \delta \quad (21.31)$$

Zerlege Q (z.B. rastermässig) in disjunkte, teilweise offene Quader Q'_1, \dots, Q'_k so, dass für alle i

$$\text{diam}(Q'_i) \left(:= \sup\{\|x - y\| \mid x, y \in Q'_i\} \right) < \delta \quad (21.32)$$

Wähle Stützstellen $x_i \in Q'_i$ für $i = 1, \dots, k$ und setze

$$\varphi = \sum_{i=1}^k f(x_i) \chi_{Q'_i} \quad (21.33)$$

Dann gilt $\text{supp}(\varphi) \subset Q$ und da die Q'_i eine disjunkte Zerlegung von Q bilden, existiert für jedes $x \in Q$ genau ein $i_x \in \{1, \dots, k\}$ mit $x \in Q'_{i_x}$. Es folgt für alle $x \in Q$

$$|f(x) - \varphi(x)| = |f(x) - f(x_{i_x})| < \epsilon \quad (21.34)$$

wegen (21.31) und da $\|x - x_{i_x}\| \leq \text{diam}(Q'_{i_x}) < \delta$. Da f und φ ausserhalb von Q verschwinden, folgt (21.30). □

Damit sind wir bereit für die ‘Verallgemeinerung’¹⁵ von 13.3.

¹⁴ $|\varphi|(x) := |\varphi(x)|$

¹⁵in Anführungszeichen, wegen der Fussnote auf S. 146

§ 21. INTEGRATION IM \mathbb{R}^n

Proposition/Definition 21.4. *Es sei $f \in \mathcal{C}_c(\mathbb{R}^n)$. Wähle einen nicht-entarteten Quader $Q \supset \text{supp}(f)$, und eine Folge $(\varphi_k)_{k \in \mathbb{N}} \subset \mathcal{T}(\mathbb{R}^n)$ von Treppenfunktionen mit $\text{supp}(\varphi_k) \subset Q$ und $\lim_{k \rightarrow \infty} \|f - \varphi_k\|_\infty = 0$. Dann ist $(I_{\mathcal{T}}(\varphi_k))_{k \in \mathbb{N}}$ eine Cauchy-Folge (von reellen Zahlen) und das Integral*

$$I_{\mathcal{C}_c}(f) := \lim_{k \rightarrow \infty} I_{\mathcal{T}}(\varphi_k) \quad (21.35)$$

hängt weder von dem Quader Q noch von der Folge (φ_k) ab.

Beweis. Die Existenz einer Folge mit den geforderten Eigenschaften folgt aus dem Lemma 21.3. Für jedes $\epsilon > 0$ existiert dann ein K so, dass für alle $k, l \geq K$ und für alle $x \in Q$:

$$\begin{aligned} |\varphi_k(x) - \varphi_l(x)| &\leq |\varphi_k(x) - f(x)| + |f(x) - \varphi_l(x)| \\ &\leq \|f - \varphi_k\|_\infty + \|f - \varphi_l\|_\infty < \frac{\epsilon}{\text{vol}(Q)} \end{aligned} \quad (21.36)$$

Für $x \notin Q$ ist $\varphi_k(x) - \varphi_l(x) = 0$. Mit Hilfe von Lemma 21.2 und $I_{\mathcal{T}}(\chi_Q) = \text{vol}(Q)$ folgt für $k, l \geq K$

$$\begin{aligned} |I_{\mathcal{T}}(\varphi_k) - I_{\mathcal{T}}(\varphi_l)| &= |I_{\mathcal{T}}(\varphi_k - \varphi_l)| \leq I_{\mathcal{T}}(|\varphi_k - \varphi_l|) \\ &\leq I_{\mathcal{T}}(\|\varphi_k - \varphi_l\|_\infty \cdot \chi_Q) < \frac{\epsilon}{\text{vol}(Q)} \cdot \text{vol}(Q) = \epsilon \end{aligned} \quad (21.37)$$

$(I_{\mathcal{T}}(\varphi_k))$ ist also eine Cauchy-Folge, daher konvergent. Sind \tilde{Q} und $(\tilde{\varphi}_k)$ ein anderer Quader und Folge mit $\text{supp}(f), \text{supp}(\tilde{\varphi}_k) \subset \tilde{Q}$ und $\|f - \tilde{\varphi}_k\|_\infty \rightarrow 0$, so sei $\tilde{\tilde{Q}}$ ein Quader mit $Q \cap \tilde{Q} \subset \tilde{\tilde{Q}}$. Dann folgt aus $\|\varphi_k - \tilde{\varphi}_k\|_\infty \rightarrow 0$ mit einer Abschätzung auf $\tilde{\tilde{Q}}$, dass $\lim I_{\mathcal{T}}(\tilde{\varphi}_k) = \lim I_{\mathcal{T}}(\varphi_k)$. \square

· Man überzeuge sich davon, dass im Falle $n = 1$ mit (21.6) $I_{\mathcal{C}_c}(f)$ mit der ursprünglichen Definition (21.1) übereinstimmt.

Proposition 21.5 (Eigenschaften des Integrals). $I_{\mathcal{C}_c} : \mathcal{C}_c(\mathbb{R}^n) \rightarrow \mathbb{R}$ ist linear, monoton, und für $f \in \mathcal{C}_c(\mathbb{R}^n)$ mit $\text{supp}(f) \subset Q$ für einen Quader Q gilt

$$|I_{\mathcal{C}_c}(f)| \leq \|f\|_\infty \cdot \text{vol}(Q) \quad (21.38)$$

Beweis. Wende die Rechenregeln für Folgen auf die entsprechenden Eigenschaften von $I_{\mathcal{T}}$ an. \square

· Wir geben nun das erste Resultat zur prinzipiellen Berechnung des mehrdimensionalen Integrals, nämlich durch sukzessive Rückführung auf das ein-dimensionale Integral, welches sich wegen (21.6) im Prinzip mittels Stammfunktionen lösen lässt (vgl. Thm. 13.2).

Proposition 21.6. *Sei $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ stetig mit kompaktem Träger.*

(i) *Für jedes $y \in \mathbb{R}^m$ ist die Funktion*

$$\mathbb{R}^n \ni x \mapsto f^{(y)}(x) := f(x, y) \in \mathbb{R} \quad (21.39)$$

stetig mit kompaktem Träger.

(ii) Die Funktion

$$\mathbb{R}^m \in y \mapsto g(y) := I_{C_c}(f^{(y)}) \quad (21.40)$$

ist stetig mit kompaktem Träger.

(iii) Es gilt: $I_{C_c}(g) = I_{C_c}(f)$

Bemerkungen. Beachte, dass wir in der Notation von I_{C_c} nicht zwischen den verschiedenen Dimensionen unterschieden haben, da wir diese aus dem jeweiligen Argument ablesen können. Mit der (klobigeren) üblichen Notation lautet die Aussage:

$$\int_{\mathbb{R}^m} \left(\int_{\mathbb{R}^n} f(x, y) dx \right) dy = \int_{\mathbb{R}^{n+m}} f(x, y) d(x, y) \quad (21.41)$$

(und aus Symmetriegründen natürlich auch $= \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^m} f(x, y) dy \right) dx$.)

Beweis. · Zunächst halten wir fest, dass die analogen Aussagen für das Integral $I_{\mathcal{T}}$ auf Treppenfunktionen gilt. Dazu reicht es wegen der Linearität der beteiligten Abbildungen ($I_{\mathcal{T}}, \varphi \mapsto \varphi^{(y)}, \varphi \mapsto I_{\mathcal{T}}(\varphi^{(\cdot)}) \in \mathcal{T}(\mathbb{R}^n)$) aus, sie für das Erzeugendensystem (21.25) der charakteristischen Funktionen abgeschlossener Quader zu verifizieren. Offenbar ist aber jeder Quader in \mathbb{R}^{n+m} ein Produkt $Q \times P$ von Quadern in \mathbb{R}^n und \mathbb{R}^m und es gilt

$$I_{\mathcal{T}}(\chi_{Q \times P}) = \text{vol}(Q \times P) = \text{vol}(Q) \cdot \text{vol}(P) = I_{\mathcal{T}}(\underbrace{I_{\mathcal{T}}(\chi_{Q \times P}^{(\cdot)})}_{=\text{vol}(Q) \cdot \chi_P}) \quad (21.42)$$

Für stetige Funktionen mit kompaktem Träger sind die Aussage (i) und die Beschränktheit des Träger von g offensichtlich. Die Stetigkeit von g ist ein Spezialfall der “stetigen/differenzierbaren Abhängigkeit von Parametern”, deren Beweis für $n = 1$ aus 13.7 sich leicht übertragen lässt.

Zum Nachweis von (iii) sei $Q \times P$ ein Quader in $\mathbb{R}^n \times \mathbb{R}^m$ mit $\text{supp}(f) \subset Q \times P$ und für $\epsilon > 0$ sei $\varphi \in \mathcal{T}(\mathbb{R}^n \times \mathbb{R}^m)$ mit $\text{supp}(\varphi) \subset Q \times P$ und $\|f - \varphi\|_{\infty} < \epsilon$. Dann gilt¹⁶

$$|I_{C_c}(f) - I_{\mathcal{T}}(\varphi)| < \epsilon \cdot \text{vol}(Q \times P) \quad (21.43)$$

Ausserdem gilt natürlich $\forall y \in \mathbb{R}^m \ \|f^{(y)} - \varphi^{(y)}\|_{\infty} \leq \|f - \varphi\|_{\infty} < \epsilon$ und wegen $\text{supp}(f^{(y)}), \text{supp}(\varphi^{(y)}) \subset Q$ mit $g(y) := I_{C_c}(f^{(y)})$, $\psi(y) := I_{\mathcal{T}}(\varphi^{(y)})$ daher

$$\|g - \psi\|_{\infty} < \epsilon \cdot \text{vol}(Q) \quad (21.44)$$

Es folgt wegen $\text{supp}(g), \text{supp}(\psi) \subset P$

$$|I_{C_c}(g) - I_{\mathcal{T}}(\psi)| < \epsilon \cdot \text{vol}(Q) \cdot \text{vol}(P) \quad (21.45)$$

und wegen $I_{\mathcal{T}}(\psi) = I_{\mathcal{T}}(\varphi)$ zusammen

$$|I_{C_c}(g) - I_{C_c}(f)| \leq |I_{C_c}(g) - I_{\mathcal{T}}(\psi)| + |I_{\mathcal{T}}(\varphi) - I_{C_c}(f)| < 2\epsilon \cdot \text{vol}(Q) \cdot \text{vol}(P) \quad (21.46)$$

Mit $\epsilon \rightarrow 0$ folgt die Behauptung. □

¹⁶Betrachte $\lim |I_{\mathcal{T}}(\varphi_k) - I_{\mathcal{T}}(\varphi)|$ für $\|f - \varphi_k\|_{\infty} \rightarrow 0$.

§ 21. INTEGRATION IM \mathbb{R}^n

Damit haben wir also das (“euklidische Standard-”) Integral für Treppenfunktionen sowie für stetige Funktionen mit kompaktem Träger definiert und unter Kontrolle. Die intuitive Begründung der weiteren Verallgemeinerung (auf “möglichst alle”¹⁷ Funktionen) ist dann ähnlich wie die der Vervollständigung der rationalen zu den reellen Zahlen: Lösungen gewisser Probleme lassen sich durch Folgen in \mathcal{C}_c (numerisch häufig auch \mathcal{T}) “beliebig annähern”, aber nie exakt finden. Dabei können sowohl nicht-kompakte Träger als auch Unstetigkeiten auftreten.¹⁸ Zur Sicherung der Konvergenz erweitert man wieder den Funktionenraum.

Als Beispiel ist es insbesondere wünschenswert, stetige Funktionen $f \in \mathcal{C}(\mathbb{R}^n)$ über kompakte (oder allgemeinere) Menge X zu integrieren, die nicht den gesamten Träger $\text{supp}(f)$ umfassen. Solche Integrale sollten sich wiederum schreiben lassen als

$$\int_X f(x) dx \stackrel{!}{=} \int_{\mathbb{R}^n} \chi_X(x) f(x) dx \quad (21.47)$$

und $\chi_X \cdot f$ ist im Allgemeinen weder stetig noch eine Treppenfunktion.

Im Prinzip sind mehrere Vervollständigungen denkbar, in Abhängigkeit von dem benutzten Kriterium für die Güte einer Näherung. Tatsächlich sind aber die Ergebnisse meistens vergleichbar. Eine Heidelberger Tradition, der wir hier folgen wollen,¹⁹ ist die “punktweise Anwendung des Prinzips der monotonen Konvergenz” Thm. 3.9:

Eine monotone Folge reeller Zahlen ist entweder unbeschränkt oder konvergent.

Dazu untersuchen wir zunächst die von der Anordnung von \mathbb{R} auf $\mathcal{F}(\mathbb{R}^n)$ induzierten Strukturen, und erweitern sie leicht.

Verabredung: · Für $f, g \in \mathcal{F}(\mathbb{R}^n)$ schreiben wir $f \geq g$, falls $f(x) \geq g(x) \forall x \in \mathbb{R}^n$. (Gibt es sowohl x mit $f(x) > g(x)$ als auch solche mit $f(x) < g(x)$, so sagen wir gar nix.)

· Eine Folge $(f_k) \subset \mathcal{F}(\mathbb{R}^n)$ heisst monoton wachsend (fallend), falls $f_{k+1} \geq f_k \forall k$ ($f_{k+1} \leq f_k \forall k$).

Proposition 21.7. *Sei (f_k) eine monotone Folge stetiger Funktionen mit kompaktem Träger. Angenommen, die punktweise Grenzfunktion $f(x) := \lim_{k \rightarrow \infty} f_k(x)$ ist ebenfalls stetig mit kompaktem Träger. Dann konvergiert $I_{\mathcal{C}_c}(f_k)$ monoton gegen $I_{\mathcal{C}_c}(f)$.*

Wir schreiben jetzt wieder $f_k \uparrow f$ (bzw. $f_k \downarrow f$) für monotone Konvergenz (von Funktionen etc.). Die in der Prop. behauptete Eigenschaft

$$I_{\mathcal{C}}(f_k) \uparrow I_{\mathcal{C}_c}(f) \text{ für } (f_k) \subset \mathcal{C}_c(\mathbb{R}^n), f_k \uparrow f \text{ und } f \in \mathcal{C}_c(\mathbb{R}^n), \quad (21.48)$$

auch geschrieben als $\sup(I_{\mathcal{C}_c}(f_k)) = I_{\mathcal{C}_c}(\sup f_k)$, bezeichnen wir als *Ordnungsstetigkeit* von $I_{\mathcal{C}_c}$.

¹⁷Tatsächlich lässt sich ein Integral mit vernünftigen Eigenschaften nicht auf ganz $\mathcal{F}(\mathbb{R}^n)$ definieren. Der Unterraum, den wir beschreiben, ist aber in einem gewissen Sinne “so gross wie möglich”.

¹⁸Tatsächlich sind die physikalisch relevanten Lösungen häufig wieder stetig (sogar glatt). Es verhält sich dann ähnlich wie beim Übergang von \mathbb{R} auf \mathbb{C} : Komplexe Zahlen sind aus formalen Gründen unabdingbar, physikalische Messergebnisse stets reell.

¹⁹Ich lehne mich an das Weissauersche Skript an, frühere Versionen stammen von Prof. Freitag.

Beweis von 21.7. Wegen $f \in \mathcal{C}_c(\mathbb{R}^n)$ ist auch die Folge $(f_k - f)_{k \in \mathbb{N}}$, welche monoton gegen Null konvergiert, in $\mathcal{C}_c(\mathbb{R}^n)$. Nach Übergang auf $\pm(f_k - f)$ können wir also oBdA annehmen, dass $f_k \downarrow 0$. Die Monotonie von $(I_{\mathcal{C}_c}(f_k))$ folgt dann wegen 21.5 aus der Monotonie der Funktionenfolge. Ausserdem existiert ein Quader Q so, dass $\text{supp}(f_k) \subset \text{supp}(f_1) \subset Q$. Zum Nachweis, dass $I_{\mathcal{C}_c}(f_k) \rightarrow 0$, behaupten wir zunächst: Beh.: Aus der Tatsache, dass $f \in \mathcal{C}_c(\mathbb{R}^n)$ und f_k monoton, folgt bereits, dass (f_k) *gleichmässig* konvergiert (Satz von Dini).

Bew.: Wäre die Konvergenz nicht gleichmässig, so gäbe es ein $\epsilon_* > 0$ mit einer Folge von Punkten $x_k \in Q$ so, dass $|f_k(x_k) - 0| = f_k(x_k) > \epsilon_*$ für alle k . Wegen der Folgenkompaktheit von Q können wir ggbfs. nach Übergang zu einer Teilfolge annehmen, dass $x_k \rightarrow x_* \in Q$ für $k \rightarrow \infty$. Sei dann K so gross, dass $f_K(x_*) < \frac{\epsilon_*}{2}$ (Konvergenz von $(f_k(x_*))$), und $\delta > 0$ so, dass $|f_K(x) - f_K(x_*)| < \frac{\epsilon_*}{2}$ für $\|x - x_*\| < \delta$ (Stetigkeit von f_K). Für $k \geq K$ gross genug ist dann $\|x_k - x_*\| < \delta$ und es folgt

$$f_k(x_k) \leq f_K(x_k) \leq f_K(x_*) + \frac{\epsilon_*}{2} < \frac{\epsilon_*}{2} + \frac{\epsilon_*}{2} = \epsilon_* \quad (21.49)$$

\uparrow
 Monotonie!

im Widerspruch zu $f_k(x_k) > \epsilon_*$.

· Mit der gleichmässigen Konvergenz von (f_k) folgt dann aber $I_{\mathcal{C}_c}(f_k) \rightarrow 0$ einfach aus der Abschätzung

$$|I(f_k)| < \|f_k\|_\infty \cdot \text{vol}(Q) \quad (21.50)$$

□

Monotone Hüllen

Wir vervollständigen nun zunächst unser Integral $I_{\mathcal{C}_c}$ unter *punktweise* monotoner Konvergenz. Dazu benutzen wir ausser der partiellen Ordnung auf $\mathcal{C}_c(\mathbb{R}^n)$ und der Prop. 21.7 noch die folgenden, ebenfalls von der Anordnung in \mathbb{R} mit der Stetigkeit des Absolutbetrages induzierten Operationen.

Fakten/Definition 21.8. $\forall f, g \in \mathcal{C}_c(\mathbb{R}^n), \lambda \in \mathbb{R}$ sind auch $\lambda f + g, |f|, |g|$ (punktweise definiert) und daher auch

$$\begin{aligned} \max(f, g) &= \frac{f + g + |f - g|}{2} \\ \min(f, g) &= \frac{f + g - |f - g|}{2} \end{aligned} \quad (21.51)$$

stetig mit kompaktem Träger. Man nennt einen Raum von reellwertigen Funktionen, welcher auf diese Weise unter Addition, Skalarmultiplikation und \min & \max abgeschlossen ist, einen (*Vektor-Verband*).²⁰ Wir schreiben ab jetzt \mathcal{V} für den Verband $\mathcal{C}_c(\mathbb{R}^n)$.

· Eine reellwertige Funktion (man sagt auch Funktional) auf einem Verband, welche (wie $I_{\mathcal{C}_c}$) linear, monoton, und ordnungstetig ist, heisst *Daniell-Integral*.

²⁰Eine andere Bezeichnung für Vektorverbände ist "Riesz-Raum". Abstrakte Verbände spielen in einem axiomatischen Zugang zum quantenmechanischen Messprozess eine fundamentale Rolle. Die Details wären in einem (Pro-)Seminar auszuarbeiten.

§ 21. INTEGRATION IM \mathbb{R}^n

Im nächsten Schritt ist es zweckmässig, den Wertebereich unserer Funktionen so zu erweitern, dass Monotonie tatsächlich stets Konvergenz impliziert. Um für eine unbeschränkte monoton wachsende (fallende) Folge (x_k)

$$\lim_{k \rightarrow \infty} x_k = \sup(x_k) = \infty \quad \left(\lim_{k \rightarrow \infty} x_k = \inf(x_k) = -\infty \right) \quad (21.52)$$

schreiben zu können, erweitern wir (zunächst als Mengen)

$$\mathbb{R}^+ := \mathbb{R} \cup \{\infty\} \quad (\mathbb{R}^- := \mathbb{R} \cup \{-\infty\}) \quad (21.53)$$

(Achtung: Wir erweitern zunächst noch nicht gleichzeitig mit ∞ und $-\infty$.) Ordnung und Addition setzen sich sinnvoll auf \mathbb{R}^\pm fort, die Multiplikation aber *nur mit nicht-negativen reellen Skalaren*, d.h. für $x \in \mathbb{R}^+$, $\lambda \in \mathbb{R}$ mit $\lambda \geq 0$ ist $\lambda \cdot x \in \mathbb{R}^+$ wohldefiniert. Additive Inverse existieren in \mathbb{R}^\pm aber im Allgemeinen nicht mehr. Man sagt: \mathbb{R}^\pm ist ein *konvexer Kegel* über dem angeordneten Körper \mathbb{R} . max & min machen ebenfalls Sinn, der Zusammenhang (21.51) zum Absolutbetrag setzt sich aber nicht auf \mathbb{R}^+ fort.

Fakten/Definition 21.9. Unter Benutzung dieses erweiterten Konvergenzbegriffs führen wir die *monotonen Hüllen* des Verbandes \mathcal{V} ein:

$$\begin{aligned} \mathcal{V}^+ &:= \{f : \mathbb{R}^n \rightarrow \mathbb{R}^+ := \mathbb{R} \cup \{\infty\} \mid \exists (f_k) \subset \mathcal{V} \mid f_k \uparrow f\} \\ \mathcal{V}^- &:= \{f : \mathbb{R}^n \rightarrow \mathbb{R}^- := \mathbb{R} \cup \{-\infty\} \mid \exists (f_k) \subset \mathcal{V} \mid f_k \downarrow f\} \end{aligned} \quad (21.54)$$

Dann gilt: \mathcal{V}^\pm sind unter Addition und min & max abgeschlossen, sowie unter Multiplikation mit *nicht-negativen* reellen Skalaren. D.h. für $\lambda \geq 0$ und $f \in \mathcal{V}^+$ ist auch $\lambda \cdot f \in \mathcal{V}^+$. Dagegen liegen $-f$ sowie $|f|$ i.A. nicht in \mathcal{V}^+ (siehe Übungen für Beispiele). Wir wollen einen solchen Funktionenraum einen *Kegelverband* nennen.²¹

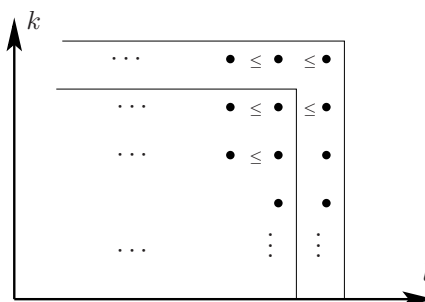
Beh.: Die monotone Hülle \mathcal{V}^+ ist monoton abgeschlossen, d.h.: Ist $(f_k) \subset \mathcal{V}^+$ eine monoton wachsende Folge von Funktionen in \mathcal{V}^+ , so ist $f := \lim_{k \rightarrow \infty} f_k \in \mathcal{V}^+$ (und sinngemäss für \mathcal{V}^-).

Beweis der Abgeschlossenheit von \mathcal{V}^+ . Definitionsgemäss existiert für jedes k eine monoton wachsende Folge $(f_{kl})_{l \in \mathbb{N}} \subset \mathcal{V}$ mit $\sup_l (f_{kl}) = f_k$, d.h. $\forall k$ gilt $f_{k,l+1} \geq f_{kl}$ und $f_{kl} \uparrow f_k$ für $l \rightarrow \infty$. Setze für $l \in \mathbb{N}$

$$g_l := \max\{f_{kl} \mid k \leq l\} \quad (21.55)$$

Dann gilt $g_l \in \mathcal{V}$ und

$$\begin{aligned} g_{l+1} &= \max\{f_{k,l+1} \mid k \leq l+1\} \\ &\geq \max\{f_{kl} \mid k \leq l+1\} \\ &\geq \max\{f_{kl} \mid k \leq l\} \\ &= g_l \end{aligned} \quad (21.56)$$



²¹Wir könnten auch "Riesz-Kegel" sagen.

Ausserdem ist

$$\begin{aligned} \sup(g_l) &= \sup\{f_{kl} \mid k \leq l, l \in \mathbb{N}\} \\ &= \sup\{f_{kl} \mid k, l \in \mathbb{N}\} = \sup\{f_k \mid k \in \mathbb{N}\} = f \end{aligned} \quad (21.57)$$

d.h. $g_l \uparrow f$ und damit ist $f \in \mathcal{V}^+$. □

Beispiel 21.10. Gemäss Aufgabe 6.3 aus der Höheren Mathematik II ist für eine nicht-leere Menge $X \subset \mathbb{R}^n$ die Funktion $d(\cdot, X) : \mathbb{R}^n \rightarrow \mathbb{R}$,

$$d(x, X) := \inf\{\|x - y\| \mid y \in X\} \quad (21.58)$$

stetig, und es gilt

$$d(x, X) = 0 \Leftrightarrow x \in \overline{X} \quad (21.59)$$

Für beschränktes X ist dann $f_k \in \mathcal{F}(\mathbb{R}^n)$ mit

$$f_k(x) := (\max(1 - d(x, X), 0))^k \quad (21.60)$$

stetig mit kompaktem Träger. Die Folge (f_k) fällt monoton und konvergiert für $k \rightarrow \infty$ punktweise gegen die charakteristische Funktion $\chi_{\overline{X}}$, welche i.A. nicht stetig ist. Es gilt also $\chi_{\overline{X}} \in \mathcal{V}^-$ für \overline{X} kompakt.

Übungsaufgabe: Man zeige, dass charakteristische Funktionen von offenen (beschränkten und unbeschränkten) Mengen in \mathcal{V}^+ , aber nicht in \mathcal{V}^- liegen.

Proposition 21.11. Sei $J : \mathcal{V} \rightarrow \mathbb{R}$ ein Daniell-Integral auf dem Verband \mathcal{V} . Dann existiert eine eindeutige Fortsetzung zu einem ordnungsstetigen Funktional $J^+ : \mathcal{V}^+ \rightarrow \mathbb{R}^+$ auf dem Kegelveband \mathcal{V}^+ , d.h.

(iii) $J^+(f) = J(f) \forall f \in \mathcal{V} \subset \mathcal{V}^+$ und $J^+(f_k) \uparrow J^+(f)$ falls $(f_k)_{k \in \mathbb{N}} \subset \mathcal{V}^+$ mit $f_k \uparrow f \in \mathcal{V}^+$.

J^+ ist

(i) *kegellinear*, d.h. $J^+(\lambda f + g) = \lambda J^+(f) + J^+(g) \forall f, g \in \mathcal{V}^+, \lambda \in \mathbb{R}$ mit $\lambda \geq 0$.

(ii) *monoton*, d.h. $f \geq g \Rightarrow J^+(f) \geq J^+(g)$.

Bemerkungen. Beachte, dass bei Fehlen von additiven Inversen die Monotonie in der Form (ii) nicht aus der schwächeren Aussage $f \geq 0 \Rightarrow J^+(f) \geq 0$ folgt.

· Eine analoge Aussage gilt natürlich für \mathcal{V}^- .

Lemma 21.12. Sei $(f_k)_{k \in \mathbb{N}} \subset \mathcal{V}$ eine monoton wachsende Folge, und $g \in \mathcal{V}$ mit $g \leq \sup f_k$. Dann gilt $J(g) \leq \sup J(f_k)$.

Beweis. Betrachte $h_k := \min(g, f_k) \in \mathcal{V}$. Dann gilt einerseits $h_k \uparrow g \in \mathcal{V}$ und daher wegen der Ordnungsstetigkeit von J , dass $J(h_k) \uparrow J(g)$. Andererseits ist $h_k \leq f_k$ und daher wegen der Monotonie von J $J(h_k) \leq J(f_k) \leq \sup_k J(f_k)$. Zusammen folgt $J(g) \leq \sup_k J(f_k)$. □

Beweis von 21.11. Wir halten zunächst fest, dass falls so ein Funktional J^+ mit den behaupteten Eigenschaften existiert, es bereits eindeutig festliegt. Ist nämlich $f \in \mathcal{V}^+$, so existiert eine Folge $(f_k) \subset \mathcal{V}$ mit $f_k \uparrow f$ und da wegen der Monotonie

§ 21. INTEGRATION IM \mathbb{R}^n

von J die Folge $(J(f_k))_{k \in \mathbb{N}}$ monoton wächst, existiert $\lim_{k \rightarrow \infty} J(f_k) \in \mathbb{R}^+$. Wegen $J(f_k) = J^+(f_k)$ und der Stetigkeit von J^+ muss dann

$$\lim_{k \rightarrow \infty} J(f_k) = \lim_{k \rightarrow \infty} J^+(f_k) = J^+(f) \quad (21.61)$$

sein. Um andererseits J^+ durch die Formel (21.61) definieren zu können, müssen wir nachweisen, dass der Wert nicht von der gewählten Folge abhängt. Ist nun (\tilde{f}_k) eine weitere Folge mit $\tilde{f}_k \uparrow f$ so gilt $\forall k$

$$\tilde{f}_k \leq f = \sup_l (f_l) \stackrel{21.12}{\implies} J(\tilde{f}_k) \leq \sup_l J(f_l) \quad (21.62)$$

und daher $\sup_k J(\tilde{f}_k) \leq \sup_l J(f_l)$. Ebenso gilt $\sup_l J(f_l) \leq \sup_k J(\tilde{f}_k)$, und daher $\sup_l J(f_l) = \sup_k J(\tilde{f}_k)$. Es macht also Sinn, für die Definition (21.61) die geforderten Eigenschaften nachzuweisen.

(i) Kegellinearität folgt aus der Linearität von J :

$$J^+(\lambda f + g) = \sup J(\lambda f_k + g_k) = \lambda \sup J(f_k) + \sup J(g_k) = \lambda J^+(f) + J^+(g) \quad (21.63)$$

(ii) Aus $g \leq f$ folgt für $f_k \uparrow f$ und $g_l \uparrow g$, dass $g_l \leq \sup f_k \forall l$ und daher wieder mit Lemma 21.12

$$J(g_l) \leq \sup J(f_k) \implies J^+(g) \leq J^+(f) \quad (21.64)$$

Für (den Rest von) (iii) seien $(f_{kl})_{l \in \mathbb{N}} \subset \mathcal{V}$ mit $f_{kl} \uparrow f_k \in \mathcal{V}^+$ und $f_k \uparrow f \in \mathcal{V}^+$. Dann erfüllt $g_l := \max\{f_{kl} \mid k \leq l\}$ (vgl. Beweis von 21.9) $g_l \uparrow f$ und daher gilt

$$\sup J^+(f_k) = \sup J(f_{kl}) = \sup J(g_l) = J^+(f) \quad (21.65)$$

□

Die Räume \mathcal{V}^\pm enthalten schon deutlich mehr Funktionen als $\mathcal{C}_c(\mathbb{R}^n)$, und die Funktionale J^\pm erfüllen schon einen Teil unserer Wunschvorstellungen.

Beispiel 21.13. Sei $Q \subset \mathbb{R}^n$ ein abgeschlossener Quader. Dann ist $\chi_Q \in \mathcal{V}^-$ (s. 21.10) und es gilt $J^-(\chi_Q) = I_{\mathcal{T}}(\chi_Q) = \text{vol}(Q)$.

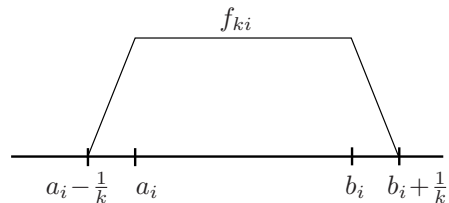
Bew.: Mit $Q = [a_1, b_1] \times \dots \times [a_n, b_n]$ sei für $k = 1, 2, \dots$ und $i = 1, \dots, n$, $x \in \mathbb{R}$:

$$f_{ik}(x) := \begin{cases} 0 & x < a_i - \frac{1}{k} \\ k(x - a_i + \frac{1}{k}) & x \in [a_i - \frac{1}{k}, a_i] \\ 1 & x \in [a_i, b_i] \\ 1 - k(x - b_i) & x \in [b_i, b_i + \frac{1}{k}] \\ 0 & x > b_i + \frac{1}{k} \end{cases}$$

(21.66)

Dann gilt $f_{ki} \in \mathcal{C}_c(\mathbb{R})$ und $f_{ki} \downarrow \chi_{[a_i, b_i]}$ für $k \rightarrow \infty$. Daher erfüllen die Funktionen f_k , für $x = (x^1, \dots, x^n)^T \in \mathbb{R}^n$ definiert durch,

$$f_k(x) := \prod_{i=1}^n f_{ki}(x^i) \quad (21.67)$$



$f_k \in \mathcal{C}_c(\mathbb{R}^n)$ und $f_k \downarrow \chi_Q$. Mit Prop. 21.6 gilt

$$\begin{aligned} I_{\mathcal{C}_c}(f_k) &= \int_{\mathbb{R}^n} f_k(x) dx = \prod_{i=1}^n \left(\int_{\mathbb{R}} f_{ki}(x^i) dx^i \right) \\ &= \prod_{i=1}^n \left(b_i - a_i + \frac{1}{k} \right) \rightarrow \text{vol}(Q) \text{ für } k \rightarrow \infty \end{aligned} \quad (21.68)$$

Ähnlich zeigt man $\chi_{\overset{\circ}{Q}} \in \mathcal{V}^+$ und $J^+(\chi_{\overset{\circ}{Q}}) = I_{\mathcal{T}}(\chi_{\overset{\circ}{Q}}) = \text{vol}(Q)$.

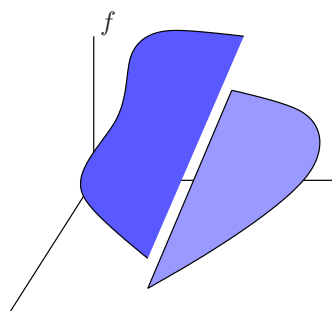
Integrierbare Funktionen

Klarerweise sind wir aber mit (\mathcal{V}^\pm, J^\pm) noch nicht am Ziel — Die \mathcal{V}^\pm sind keine Vektorräume, keine der neuen Funktionen liegen sowohl in \mathcal{V}^+ als auch in \mathcal{V}^- und \mathcal{V}^+ (bzw. \mathcal{V}^-) ist auch nicht nach unten (bzw. oben) monoton abgeschlossen. Teile dieser Aussagen folgen aus dem

Lemma 21.14. *Sei (für $\mathcal{V} = \mathcal{C}_c(\mathbb{R}^n)$) $f \in \mathcal{V}^+$. Dann ist f unterhalbstetig: $\forall x \in \mathbb{R}^n$ gilt: Ist $f(x) > C$ für ein $C \in \mathbb{R}$, so existiert ein $\delta > 0$ so, dass*

$$f(y) > C \quad \text{falls } \|x - y\| < \delta \quad (21.69)$$

Bildlich: f springt bei Annäherung an einen Punkt $x \in \mathbb{R}^n$ allenfalls runter, aber nie hoch.



Beweis. Für $f(x) < \infty$ ist die Behauptung: $\forall \epsilon > 0 \exists \delta > 0$ so, dass

$$f(y) > f(x) - \epsilon \quad \text{falls } \|x - y\| < \delta \quad (21.70)$$

Sei $(f_k) \subset \mathcal{C}_c(\mathbb{R}^n)$ mit $f_k \uparrow f$. Wegen $f_k(x) \uparrow f(x)$ existiert ein K so gross, dass $f_K(x) > f(x) - \frac{\epsilon}{2}$. Sodann existiert wegen der Stetigkeit von f_K ein $\delta > 0$ so, dass

$$|f_K(x) - f_K(y)| < \frac{\epsilon}{2} \quad \text{falls } \|x - y\| < \delta \quad (21.71)$$

d.h. insbesondere

$$f_K(y) > f_K(x) - \frac{\epsilon}{2} \quad (21.72)$$

Zusammen folgt

$$f(y) \geq f_K(y) > f_K(x) - \frac{\epsilon}{2} > f(x) - \epsilon \quad \text{falls } \|x - y\| < \delta \quad (21.73)$$

· Für $f(x) = \infty$ ist die Behauptung: $\forall N \in \mathbb{N} \exists \delta > 0$ so, dass

$$f(y) > N \quad \text{falls } \|x - y\| < \delta \quad (21.74)$$

und der Beweis ist sehr ähnlich. □

§ 21. INTEGRATION IM \mathbb{R}^n

· In Umkehrung gilt: Ist $f \in \mathcal{F}^+(\mathbb{R}^n)$ unterhalbstetig, und $f(x) \geq 0$ für x ausserhalb einer kompakten Menge, so ist $f \in \mathcal{V}^+$. Hat f beispielsweise kompakten Träger, so liegt die Folge (f_k) mit

$$f_k(x) := \inf\{f(y) + k\|x - y\| \mid y \in \mathbb{R}^n\} \quad (21.75)$$

in $\mathcal{C}_c(\mathbb{R}^n)$ und konvergiert monoton wachsend gegen f . (Übungsaufgabe)

· Man muss auch gar nicht unbedingt Folgen benutzen: Es gilt

$$f \in \mathcal{V}^+ \Leftrightarrow f = \sup\{g \mid g \in \mathcal{C}_c(\mathbb{R}^n), g \leq f\} \quad (21.76)$$

· Die Beziehung zwischen \mathcal{V}^- und *oberhalbstetigen* Funktionen ist natürlich ähnlich. Ausserdem gelten Aussagen der Art:

- Für $A \subset \mathbb{R}^n$ abgeschlossen ist χ_A oberhalbstetig.
- Für $U \subset \mathbb{R}^n$ offen ist χ_U unterhalbstetig.
- f ist ober- und unterhalbstetig $\Leftrightarrow f$ ist stetig.

· Weiterhin gilt:

$$f \in \mathcal{V}^+ \Leftrightarrow -f \in \mathcal{V}^- \quad (21.77)$$

(mit geeigneter Deutung $-1 \cdot \infty = -\infty$, s. (21.81)), und für $f \in \mathcal{V}^+$ ist

$$J^-(-f) = -J^+(f) \quad (21.78)$$

· Zuletzt bemerken wir, dass konstante positive Funktionen, d.h. $f(x) = f(0) > 0 \forall x \in \mathbb{R}^n$ in \mathcal{V}^+ liegen: Dies folgt beispielsweise aus

$$f(0) \cdot \chi_{B_k(0)} \uparrow f(0) \quad \text{für } k \rightarrow \infty, \quad (21.79)$$

$\chi_{B_k(0)} \in \mathcal{V}^+$, und der monotonen Abgeschlossenheit von \mathcal{V}^+ . Insbesondere halten wir fest, dass die konstante Funktion $f(x) = \infty \forall x$ in \mathcal{V}^+ liegt: Die konstanten Funktionen $f_k(x) = k \forall x$ wachsen monoton gegen ∞ . (Ebenso gilt $-\infty \in \mathcal{V}^-$.)

Prinzipiell ist es nicht weiter besorgniserregend, dass im Gegensatz zum totalgeordneten Körper \mathbb{Q} , den wir mit "einem einzigen Satz Dedekindscher Schnitte" (s. 1.11) zu \mathbb{R} vervollständigen konnten, $\mathcal{C}_c(\mathbb{R}^n)$ sich nicht in einem monotonen Schritt ordentlich abschliessen lässt. Überraschend ist vielmehr, dass wir mit Unterstützung von J^\pm zum Grössenvergleich im Raum *aller* Funktionen auf \mathbb{R}^n einen vollständigen Unterraum sogenannter *integrierbarer Funktionen* auszeichnen können, auf dem sich dann das Integral aller wünschenswerten Eigenschaften erfreut. Dazu erweitern wir nun unseren Wertebereich gleichzeitig in beide Richtungen. Es sei

$$\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\} \quad (21.80)$$

Mit den Verabredungen:

$$\begin{aligned} \infty + x &= \infty \quad \forall x \in \mathbb{R}^+, & -\infty + x &= -\infty \quad \forall x \in \mathbb{R}^- \\ x \cdot \infty &= \infty \text{ und } x \cdot (-\infty) = -\infty \quad \forall x \in \mathbb{R}, x > 0 & \text{ sowie } -1 \cdot \infty &= -\infty \\ \infty \cdot \infty &= \infty, & \infty \cdot (-\infty) &= -\infty, & (-\infty) \cdot (-\infty) &= \infty \end{aligned} \quad (21.81)$$

gelten dann die üblichen Rechenregeln "solange sie Sinn machen". Dies bedeutet insbesondere, dass man Bildungen wie $0 \cdot \infty$ und $\infty - \infty$ nicht ohne algebraische

Widersprüche einen Wert zuordnen kann. In der Integrationstheorie treten solche Ausdrücke zwar auf, aber nur so punktuell, dass ihr Wert keinen Einfluss auf das Ergebnis hat. I'm not happy, but if in doubt,

$$0 \cdot \infty = 0, \quad \infty - \infty = -\frac{1}{12} \tag{21.82}$$

Definition 21.15. Es seien wie oben J ein Daniell-Integral auf dem Funktionenverband $\mathcal{V} \subset \mathcal{F}(\mathbb{R}^n)$ und (\mathcal{V}^\pm, J^\pm) die monotonen Hüllen. Für $f \in \overline{\mathcal{F}}(\mathbb{R}^n) := \mathcal{F}(\mathbb{R}^n, \overline{\mathbb{R}})$ ²² heisst dann

$$\bar{J}^+(f) := \inf\{J^+(g) \mid g \in \mathcal{V}^+, g \geq f\} \tag{21.83}$$

das *Oberintegral* von f (bzgl. (\mathcal{V}, J)), und

$$\bar{J}^-(f) := \sup\{J^-(h) \mid h \in \mathcal{V}^-, h \leq f\} \tag{21.84}$$

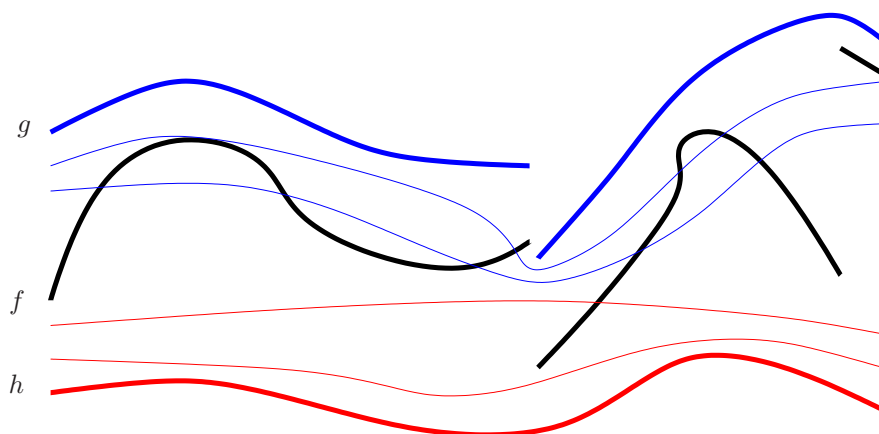
das *Unterintegral* von f .²³

• Eine Funktion $f \in \overline{\mathcal{F}}(\mathbb{R}^n)$ heisst *integrierbar* bzgl. (\mathcal{V}, J) , falls ihr Ober- und Unterintegral gleich und endlich sind. Dann heisst

$$\bar{J}(f) := \bar{J}^+(f) = \bar{J}^-(f) \in \mathbb{R} \tag{21.86}$$

ihr *Integral*. Wir schreiben $\bar{\mathcal{V}}$ für die Menge der integrierbaren Funktionen. Für $(\mathcal{V}, J) = (\mathcal{C}_c(\mathbb{R}^n), I_{\mathcal{C}_c})$ sagen wir auch “Lebesgue-integrierbar” und schreiben $\bar{\mathcal{L}}(\mathbb{R}^n)$ statt $\bar{\mathcal{V}}$ sowie

$$\int_{\mathbb{R}^n} f(x) dx \quad \left(\text{manchmal auch } \int_{\mathbb{R}^n} f(x) d^n x \right) \quad \text{statt } \bar{J}(f) \tag{21.87}$$



²²Beachte, dass ohne (21.82) $\overline{\mathcal{F}}(\mathbb{R}^n)$ kein Vektorraum ist.

²³Wegen $\pm\infty \in \mathcal{V}^\pm$, s. (21.79), sind die betroffenen Mengen nicht leer. Andernfalls macht es aber auch einen gewissen Sinn, jede Zahl in $\overline{\mathbb{R}}$ als untere/obere Schranke der leeren Menge anzusehen, und dann wegen $\infty > x > -\infty \forall x \in \mathbb{R}$

$$\sup(\emptyset) = \text{“kleinste obere Schranke”} = -\infty \quad \text{und} \quad \inf(\emptyset) = \infty \tag{21.85}$$

zu setzen. (Für eine nicht-leere Menge ist allerdings $\sup X \geq \inf(X)$ üblich.) Damit sind \bar{J}^\pm jedenfalls wohl-definiert.

§ 21. INTEGRATION IM \mathbb{R}^n

Lemma 21.16. (i) $\forall f \in \overline{\mathcal{F}}(\mathbb{R}^n)$ gilt $\bar{J}^+(f) \geq \bar{J}^-(f)$
(ii) $\forall f \in \mathcal{V}^+$ gilt $\bar{J}^+(f) = J^+(f) = \bar{J}^-(f)$. $f \in \mathcal{V}^+$ ist also genau dann integrierbar wenn $J^+(f) < \infty$. (Entsprechendes gilt für \mathcal{V}^- .)
(iii) $f \in \overline{\mathcal{F}}(\mathbb{R}^n)$ ist genau dann integrierbar, falls gilt: $\forall \epsilon > 0 \exists g_\epsilon \in \mathcal{V}^+, h_\epsilon \in \mathcal{V}^-$ mit $h_\epsilon \leq f \leq g_\epsilon$ und

$$J^+(g_\epsilon) - J^-(h_\epsilon) < \epsilon \quad (21.88)$$

Beweis. Hilfslemma: $\forall g \in \mathcal{V}^+, h \in \mathcal{V}^-$ mit $g \geq h$ gilt $J^+(g) \geq J^-(h)$

Bew.: Gemäss (21.78) ist $-h \in \mathcal{V}^+$ und $-J^-(h) = J^+(-h)$. Mit der Kegellinearität und der Monotonie von J^+ , s. 21.11, folgt

$$J^+(g) - J^-(h) = J^+(g) + J^+(-h) = J^+(\underbrace{g - h}_{\geq 0}) \geq 0 \quad (21.89)$$

und daraus die Behauptung.

(i) Aus dem Hilfslemma folgt

$$\{J^+(g) \mid g \in \mathcal{V}^+, g \geq f\} \geq \{J^-(h) \mid h \in \mathcal{V}^-, h \leq f\} \quad (21.90)$$

und durch inf/sup-Bildung daraus die Behauptung.

(ii) In Anbetracht von (i) genügt es zu zeigen, dass für $f \in \mathcal{V}^+$

$$\bar{J}^+(f) = J^+(f) \quad \text{und} \quad \bar{J}^-(f) \geq J^+(f) \quad (21.91)$$

Die erste Aussage folgt aus der Ungleichungskette

$$J^+(f) \leq \inf\{J^+(g) \mid g \in \mathcal{V}^+, g \geq f\} \leq J^+(f) \quad (21.92)$$

\uparrow \uparrow
 Monotonie $f \geq f$

Für die zweite sei $(f_k) \subset \mathcal{V}$ mit $f_k \uparrow f$. Dann gilt $f_k \leq f$ und $f_k \in \mathcal{V}^-$ und daher

$$\bar{J}^-(f) = \sup\{J^-(h) \mid h \in \mathcal{V}^-, h \leq f\} \geq \sup_k J^-(f_k) = \sup_k J(f_k) = J^+(f) \quad (21.93)$$

(iii) “ \Rightarrow ”: Für gegebenes $f \in \overline{\mathcal{F}}$, und $\epsilon > 0$ seien g_ϵ und h_ϵ wie angegeben. Dann folgt

$$\bar{J}^+(f) = \inf\{J^+(g) \mid g \in \mathcal{V}^+, g \geq f\} \leq J^+(g_\epsilon) < J^-(h_\epsilon) + \epsilon < \infty \quad (21.94)$$

Zusammen mit

$$\bar{J}^-(f) \geq J^-(h_\epsilon) \quad (21.95)$$

folgt

$$\bar{J}^+(f) \leq \bar{J}^-(f) + \epsilon \quad (21.96)$$

Mit $\epsilon \downarrow 0$ folgt (zusammen mit (i)!), dass $\bar{J}^+(f) = \bar{J}^-(f)$ und damit die Behauptung, dass $f \in \overline{\mathcal{V}}$.

“ \Leftarrow ”: Übung □

§ 22 Eigenschaften des Lebesgue-Integrals

Die technische Konstruktion des Integrals von reellwertigen Funktionen auf \mathbb{R}^n liegt nun grösstenteils hinter uns. Wir stellen in diesem § seine charakteristischen Stärken zusammen, machen ein paar Bemerkungen zu seiner Einzigartigkeit und Verallgemeinerungen, und geben eine wichtige Formel zu seiner Berechnung (nämlich, durch Rückführung auf “Quadraturen”, d.h. das Auffinden von Stammfunktionen, und andere bereits bekannte Integrale). Zunächst halten wir fest, in welchem Sinne wir mit dem Funktional \bar{J} auf $\bar{\mathcal{V}}$ unser Ziel erreicht haben.

Proposition/Definition 22.1. *Der Raum*

$$\mathcal{L}^1(\mathbb{R}^n) := \{f \in \bar{\mathcal{L}}(\mathbb{R}^n) \mid f(x) \in \mathbb{R} \forall x \in \mathbb{R}^n\} \quad (22.1)$$

der reellwertigen integrierbaren Funktionen ist abgeschlossen unter Addition, Multiplikation mit reellen Skalaren, sowie unter \max & \min und Absolutbetrag. M.a.W., $\mathcal{L}^1(\mathbb{R}^n) \subset \mathcal{F}(\mathbb{R}^n)$ ist ein Untervektorverband. $\bar{J}|_{\mathcal{L}^1(\mathbb{R}^n)}$ definiert auf $\mathcal{L}^1(\mathbb{R}^n)$ ein lineares und monotones Funktional, welches wir als

$$\mathcal{L}^1(\mathbb{R}^n) \ni f \mapsto \int_{\mathbb{R}^n} f(x) dx := \bar{J}(f) \quad (22.2)$$

schreiben.

Beweis. Für $f_1, f_2 \in \mathcal{L}^1(\mathbb{R}^n)$ und $g_1, g_2 \in \mathcal{V}^+$ mit $g_1 \geq f_1, g_2 \geq f_2$ ist $g_1 + g_2 \in \mathcal{V}^+$ und $g_1 + g_2 \geq f_1 + f_2$. Daraus folgt

$$\begin{aligned} \bar{J}^+(f_1) + \bar{J}^+(f_2) &= \inf\{J^+(g_1) \mid g_1 \in \mathcal{V}^+, g_1 \geq f_1\} + \inf\{J^+(g_2) \mid g_2 \in \mathcal{V}^+, g_2 \geq f_2\} \\ &= \inf\{J^+(g_1 + g_2) \mid g_1, g_2 \in \mathcal{V}^+, g_1 + g_2 \geq f_1 + f_2\} \\ &\geq \inf\{J^+(g) \mid g \in \mathcal{V}^+, g \geq f_1 + f_2\} = \bar{J}^+(f_1 + f_2) \end{aligned} \quad (22.3)$$

Ebenso folgt $\bar{J}^-(f_1 + f_2) \geq \bar{J}^-(f_1) + \bar{J}^-(f_2)$ und daraus die Ungleichungskette

$$\bar{J}^-(f_1 + f_2) \geq \bar{J}(f_1) + \bar{J}(f_2) \geq \bar{J}^+(f_1 + f_2) \quad (22.4)$$

Mit Lemma 21.16 (i) folgt $f_1 + f_2 \in \bar{\mathcal{V}}$ und $\bar{J}(f_1 + f_2) = \bar{J}(f_1) + \bar{J}(f_2)$.

• Der Beweis der Abgeschlossenheit unter Multiplikation mit nicht-negativen Skalaren folgt aus der von \mathcal{V}^\pm . Für die Multiplikation mit -1 beachte, dass mit (21.78)

$$\begin{aligned} \bar{J}^+(f) &= \inf\{J^+(g) \mid g \in \mathcal{V}^+, g \geq f\} = -\sup\{J^-(g) \mid g \in \mathcal{V}^+, g \geq f\} \\ &= -\sup\{J^-(h) \mid h \in \mathcal{V}^-, h \leq -f\} = -\bar{J}^-(f) \end{aligned} \quad (22.5)$$

und entsprechend $\bar{J}^-(f) = -\bar{J}^+(-f)$.

• Ist $f \in \bar{\mathcal{V}}$ mit $f \geq 0$ so folgt $0 \leq f$, d.h. $\bar{J}^-(f) \geq 0$ und daraus $\bar{J}(f) \geq 0$. Dies bedeutet, dass \bar{J} monoton ist.

• Für die Verbandsstruktur genügt es wegen der Linearität, der Beziehungen (21.51),

$$|f| = \max(f, 0) - \min(f, 0) \quad (22.6)$$

§ 22. EIGENSCHAFTEN DES LEBESGUE-INTEGRALS

sowie $\min(f, 0) = -\max(-f, 0)$ zu zeigen, dass mit f auch f_+ in $\mathcal{L}^1(\mathbb{R}^n)$ liegt. Wir tun dies in der Formulierung 21.16 (iii).

· Für $\epsilon > 0$ seien $g \in \mathcal{V}^+$, $h \in \mathcal{V}^-$ mit $h \leq f \leq g$ und $J^+(g) - J^-(h) < \epsilon$. Dann folgt aus der Kegolverbandsstruktur von \mathcal{V}^+ , dass $g_+ := \max(g, 0)$, $g_- := \min(g, 0) \in \mathcal{V}^+$ und ebenso $h_{\pm} \in \mathcal{V}^-$. Wegen der Kegellinearität und (21.78) folgt

$$\begin{aligned} \epsilon > J^+(g) - J^-(h) &= J^+(g_+) + J^+(g_-) + J^+(-h_+) + J^+(-h_-) \\ &= J^+(g_+ - h_+) + J^+(g_- - h_-) \end{aligned} \quad (22.7)$$

Aus $g \geq f \geq h$ folgt $g_{\pm} - h_{\pm} \geq 0$ und daraus wegen der Monotonie von J^+ , dass beide Ausdrücke auf der rechten Seite von (22.7) nicht-negativ und daher *getrennt* $< \epsilon$ sind. Damit gilt (wieder mit (21.78))

$$J^+(g_+) - J^-(h_+) < \epsilon \quad (22.8)$$

Wegen $g_+ \geq f_+ \geq h_+$ bedeutet dies gerade, dass $f_+ \in \mathcal{V}^+$. □

Bemerkungen. Die Abschätzung (22.3) und (22.4) gelten offenbar weiterhin für integrierbare Funktionen $f_1, f_2 \in \overline{\mathcal{L}}(\mathbb{R}^n)$ mit Werten in den erweiterten reellen Zahlen, wenn $f_1 + f_2$ an den Unendlichkeitsstellen mit den Regeln (21.81) bzw. für $\infty - \infty$ beliebig definiert ist. Jede solche Funktion ist also integrierbar und hat das gleiche Integral. In einem ähnlichen Sinne ist $\overline{\mathcal{L}}(\mathbb{R}^n)$ auch unter reeller Skalarmultiplikation und unter \min / \max abgeschlossen.

Korollar 22.2. Die Räume der Treppenfunktionen $\mathcal{T}(\mathbb{R}^n) \subset \mathcal{L}^1(\mathbb{R}^n)$, und der stetigen Funktionen mit kompaktem Träger $\mathcal{C}_c(\mathbb{R}^n) \subset \mathcal{L}^1(\mathbb{R}^n)$, und es gilt

$$\int_{\mathbb{R}^n} \varphi(x) dx = I_{\mathcal{T}}(\varphi) \quad \text{und} \quad \int_{\mathbb{R}^n} f(x) dx = I_{\mathcal{C}_c}(f) \quad (22.9)$$

Beweis. Für $\mathcal{C}_c(\mathbb{R}^n)$ folgen die Aussagen aus $\mathcal{C}_c(\mathbb{R}^n) = \mathcal{V} \subset \mathcal{V}^{\pm}$ und Lemma 21.16. Für $\mathcal{T}(\mathbb{R}^n)$ folgen die Aussagen aus dem Beispiel 21.13 und Gl. (21.25). □

Bemerkungen. Das Lebesgue-Integral wird mühelos mit nicht-beschränkten Funktionen und/oder Definitionsbereichen fertig (s. Übungen). Es ist etwas schwächer als das “uneigentliche Integral” im Sinne von (13.31), falls Letzteres nur bedingt und nicht absolut konvergiert. Beispiel:

$$\int_{-\infty}^{\infty} \frac{\sin x}{x} dx = \lim_{\alpha \rightarrow \infty} \int_0^{\alpha} + \lim_{\beta \rightarrow -\infty} \int_{\beta}^0 = \pi \quad (22.10)$$

existiert. (Das Integral lässt sich recht elegant mit komplexer Integration auswerten.) Der Absolutbetrag ist aber nicht integrierbar:

$$\int_0^{N\pi} \left| \frac{\sin x}{x} \right| dx \geq \sum_{k=0}^{N-1} \frac{1}{k+1} \int_{k\pi}^{(k+1)\pi} |\sin x| dx = \sum_{k=0}^{N-1} \frac{2}{k+1} \rightarrow \infty \text{ für } N \rightarrow \infty \quad (22.11)$$

Die Verträglichkeit des Lebesgue-Integrals mit anderen Grenzprozessen ist Gegenstand der folgenden

Vertauschungssätze

Hier betrifft das erste Resultat die Vollständigkeit von $\bar{\mathcal{V}} = \bar{\mathcal{L}}(\mathbb{R}^n)$ und die Ordnungsstetigkeit von \bar{J} , in Erfüllung des Prinzips der monotonen Konvergenz.

Theorem 22.3 (Satz von Beppo Levi). *Es sei $(f_k)_{k \in \mathbb{N}} \subset \bar{\mathcal{V}}$ eine monotone Folge integrierbarer Funktionen. Angenommen, die Folge $(\bar{J}(f_k))_{k \in \mathbb{N}}$ der Integrale ist beschränkt. Dann ist die punktweise gebildete Grenzfunktion $f(x) := \lim_{k \rightarrow \infty} f_k(x)$ integrierbar mit $\bar{J}(f) = \lim_{k \rightarrow \infty} \bar{J}(f_k)$.*

Beweis. Wir behandeln den Fall $f_k \uparrow f$ ($f_k \downarrow f$ geht analog). Wegen der Monotonie von \bar{J} (s. 22.1) ist dann die Folge $(\bar{J}(f_k))_{k \in \mathbb{N}}$ monoton wachsend und nach Voraussetzung beschränkt. Wir bezeichnen mit $S = \lim_{k \rightarrow \infty} \bar{J}(f_k) \in \mathbb{R}$ ihren Grenzwert und zeigen zwei Ungleichungen.

1. $\bar{J}^-(f) \geq S$: Dies folgt aus

$$\begin{aligned} \bar{J}^-(f) &= \sup\{J^-(h) \mid h \in \mathcal{V}^-, h \leq f\} \\ &\geq \sup\{J^-(h) \mid h \in \mathcal{V}^-, h \leq f_k \text{ für ein } k \in \mathbb{N}\} \\ &= \sup_k \bar{J}^-(f_k) = \sup_k \bar{J}(f_k) = S \end{aligned} \quad (22.12)$$

(Wegen $f_k \leq f \forall k$ steht in der ersten Zeile eine Obermenge der zweiten.)

2. $\bar{J}^+(f) \leq S$: Die Folge $(\delta_k)_{k \in \mathbb{N}} \subset \bar{\mathcal{V}}$ mit $\delta_1 := f_1, \delta_k := f_k - f_{k-1}$ für $k > 1$ erfüllt

$$\delta_k \geq 0 \quad \forall k > 1, \quad \sum_{k=1}^{\infty} \delta_k = f \quad \text{und} \quad \sum_{k=1}^{\infty} \bar{J}(\delta_k) = S. \quad (22.13)$$

Für gegebenes $\epsilon > 0$ sei für jedes k $g_k \in \mathcal{V}^+$ mit $g_k \geq \delta_k$ und $J^+(g_k) \leq \bar{J}(\delta_k) + \frac{\epsilon}{2^k}$.

· Aus $g_k \geq 0$ für $k > 1$ folgt wegen der monotonen Abgeschlossenheit von \mathcal{V}^+ , dass $\sum g_k \in \mathcal{V}^+$, und aus $g_k \geq \delta_k \forall k$, dass $\sum g_k \geq \sum \delta_k = f$. Mit der Ordnungsstetigkeit von J^+ und obigen Feststellungen folgt

$$\bar{J}^+(f) \leq J^+\left(\sum_{k=1}^{\infty} g_k\right) = \sum_{k=1}^{\infty} J^+(g_k) \leq \sum_{k=1}^{\infty} \left(\bar{J}(\delta_k) + \frac{\epsilon}{2^k}\right) = S + \epsilon \quad (22.14)$$

(geometrische Reihe!). Da dies für jedes $\epsilon > 0$ gilt, folgt die behauptete Ungleichung. Zusammen mit 21.16 (i) folgt aus den beiden Ungleichungen $\bar{J}^+(f) = \bar{J}^-(f) = S$ wie behauptet. \square

Ebenso häufig verwendet wird die Integralversion der “dominierten Konvergenz” (vgl. Majorantenkriterium auf S. 36).

Theorem 22.4 (Satz von Lebesgue). *Sei $(f_k)_{k \in \mathbb{N}} \subset \bar{\mathcal{L}}(\mathbb{R}^n)$ eine punktweise (aber nicht notwendig monotone) konvergente Folge integrierbarer Funktionen. Angenommen, es existiert eine integrierbare Funktion $F \in \bar{\mathcal{L}}(\mathbb{R}^n)$ mit $|f_k| \leq F \forall k$.²⁴ Dann ist die punktweise gebildete Grenzfunktion $f(x) := \lim_{k \rightarrow \infty} f_k(x)$ integrierbar mit*

$$\int_{\mathbb{R}^n} f(x) dx = \lim_{k \rightarrow \infty} \int_{\mathbb{R}^n} f_k(x) dx \quad (22.15)$$

Beweis. · Die Folge $(h_k)_{k \in \mathbb{N}}$ mit $h_k := \inf\{f_l \mid l \geq k\}$ ist monoton wachsend mit $h_k \uparrow f$. Wir behaupten zunächst, dass $h_k \in \bar{\mathcal{L}}(\mathbb{R}^n)$.

²⁴Es genügt offenbar $F \in \bar{\mathcal{F}}(\mathbb{R}^n)$ mit $|f_k| \leq F$ und $\bar{J}^+(F) < \infty$.

§ 22. EIGENSCHAFTEN DES LEBESGUE-INTEGRALS

Bew.: Für festes k ist die Folge $(h_{kj})_{j \in \mathbb{N}}$ mit $h_{kj} := \min\{f_k, f_{k+1}, \dots, f_{k+j-1}\}$ monoton fallend mit $h_{kj} \downarrow h_k$. Als Minimum endlich vieler integrierbarer Funktionen ist $h_{kj} \in \overline{\mathcal{L}}(\mathbb{R}^n)$. ($\overline{\mathcal{L}}(\mathbb{R}^n)$ ist ein Verband.) Wegen $|f_k| \leq F$ gilt $f_k \geq -F \forall k$, und daraus folgt $h_{kj} \geq -F$ sowie $\bar{J}(h_{kj}) \geq \bar{J}(-F) > -\infty$. Die Behauptung folgt dann aus dem Satz 22.3 von Beppo Levi.

· Wegen $f_k \leq F$ ist auch $h_k \leq F \forall k$ und daher $\bar{J}(h_k) \leq \bar{J}(F) < \infty$ (Monotonie von \bar{J}). Wieder mit Beppo Levi folgt daher aus $h_k \uparrow f$, dass f integrierbar ist mit $\bar{J}(h_k) \uparrow \bar{J}(f)$.

· Ebenso gilt, dass die monoton fallende Folge $(g_k)_{k \in \mathbb{N}}$ mit $g_k := \sup_l \{f_l \mid l \geq k\}$ integrierbar ist mit $g_k \downarrow f$ und $\bar{J}(g_k) \downarrow \bar{J}(f)$.

· Es bleibt noch zu zeigen, dass $\lim_{k \rightarrow \infty} \bar{J}(f_k) = f$. Dies folgt aber aus $h_k \leq f_k \leq g_k$ wegen der Monotonie von \bar{J} und des Sandwich-Lemmas 3.6. \square

Eine nützliche Anwendung der Vertauschungssätze ist die Verallgemeinerung der “Integration über Teilmengen” auf \mathbb{R}^n . (Für $n = 1$ war die Integration über kompakte Intervalle unser Ausgangspunkt gewesen, die Integration über offene Intervalle hatten wir in 13.31 “uneigentlich” definiert.) Wir werden weiterhin nicht viel mehr brauchen als die Aussage, dass für $f \in \mathcal{F}^1(\mathbb{R}^n)$ und eine offene Menge $U \subset \mathbb{R}^n$ auch $f \cdot \chi_U \in \mathcal{L}^1(\mathbb{R}^n)$, und für eine abgeschlossene Menge $A \subset \mathbb{R}^n$ auch $f \cdot \chi_A \in \mathcal{L}^1(\mathbb{R}^n)$, und wir schreiben natürlich

$$\int_U f(x) dx := \int_{\mathbb{R}^n} f(x) \cdot \chi_U(x) dx \quad \text{und} \quad \int_A f(x) dx := \int_{\mathbb{R}^n} f(x) \cdot \chi_A(x) dx. \quad (22.16)$$

Zur Begründung bemerken wir erstens, dass es wegen der Linearität und der Verbandseigenschaft von $\mathcal{L}^1(\mathbb{R}^n)$ genügt, die Aussage für $f \geq 0$ zu zeigen, und zweitens, dass $\chi_U \in \mathcal{V}^+$ (s. Diskussion auf S. 159). Es existiert also eine Folge $(\gamma_k) \subset \mathcal{C}_c(\mathbb{R}^n)$ mit $\gamma_k \geq 0$ und $\gamma_k \uparrow \chi_U$. Die Behauptung folgt dann aus $f \cdot \gamma_k \in \mathcal{F}^1(\mathbb{R}^n)$, $f \cdot \gamma_k \uparrow f \cdot \chi_U$, und $\int_{\mathbb{R}^n} (f \cdot \gamma_k)(x) dx \leq \int_{\mathbb{R}^n} f(x) dx$. Abgeschlossene Mengen gehen ähnlich.

Die Definition (22.16) greift auch, wenn f a priori gar nicht auf ganz \mathbb{R}^n definiert ist. Wir setzen dazu einfach $(f \cdot \chi_U)(x) := 0$ für $x \notin U$.

Ausblick

Das Resultat des hier vorgestellten sog. Daniell-Prozesses zur Definition des Integrals auf \mathbb{R}^n ist äquivalent zu dem üblichen, und physikalisch vielleicht durchaus leichter zu motivierenden (dafür aber noch(!) längeren), Zugang über die *Maßtheorie*.

· Die Maßtheorie beschäftigt sich ganz allgemein mit dem Problem, ein gewisses System $\mathcal{A} \subset \mathcal{P}(X)$ von Teilmengen einer beliebigen Menge X als “messbare Mengen” auszuzeichnen²⁵, und ihnen ein gewisses “Gewicht” oder “Volumen” (das *Maß*) $\mu(A) \in [0, \infty]$ für $A \in \mathcal{A}$ zuzuweisen. Dabei sind \mathcal{A} und μ gewissen natürlichen Kompatibilitätsbedingungen unterworfen. Insbesondere sollte $\emptyset \in \mathcal{A}$ mit $\mu(\emptyset) = 0$ sein und für $A_1, A_2 \in \mathcal{A}$ sollten auch $A_i^c := X \setminus A_i$, $A_1 \cap A_2$ sowie $A_1 \cup A_2$ messbar sein und es sollte gelten

$$\mu(A_1 \cup A_2) = \mu(A_1) + \mu(A_2) - \mu(A_1 \cap A_2) \quad (22.17)$$

²⁵ $\mathcal{P}(X)$ ist die “Potenzmenge”, d.h. die Menge aller Teilmengen, von X .

Man verlangt auch Verallgemeinerungen dieser Bedingungen auf *abzählbare*²⁶ Systeme von messbaren Teilmengen.

· Dieser Rahmen ist zunächst sehr allgemein und erlaubt auch im Falle des \mathbb{R}^n eine Vielzahl von verschiedenen Konstruktionen. Nicht-triviale Beispiele liefern bereits das sog. Zählmaß ($\mu(A) = |A|$ falls A endlich und $\mu(A) = \infty$ sonst) und das sog. Punkt- oder Diracmaß bei 0 mit $\mu(A) = 1$ falls $0 \in A$, $\mu(A) = 0$ sonst). Verlangt man jedoch,

1. dass das Maß *translationsinvariant* ist, d.h.

$$\mu(A + x) = \mu(A) \quad \forall A \in \mathcal{A}, x \in \mathbb{R}^n \quad (22.18)$$

und

2. dass abgeschlossene Quader messbar sind, und normiert man das Volumen des Einheitswürfels auf

$$\mu([0, 1] \times \cdots \times [0, 1]) = 1 \quad (22.19)$$

so kann man zeigen, dass diese Bedingungen das sog. Lebesgue-Maß auf \mathbb{R}^n eindeutig festlegen. Beispielsweise zeigt man durch “Teilung” und “Ausschöpfung”, dass wie erwartet $\mu(Q) = \text{vol}(Q)$ für einen beliebigen (kompakten) Quader Q .

· Zur Definition eines Integrals auf einer Teilmenge von $\mathcal{F}(\mathbb{R}^n, \mathbb{R})$ ordnet man dann zunächst den charakteristischen Funktionen von $A \in \mathcal{A}$ das Maß $\mu(A)$ zu und setzt dies linear auf endliche Linearkombinationen fort. “Messbare” und “integrierbare” Funktionen sind dann solche, die sich “unter dem Integral” durch diese endlichen Linearkombinationen approximieren lassen mit einem endlichen Wert für ihr Integral. (In der Moral recht ähnlich zur Definition 21.15.)

· Scheinbar erfordert diese Konstruktion vor der eigentlichen Definition des Integrals einen unabhängigen Grenzprozess zur Definition von Volumina der “messbaren” Mengen. Der Vorteil der Maßtheorie ist die grössere Flexibilität, dass sich also beispielsweise Punkt- und Zählmaß “auf einer Stufe” behandeln lassen. Interessant sind insbesondere *Wahrscheinlichkeitsmaße* auf beliebigen Mengen X , d.h. Maße mit der Eigenschaft, dass $\mu(X) = 1$.

· Hier wollen wir nur festhalten, wie man aus unserem Integralbegriff im Prinzip das Lebesgue-Maß zurückgewinnen kann.

Definition 22.5. Eine Teilmenge $X \subset \mathbb{R}^n$ heisst *endlich messbar* oder *integrierbar*, falls $\chi_X \in \mathcal{L}^1(\mathbb{R}^n)$, und wir setzen

$$\text{vol}(X) := \int_{\mathbb{R}^n} \chi_X(x) dx \quad (22.20)$$

· Dass beschränkte abgeschlossene und offene Mengen endlich messbar sind, hatten wir bereits festgehalten. Die Eigenschaft (22.17) folgt aus der Linearität des Integrals und der Identität

$$\chi_{A_1} + \chi_{A_2} = \chi_{A_1 \cap A_2} + \chi_{A_1 \cup A_2} \quad (22.21)$$

²⁶Eine (nicht-leere) Menge X heisst (höchstens) abzählbar, falls es eine Folge $(x_k)_{k \in \mathbb{N}} \subset X$ gibt, die jedes Element von X mindestens einmal trifft. Die rationalen Zahlen \mathbb{Q} sind abzählbar, die reellen Zahlen sind nicht abzählbar, s. 3.17.

§ 22. EIGENSCHAFTEN DES LEBESGUE-INTEGRALS

bzw. verbandstheoretisch aus (21.51) zusammen mit Beziehungen der Art

$$\begin{aligned}\chi_{A_1 \cup A_2} &= \max(\chi_{A_1}, \chi_{A_2}), & \chi_{A_1 \cap A_2} &= \min(\chi_{A_1}, \chi_{A_2}) = \chi_{A_1} \cdot \chi_{A_2} \\ |\chi_{A_1} - \chi_{A_2}| &= \chi_{A_1 \setminus A_2} + \chi_{A_2 \setminus A_1}\end{aligned}\quad (22.22)$$

und dergleichen mehr. (Achtung: Wir haben über die Multiplikation von integrierbaren Funktionen noch nicht allgemein gesprochen!) Die Translationsinvarianz folgt über unsere Grenzprozesse aus der Translationsinvarianz des Integrals auf $\mathcal{C}_c(\mathbb{R}^n)$ (s. auch 22.10).

· Beim Aufbau einer Maß- und Integrationstheorie spielen diejenigen Mengen eine besondere Rolle, die “von dem Maß gar nicht gesehen” werden.

Definition 22.6. (i) Eine Funktion $f \in \overline{\mathcal{F}}(\mathbb{R}^n)$ heisst Nullfunktion, falls ihr Oberintegral $\bar{J}^+(|f|) = 0$.

(ii) Eine Teilmenge $N \subset \mathbb{R}^n$ heisst Nullmenge, falls χ_N eine Nullfunktion ist.

(Wegen $\chi_N = |\chi_N|$ und $0 \leq \bar{J}^-(\chi_N) \leq \bar{J}^+(\chi_N)$ ist die charakteristische Funktion χ_N einer Nullmenge N selbst auch integrierbar, während allgemein die Implikation $f \in \overline{\mathcal{L}}(\mathbb{R}^n) \Rightarrow |f| \in \overline{\mathcal{L}}(\mathbb{R}^n)$ nur in einer Richtung gilt.)

Sprechweise: Man sagt, eine Aussage gelte “fast überall” (f.ü.), falls sie ausserhalb einer Nullmenge wahr ist. (Man vergleiche dies in Anbetracht der unten gegebenen Charakterisierungen von Nullmengen mit Aussagen, welche “für fast alle k ” gelten, wenn sie höchstens für endlich viele k falsch sind.)

Proposition 22.7. *Es sei $f \in \overline{\mathcal{L}}(\mathbb{R}^n)$.*

(i) *Die Menge der Unendlichkeitsstellen von f , d.h.*

$$N_\infty := \{x \in \mathbb{R}^n \mid f(x) = \infty \text{ oder } -\infty\} \quad (22.23)$$

ist eine Nullmenge.

(ii) *Ist $g \in \overline{\mathcal{F}}(\mathbb{R}^n)$ eine weitere Funktion, welche fast überall mit f übereinstimmt, so ist g ebenfalls integrierbar mit $\bar{J}(f) = \bar{J}(g)$.*

Beweis. (i) Mit den Vereinbarungen (21.82) ist $\chi_{N_\infty} = -12(f + (-f))$ und nach den Bemerkungen zu Theorem 22.1 bzw. aus der Abschätzung (22.3) folgt $\bar{J}^+(\chi_{N_\infty}) \leq \bar{J}^+(-12f) + \bar{J}^+(12f) = -12\bar{J}(f) + 12\bar{J}(f) = 0$.

(ii) Nach Voraussetzung ist $N := \{x \in \mathbb{R}^n \mid g(x) - f(x) \neq 0\}$ eine Nullmenge. Die Folge $(h_k)_{k \in \mathbb{N}}$ mit $h_k := k\chi_N$ ist integrierbar und monoton wachsend, mit $\bar{J}(h_k) = 0 < \infty \forall k$. Nach dem Satz 22.3 von Beppo Levi ist daher $\infty \cdot \chi_N = \sup_k h_k$ integrierbar mit $\bar{J}(\infty \cdot \chi_N) = 0$. Folglich ist $|g - f| \leq \infty \cdot \chi_N$ eine Nullfunktion. Aus $g \leq f + |g - f|$ bzw. $g \geq f - |g - f|$ folgt

$$\bar{J}^+(g) \leq \bar{J}^+(f) = \bar{J}^-(f) \leq \bar{J}^-(g) \quad (22.24)$$

und mit 21.16 (i) die Behauptung. □

Beispiele: · Entartete Quader sind Nullmengen. Nullmengen müssen aber nicht beschränkt sein. So konvergieren die charakteristischen Funktionen der entarteten Quader $\{0\} \times [-k, k]^{\times n-1}$ monoton wachsend gegen die charakteristische Funktion der Koordinatenebene $\{x_1 = 0\} \subset \mathbb{R}^n$, so dass aus 22.3 folgt, dass das n -dimensionale Volumen von \mathbb{R}^{n-1} verschwindet.

$$\int_{\mathbb{R}^n} \chi_{\{x_1=0\}} d^n x = 0 \tag{22.25}$$

· Ist $(N_k)_{k \in \mathbb{N}}$ ein abzählbares System von Nullmengen, so wächst die Folge

$$\max\{\chi_{N_l} \mid l \leq k\} \tag{22.26}$$

monoton gegen $\chi_{\cup_k N_k}$, so dass wieder aus 22.3 folgt, dass

$$\cup_{k=1}^{\infty} N_k = N_1 \cup N_2 \cup N_3 \cup \dots \tag{22.27}$$

ebenfalls eine Nullmenge ist. Daraus folgt etwa insbesondere auch, dass $\mathbb{Q} \subset \mathbb{R}$ eine Nullmenge ist. Das Überraschende daran ist, dass \mathbb{Q} in \mathbb{R} dicht liegt, d.h. $\overline{\mathbb{Q}} = \mathbb{R}$. Vielleicht noch überraschender ist allerdings die Existenz von überabzählbaren dichten Nullmengen (s. Übungen).

Beispiel einer nicht-integrierbaren beschränkten Funktion mit kompaktem Träger: Wir wählen²⁷ ein System $X \subset [0, 1]$ von Repräsentanten der “(Ir-)Rationalitätsklassen” von \mathbb{R} , d.h. X ist derart, dass für jedes $y \in \mathbb{R}$ genau ein $x \in X$ existiert mit $q := x - y \in \mathbb{Q}$. Mit anderen Worten zerlegen wir

$$\mathbb{R} = \bigcup_{x \in X} (x + \mathbb{Q}) = \bigcup_{q \in \mathbb{Q}} (q + X) \tag{22.28}$$

Beh.: χ_X ist nicht integrierbar.

Bew.: Wäre χ_X integrierbar, so folgte aus der Translationsinvarianz des Integrals/Maßes, dass auch alle χ_{q+X} integrierbar sind mit $\text{vol}(q+X) = \text{vol}(X) \forall q \in \mathbb{Q}$.

· Wäre $\text{vol}(X) = 0$, so wäre \mathbb{R} als abzählbare Vereinigung von Nullmengen ebenfalls eine Nullmenge, was natürlich nicht der Fall ist.

· Es müsste also $\text{vol}(X) > 0$ sein. Dann aber hätte die Menge

$$\bigcup_{q \in \mathbb{Q} \cap [0,1]} (q + X) \subset [0, 2] \tag{22.29}$$

einerseits als disjunkte Vereinigung unendlich vieler Mengen gleichen positiven Volumens kein endliches Volumen, andererseits als messbare Teilmenge von $[0, 2]$ Volumen kleiner als 2, ein Widerspruch.

Für die Berechnung des Integrals bleibt die Verallgemeinerung von 21.6 eines der wichtigsten Hilfsmittel.

Theorem 22.8 (Satz von Fubini). *Sei $f \in \mathcal{L}^1(\mathbb{R}^{n+m})$. Dann ist für y ausserhalb einer Nullmenge in \mathbb{R}^m die Funktion $x \mapsto f^{(y)}(x) := f(x, y)$ in $\mathcal{L}^1(\mathbb{R}^n)$, die fast*

²⁷mit dem Auswahlaxiom

§ 22. EIGENSCHAFTEN DES LEBESGUE-INTEGRALS

überall definierte Funktion $g : y \mapsto \bar{J}(f^{(y)})$ ist in $\mathcal{L}^1(\mathbb{R}^m)$ und es gilt $\bar{J}(g) = \bar{J}(f)$. Man schreibt dies nicht ganz präzise als

$$\int_{\mathbb{R}^m} \left(\int_{\mathbb{R}^n} f(x, y) dx \right) dy = \int_{\mathbb{R}^{n+m}} f(x, y) d(x, y) = \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^m} f(x, y) dy \right) dx \quad (22.30)$$

Ausserdem gilt in Verallgemeinerung von 13.7:

Theorem 22.9 (Stetige/differenzierbare Abhängigkeit von Parametern). *Es sei $U \subset \mathbb{R}^m$ offen und $f : \mathbb{R}^n \times U \rightarrow \mathbb{R}$ eine Funktion mit den Eigenschaften:*

- (i) Für jedes $y \in U$ ist $x \mapsto f(x, y)$ integrierbar.
- (ii) Für jedes $x \in \mathbb{R}^n$ ist $y \mapsto f(x, y)$ stetig.
- (iii) Es existiert eine integrierbare Majorante $F \in \mathcal{L}^1(\mathbb{R}^n)$ mit $|f(x, y)| \leq F(x) \forall (x, y) \in \mathbb{R}^n \times U$. Dann ist die Funktion

$$y \mapsto \int_{\mathbb{R}^n} f(x, y) dx \quad (22.31)$$

stetig.

· Ersetzt man in (ii) stetig durch stetig differenzierbar in Richtung $w \in \mathbb{R}^m$ und dominiert in (iii) F die Richtungsableitung $D_w f(x, y)$, so ist letztige $\forall y$ integrierbar, die Funktion in (22.31) ist stetig differenzierbar in Richtung w und es gilt

$$D_w \left(\int_{\mathbb{R}^n} f(x, y) dx \right) = \int_{\mathbb{R}^n} D_w f(x, y) dx \quad (22.32)$$

Beweis. Lassen wir beide aus. □

Stattdessen halten wir noch fest, dass sich das mehr-dimensionale Integral auch auf Funktionen mit Werten in vollständigen normierten Vektorräumen (insbesondere auch Banach-Algebren) verallgemeinern lässt. Für stetige Funktionen mit kompaktem Träger ist dies einfach die Kombination von 13.3 mit 21.4, zur Vervollständigung müsste man allerdings einen anderen Weg einschlagen als die monotone Konvergenz. Wir werden dies nur für Funktionen mit Werten in \mathbb{C} , bzw. $\bar{\mathbb{C}} := \mathbb{C} \cup \{\infty\}$ benutzen. Dann ist whatever Definition äquivalent zur Aussage:

· Eine erweitert komplexwertige Funktion $f \in \mathcal{F}(\mathbb{R}^n, \bar{\mathbb{C}})$ ist genau dann integrierbar, wenn ihre (Unendlichkeitsstellenmenge eine Nullmenge ist und die entsprechende modifizierten) Real- und Imaginärteile im Sinne von 21.15 integrierbar sind, und es gilt

$$\int_{\mathbb{R}^n} f(x) dx = \int_{\mathbb{R}^n} \operatorname{Re}(f)(x) dx + i \int_{\mathbb{R}^n} \operatorname{Im}(f)(x) dx \quad (22.33)$$

Wir schreiben $\mathcal{L}^1(\mathbb{R}^n, \mathbb{C})$ für die komplexwertigen integrierbaren Funktionen. Die Vertauschungssätze (ausser natürlich Beppo Levi) gelten sinngemäss.

Transformationsformel

Eines der wichtigsten Hilfsmittel zur Berechnung von ein-dimensionalen bestimmten Integralen ist die *Substitutionsregel*, welche wir in (13.18) als die Umkehrung der

Kettenregel festgehalten hatten: Ist $\tau : J \rightarrow I$ stetig differenzierbar und $f : I \rightarrow \mathbb{R}$ stetig, so gilt für alle $\alpha, \beta \in J$:

$$\int_{\tau(\alpha)}^{\tau(\beta)} f(s)ds = \int_{\alpha}^{\beta} f(\tau(\sigma))\tau'(\sigma)d\sigma \quad (22.34)$$

· Zur Übertragung auf unser mehr-dimensionales Integral müssen wir zunächst unserer Entscheidung Rechnung tragen, bei der Mittelwertbildung von Treppenfunktionen in (21.26) den elementaren Quadern in (21.20) *nicht-orientierte* Volumina zugeordnet zu haben.

· Ist etwa $f \in \mathcal{C}_c(\mathbb{R})$ mit $\text{supp}(f) \subset [a, b]$ ($b > a!$), so gilt nach (21.6)

$$\int_{\mathbb{R}} f(x)dx = \int_a^b f(s)ds \quad (22.35)$$

Mit der “Substitution” $s = \lambda\sigma$ für $\sigma \in \mathbb{R}$ und $\lambda \in \mathbb{R}$ können wir dies nur dann als

$$\int_a^b f(s)ds = \lambda \int_{\alpha}^{\beta} f(\lambda\sigma)d\sigma \quad (22.36)$$

schreiben, wenn $\alpha, \beta \in \mathbb{R}$ existieren mit $a = \lambda\alpha$ und $b = \lambda\beta$, m.a.W. falls $\lambda \neq 0$. Ist dann $\lambda > 0$, so erhalten wir

$$\int_a^b f(s)ds = \lambda \int_{a/\lambda}^{b/\lambda} f(\lambda\sigma)d\sigma = \lambda \int_{\mathbb{R}} f(\lambda\xi)d\xi \quad (22.37)$$

Im Fall $\lambda < 0$ ist andererseits $\alpha < \beta$ und daher wegen (21.5)

$$\lambda \int_{\alpha}^{\beta} f(\lambda\sigma)d\sigma = -\lambda \int_{\mathbb{R}} f(\lambda\xi)d\xi \quad (22.38)$$

Unser Integral transformiert sich also unter der *invertierbaren* Substitution $x = \lambda\xi$ in beiden Fällen gemäss

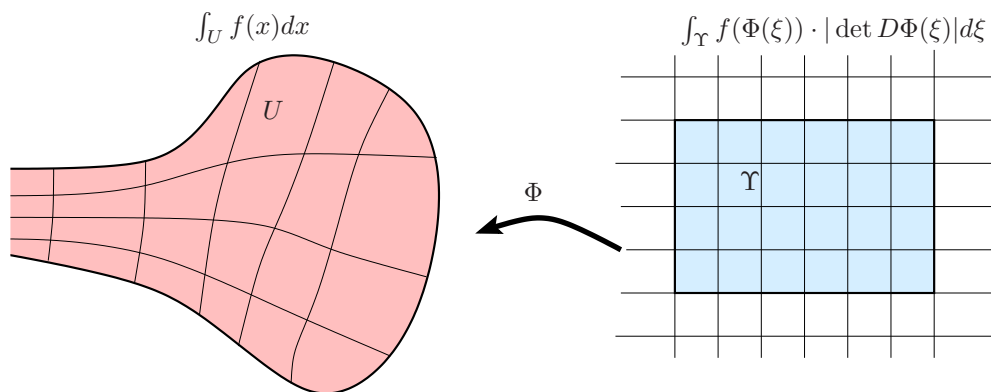
$$\int_{\mathbb{R}} f(x)dx = |\lambda| \int_{\mathbb{R}} f(\lambda\xi)d\xi \quad (22.39)$$

(Anstatt mit dem Hauptsatz der ein-dimensionalen Differential- und Integralrechnung könnte man diese Formel auch direkt über den Grenzprozess 21.4 aus der Definition (21.20) herleiten.) In n Dimensionen gilt

Theorem 22.10. *Es sei $\Phi : \Upsilon \rightarrow U$ ein Diffeomorphismus zwischen offenen Teilmengen des \mathbb{R}^n , d.h. Φ ist stetig differenzierbar und bijektiv, und das Differential $D\Phi(\xi) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ist an jedem Punkt $\xi \in \Upsilon$ invertierbar. Ferner sei $f \in \mathcal{F}(U, \mathbb{R})$ über U integrierbar, d.h. $f \cdot \chi_U \in \mathcal{L}^1(\mathbb{R}^n)$. Dann ist $f \circ \Phi \cdot |\det D\Phi|$ über Υ integrierbar und es gilt*

$$\int_U f(x)dx = \int_{\Upsilon} f(\Phi(\xi)) |\det D\Phi(\xi)| d\xi \quad (22.40)$$

§ 22. EIGENSCHAFTEN DES LEBESGUE-INTEGRALS



Beweisstrategie: Mit Hilfe von (22.39), der Translationsinvarianz (welche wir jetzt im Vorbeigehen mitbeweisen), und dem Satz 21.6 begründen wir zunächst die Transformation von I_{C_c} unter affin-linearen Substitutionen des gesamten \mathbb{R}^n . Durch verschiedene Grenzprozesse lockern wir dann schrittweise die Voraussetzungen an U , Φ , und f .

· Zur Notation bemerken wir, dass wir mit $f \cdot \chi_U$ die triviale Fortsetzung durch 0 von $f \in \mathcal{F}(U, \mathbb{R})$ zu einer Funktion in $\mathcal{F}(\mathbb{R}^n, \mathbb{R})$ meinen.²⁸ Den Integranden auf der rechten Seite von (22.40) schreiben wir zur Abkürzung als

$$f^\Phi(\xi) := f(\Phi(\xi)) \cdot |\det D\Phi(\xi)| \tag{22.41}$$

· Ausserdem halten wir fest, dass falls wir die Behauptung für (eine Klasse von) $\Phi : \Upsilon \rightarrow U$ und $\Psi : \mathfrak{U} \rightarrow \Upsilon$ bewiesen haben, sie wegen der Kettenregel und Multiplikativität der Determinanten auch für $\Phi \circ \Psi$ gilt: Aus

$$|\det D(\Phi \circ \Psi)(\mathfrak{r})| = |\det D\Phi(\Psi(\mathfrak{r}))| \cdot |\det D\Psi(\mathfrak{r})| \tag{22.42}$$

folgt $(f^\Phi)^\Psi = f^{\Phi \circ \Psi}$ für alle f . Insbesondere ist für $\Psi = \Phi^{-1}$:

$$(f^\Phi)^{\Phi^{-1}} = f \tag{22.43}$$

· Wegen der Linearität der Integrale über U bzw. Υ und der Verbandsstruktur der integrierbaren Funktionen (insbesondere, f integrierbar $\Leftrightarrow f_+ = \max(f, 0)$ und $f_- = \min(f, 0)$ integrierbar), genügt es, die Aussagen für $f \geq 0$ zu zeigen. Wir nehmen auch gelegentlich an, dass f nicht identisch verschwindet, da die Aussage ja sonst trivial ist.

1. Schritt: Es sei $U = \Upsilon = \mathbb{R}^n$, Φ eine affine Transformation, d.h. mit einer invertierbaren Matrix $A \in GL(n, \mathbb{R})$ und $b \in \mathbb{R}^n$ gilt

$$\Phi(\xi) = A\xi + b, \tag{22.44}$$

und $f \in C_c(\mathbb{R}^n)$. Beachte, dass $D\Phi = A$.

²⁸Eine vielleicht berücksichtigenswerte Alternative wäre, den Raum $C_c(U, \mathbb{R})$ dem Daniell-Prozess zu unterwerfen.

· Wegen der Bemerkung um (22.42) genügt es, die Aussage für
 (i) Translationen (d.h. $A = \text{id}$)
 und lineare Substitutionen (d.h. $b = 0$) getrennt zu zeigen. Jede invertierbare Matrix wiederum lässt sich als Produkt von sog. Elementarmatrizen schreiben, welche
 (ii) Multiplikation einer Komponente mit einer Konstanten (d.h. A diagonal mit genau einem Eintrag $\neq 1$) und
 (iii) Addieren eines Vielfachen einer Komponente zu einer anderen (d.h. $A - \text{id}$ hat nur einen nicht-verschwindenden, ausser-diagonalen Eintrag)
 umfassen. Beachte, dass nur der Typ (ii) eine nicht-triviale Determinante gibt.
 · Zum Nachweis dieser elementaren Transformationen isolieren wir mit Hilfe von 21.6 die betroffene Koordinate und benutzen dann für (i) und (ii) die Translationsinvarianz des eindimensionalen Integrals, dass also $\int_{\mathbb{R}} f(x+b)dx = \int_{\mathbb{R}} f(x) \forall b \in \mathbb{R}$ sowie für (iii) die Gleichung (22.39). Beispiel:

$$\begin{aligned} \int_{\mathbb{R}^2} f(x_1 + 2x_2, x_2) d^2x &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x_1 + 2x_2, x_2) dx_1 \right) dx_2 \\ &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x_1, x_2) dx_1 \right) dx_2 = \int_{\mathbb{R}^2} f(x_1, x_2) d^2x \end{aligned} \quad (22.45)$$

□_{1. Schritt}

2. Schritt: $U \subset \mathbb{R}^n$ ist beliebig offen, Φ weiterhin affin linear (d.h. $\Upsilon = \Phi^{-1}(U)$) und $f \in \mathcal{C}_c(\mathbb{R}^n)$. (Hier soll betont werden, dass wir f über “kleine Teile” seines Definitionsbereichs integrieren wollen. Im Unterschied zum 3. Schritt setzen wir nicht voraus, dass $\text{supp}(f) \subset U$.)

· Nach den Überlegungen auf S. 159 (s. auch 22.16) gibt es eine Folge $(f_k)_{k \in \mathbb{N}} \subset \mathcal{C}_c(\mathbb{R}_x^n)$ mit $f_k \uparrow f \cdot \chi_U$ (wir nehmen an, dass $f \geq 0$) und (per Definition)

$$\int_U f(x) dx = \lim_{k \rightarrow \infty} \int_{\mathbb{R}^n} f_k(x) dx \quad (22.46)$$

Dann konvergiert die transformierte Folge $(f_k^\Phi)_{k \in \mathbb{N}} \subset \mathcal{C}_c(\mathbb{R}_\xi^n)$ monoton wachsend gegen $f^\Phi \cdot \chi_\Upsilon$, und da wegen des 1. Schritts

$$\lim_{k \rightarrow \infty} \int_{\mathbb{R}^n} f_k(x) dx = \lim_{k \rightarrow \infty} \int_{\mathbb{R}^n} f_k^\Phi(\xi) d\xi \quad (22.47)$$

folgt die Gleichung

$$\int_U f(x) dx = \int_\Upsilon f(A\xi + b) \cdot |\det A| d\xi \quad (22.48)$$

direkt aus der Definition des Integrals von f^Φ .

□_{2. Schritt}

Grösster Schritt: $\Phi : \Upsilon \rightarrow U$ ein beliebiger Diffeomorphismus, und $f \in \mathcal{C}_c(U, \mathbb{R})$, d.h. $f \in \mathcal{C}_c(\mathbb{R}^n)$ mit $\text{supp}(f) \subset U$. (Beachte hier den feinen Unterschied zum 2. Schritt.)

· Offenbar ist $f^\Phi \in \mathcal{C}_c(\Upsilon, \mathbb{R})$, und daher steht die Integrierbarkeit nicht zu Debatte. Daraufhin bemerken wir, dass die Gleichheit (22.40) bereits aus der Gültigkeit der Ungleichung

$$\int_U f(x) dx \geq \int_\Upsilon f^\Phi(\xi) d\xi \quad (22.49)$$

§ 22. EIGENSCHAFTEN DES LEBESGUE-INTEGRALS

für alle Diffeomorphismen zwischen offenen Mengen des \mathbb{R}^n folgt. Diese Ungleichung kontrolliert nämlich dann insbesondere die Transformation von f^Φ unter der Umkehrabbildung Φ^{-1} , so dass mit (22.43) die Kette

$$\int_U f(x)dx \geq \int_\Upsilon f^\Phi(\xi)d\xi \geq \int_{\Phi^{-1}(\Upsilon)} (f^\Phi)^{\Phi^{-1}}(x)dx = \int_U f(x)dx \quad (22.50)$$

resultiert, welche nur bei Gleichheit möglich ist. Wir beweisen zunächst eine lokale Abschwächung von (22.49), und dann ein topologisches Lemma.

Lemma 22.11. *Sei $\xi_* \in \Upsilon$ mit $f^\Phi(\xi_*) > 0$, und $q > 1$ beliebig fest. Dann existiert ein $r > 0$ so, dass für jeden offenen Würfel $W_l(\eta) = \{\xi \in \mathbb{R}^n \mid \|\xi - \eta\|_\infty < l\}$ (Mittelpunkt: η , Kantenlänge: $2l$) welcher ganz in $B_r(\xi_*) \subset \Upsilon$ enthalten ist²⁹*

$$\int_{W_l(\eta)} f^\Phi(\xi)d\xi \leq q^{n+3} \int_{\Phi(W_l(\eta))} f(x)dx \quad (22.51)$$

Beweis: Die Grundidee ist, Φ durch seine affine Approximation um η ,

$$\Phi_\eta(\xi) := \Phi(\eta) + D\Phi(\eta)(\xi - \eta) \quad (22.52)$$

zu ersetzen, und dafür (22.48) zu benutzen.

· Wir machen zunächst r so klein³⁰, dass $B_r(\xi_*) \subset \Upsilon$ und $f \circ \Phi > 0$ auf $B_r(\xi_*)$. Dann stellen wir fest, dass (22.52) als Funktion von $(\eta, \xi) \in B_r(\xi_*) \times B_r(\xi_*)$ stetig ist. Daher ist auch

$$\frac{f^{\Phi_\eta}(\xi)}{f^\Phi(\xi)} \quad (22.53)$$

stetig, und zwar mit Wert 1 auf der Diagonalen $\xi = \eta$. (Beachte, dass $\Phi_\eta(\eta) = \Phi(\eta)$.) Wir können daher durch Verkleinern von r erreichen, dass

$$f^{\Phi_\eta}(\xi) > q^{-1} \cdot f^\Phi(\xi) \quad \forall \xi, \eta \in B_r(\xi_*) \quad (22.54)$$

(Beachte, dass $q > 1$.) Wegen der Monotonie des Integrals erhalten wir daraus für alle $W_l(\eta) \subset B_r(\xi_*)$

$$\int_{W_l(\eta)} f^\Phi(\xi)d\xi \leq q \int_{W_l(\eta)} f^{\Phi_\eta}(\xi)d\xi \stackrel{(22.48)}{=} q \int_{\Phi_\eta(W_l(\eta))} f(x)dx \quad (22.55)$$

Beh.: Für r klein genug ist für alle $W_l(\eta) \subset B_r(\xi_*)$

$$\Phi_\eta(W_{\frac{l}{q}}(\eta)) \subset \Phi(W_l(\eta)) \quad (22.56)$$

(“Der um q geschrumpfte Würfel wird unter Φ_η nicht weiter als der ursprüngliche Würfel unter Φ abgebildet.”)

²⁹ $B_r(\xi_*)$ ist die offene Kugel in der beliebig gewählten Norm. Dies kann natürlich auch die max-Norm sein, die in die Definition der Würfel eingeht.

³⁰aber immer noch > 0 !

Bew.: Da Φ ein Diffeomorphismus ist, ist die Behauptung äquivalent zu

$$\Phi^{-1}(\Phi_\eta(W_{\frac{l}{q}}(\eta))) \subset W_l(\eta) \quad (22.57)$$

m.a.W.

$$\|\xi - \eta\|_\infty < \frac{l}{q} \Rightarrow \|\Phi^{-1}(\Phi_\eta(\xi)) - \eta\|_\infty < l \quad (22.58)$$

Nun ist aber (für festes η) $\Phi^{-1} \circ \Phi_\eta$ stetig differenzierbar mit $D(\Phi^{-1}\Phi_\eta)(\eta) = \text{id}$, und die Abbildung $(\eta, \xi) \mapsto D(\Phi^{-1}\Phi_\eta)(\xi)$ ist ebenfalls stetig. Daher ist für r klein genug

$$\|D(\Phi^{-1}\Phi_\eta)(\xi)\| < q \quad \forall \xi, \eta \in B_r(\xi_*) \quad (22.59)$$

(Operatornorm!) Aus der Konvexität von $W_l(\eta) \subset B_r(\xi_*)$ folgt mit Hilfe des Schrankeinsatzes 10.6, dass

$$\|\Phi^{-1}(\Phi_\eta(\xi)) - \eta\|_\infty = \|\Phi^{-1}(\Phi_\eta(\xi)) - \Phi^{-1}(\Phi_\eta(\eta))\|_\infty \leq q \cdot \|\xi - \eta\|_\infty \quad (22.60)$$

Dies zeigt (22.58).

· Nun ist ähnlich wie in (22.53)

$$\frac{f^\Phi(\xi)}{f^\Phi(\eta)} \quad (22.61)$$

stetig als Funktion von (η, ξ) , und gleich 1 auf der Diagonalen, so dass, gegebenenfalls wieder durch Verkleinern von r ,

$$q^{-1} \cdot f^\Phi(\eta) \leq f^\Phi(\xi) \leq q \cdot f^\Phi(\eta) \quad \forall \xi, \eta \in B_r(\xi_*) \quad (22.62)$$

Hieraus folgt für alle $W_l(\eta) \subset B_r(\xi_*)$

$$\int_{W_{\frac{l}{q}}(\eta)} f^\Phi(\xi) d\xi \geq q^{-1} \cdot f^\Phi(\eta) \cdot \underbrace{\text{vol}(W_{\frac{l}{q}}(\eta))}_{=q^{-n} \cdot \text{vol}(W_l(\eta))} \quad (22.63)$$

und
$$\int_{W_l(\eta)} f^\Phi(\xi) d\xi \leq q \cdot f^\Phi(\eta) \cdot \text{vol}(W_l(\eta))$$

Zusammen also

$$\begin{aligned} \int_{W_l(\eta)} f^\Phi(\xi) d\xi &\stackrel{(22.63)}{\leq} q^{n+2} \int_{W_{\frac{l}{q}}(\eta)} f^\Phi(\xi) d\xi \\ (22.55) \quad &\leq q^{n+3} \int_{\Phi_\eta(W_{\frac{l}{q}}(\eta))} f(x) dx \end{aligned} \quad (22.64)$$

$$(f \geq 0 \text{ und } (22.56)) \quad \leq q^{n+3} \int_{\Phi(W_l(\eta))} f(x) dx$$

wie behauptet. □ Lemma 22.11

Lemma 22.12. *Es sei $\Omega \subset \mathbb{R}^n$ offen und nichtleer. Dann existiert eine Folge W_1, W_2, W_3, \dots von paarweise disjunkten offenen (nicht-leeren) Würfeln mit der Eigenschaft, dass $\Omega = \overline{W_1} \cup \overline{W_2} \cup \overline{W_3} \cup \dots$.*

§ 22. EIGENSCHAFTEN DES LEBESGUE-INTEGRALS

Beweis: Wir betrachten die Menge der Würfel

$$\mathfrak{W} := \left\{ \frac{\prod_{i=1}^n (m_i, m_i + 1)}{2^s} \mid m_i \in \mathbb{Z}, s \in \mathbb{N} \right\} \quad (22.65)$$

die durch wiederholtes Halbieren des Standardgitters $\mathbb{Z}^n \subset \mathbb{R}^n$ entsteht. Dann gilt “offensichtlich”:

- (i) $W_1 \cap W_2 \neq \emptyset \Rightarrow \overline{W_1} \subset \overline{W_2}$ oder $\overline{W_2} \subset \overline{W_1}$, und
- (ii) für jeden Punkt $x \in \mathbb{R}^n$ und jedes $\epsilon > 0$ existiert ein Würfel $W \in \mathfrak{W}$ mit der Eigenschaft, dass $x \in \overline{W}$ und $\text{diam}(\overline{W}) < \epsilon$.

· Für gegebenes Ω wählen wir dann gemäss (ii) für jedes $x \in \Omega$ einen Würfel $W_x \in \mathfrak{W}$ mit $x \in \overline{W_x} \subset \Omega$. Dann eliminieren wir für gegebenes s iterativ alle Würfel, die gemäss (i) bereits in grösseren Würfeln enthalten sind. Der Grösse nach sortiert geben diese Würfel die gewünschte Überdeckung von Ω . □_{Lemma 22.12}

Beweis von (22.49): Angenommen, die Ungleichung wäre falsch. Dann existiert eine Zahl $q > 1$ so, dass noch

$$q^{n+3} \int_U f(x) dx < \int_{\Upsilon} f^\Phi(\xi) d\xi \quad (22.66)$$

was wir nun zu einem Widerspruch führen.

· Aus Stetigkeitsgründen ist $\Omega := \{\xi \in \mathbb{R}^n \mid f^\Phi(\xi) \neq 0\}$ offen und nicht-leer, da $f \neq 0$. Es sei dann $\Omega = \overline{W_1} \cup \overline{W_2} \cup \overline{W_3} \cup \dots$ eine Ausschöpfung wie in 22.12.

Beh.: Für mindestens einen Würfel $W^{(0)} \in \{W_i \mid i \in \mathbb{N}\}$ gilt

$$q^{n+3} \int_{\Phi(W^{(0)})} f(x) dx < \int_{W^{(0)}} f^\Phi(\xi) d\xi \quad (22.67)$$

Bew.: Andernfalls wäre unter Ausnutzung der Diffeomorphismeigenschaft von Φ :

$$\begin{aligned} q^{n+3} \int_U f(x) dx &= q^{n+3} \int_{\Phi(\Omega)} f(x) dx \\ (f \geq 0 \text{ und } 22.3) \quad &\geq q^{n+3} \sum_{i=1}^{\infty} \int_{\Phi(W_i)} f(x) dx \\ (falls (22.67) \text{ nicht gälte}) \quad &\geq \sum_{i=1}^{\infty} \int_{W_i} f^\Phi(\xi) d\xi \\ &= \int_{\Omega} f^\Phi(\xi) d\xi = \int_{\Upsilon} f^\Phi(\xi) d\xi \end{aligned} \quad (22.68)$$

im Widerspruch zu (22.66). (In den letzten beiden Schritten haben wir ausgenutzt, dass die Vereinigung der Würfelränder eine Nullmenge ist, und dass f ausserhalb von Ω verschwindet. Eine notorische Schwierigkeit beim Beweis des Transformationssatzes ist der Nachweis, dass diffeomorphe Bilder von Nullmengen wieder Nullmengen sind.³¹ Wir haben beim Übergang von der ersten auf die zweite Zeile statt-

³¹Die Topologie reicht nicht aus: Es gibt offene Mengen, deren Rand keine Nullmenge ist. Beispiel:

$$\bigcup_{k=1}^{\infty} \left(r_k - \frac{\delta}{2^k}, r_k + \frac{\delta}{2^k} \right) \subset (0, 1) \quad (22.69)$$

wobei (r_k) eine (geeignete) Abzählung von $\mathbb{Q} \cap (0, 1)$ ist. $\overline{U} = [0, 1]$, $\text{vol}(U) \leq \delta \Rightarrow \text{vol}(\partial U) \geq 1 - \delta$.

dessen nur die Disjunktheit der $\Phi(W_i)$ und Monotonie ausgenutzt.)

· Durch 2^n -teln und Wiederholen des Arguments erhalten wir eine absteigende Folge von nicht-leeren Würfeln $W^{(0)} \supset W^{(1)} \supset W^{(2)} \supset \dots$ mit $\overline{W^{(k)}} \subset \Omega \subset \Upsilon \forall k$ und

$$q^{n+3} \int_{\Phi(W^{(k)})} f(x) dx < \int_{W^{(k)}} f^\Phi(\xi) d\xi \quad (22.70)$$

Wegen $\text{diam } \overline{W^{(k)}} \rightarrow 0$ folgt aus dem Schachtelungsprinzip 4.16, dass

$$\bigcap_{k=0}^{\infty} \overline{W^{(k)}} = \{\xi_*\} \quad (22.71)$$

für ein $\xi_* \in \Omega$. Mit $f^\Phi(\xi_*) \neq 0$ können wir unser Lemma 22.11 anrufen. Für k gross genug ist $W^{(k)} \subset B_r(\xi_*)$, (und natürlich von der Form $W_l(\eta)$ für geeignetes l und η). Damit aber steht (22.70) im Widerspruch zu (22.51). $\square_{3. \text{ Schritt}}$

4. Schritt: Für U, Φ beliebig, $f \cdot \chi_U \in \mathcal{V}^+$ argumentiert man wie 2. Schritt. $\square_{4. \text{ Schritt}}$

5. Schritt: U, Φ und f beliebig. Die Mengen

$$\left\{ \int_U g(x) dx \mid g \in \mathcal{V}^+, g \geq f \cdot \chi_U \right\} \quad \text{und} \quad \left\{ \int_U h(x) dx \mid h \in \mathcal{V}^-, h \leq f \cdot \chi_U \right\} \quad (22.72)$$

werden nach dem 4. Schritt durch die Substitution $x = \Phi(\xi)$ offenbar bijektiv auf die Mengen

$$\left\{ \int_\Upsilon g^\Phi(\xi) d\xi \mid g^\Phi \in \mathcal{V}^+, g^\Phi \geq f^\Phi \cdot \chi_\Upsilon \right\} \quad (22.73)$$

$$\text{und} \quad \left\{ \int_\Upsilon h^\Phi(\xi) d\xi \mid h^\Phi \in \mathcal{V}^-, h^\Phi \leq f^\Phi \cdot \chi_\Upsilon \right\}$$

abgebildet. Damit folgt die Aussage unmittelbar aus der Definition des Integrals. $\square_{22.10}$

§ 23 Differentialformen und der Stokesche Satz

Bevor wir unser n -dimensionales Integral zur Normierung von Funktionenräumen benutzen, um damit weitere Beispiele von linearen und nicht-linearen Gleichungen in unendlich-dimensionalen Vektorräumen zu untersuchen, wollen wir in diesem § seine Bilanz bei der Umkehrung der gewöhnlichen Ableitung richtig stellen. Auf die folgenden noch bestehenden Unterschiede wurde in der einen oder anderen Form schon hingewiesen:

- (a) In der Transformationsformel 22.40 tritt der *Betrag* der Jacobi- (Funktional-)determinante auf, in der Substitutionsregel 22.34 hingegen die einfache Ableitung.
- (b) Der ein-dimensionale Hauptsatz verbindet in der Form

$$F' = f \quad \Rightarrow \quad \int_a^b f(s) ds = F(b) - F(a) \quad (23.1)$$

das bestimmte Integral über ein *kompaktes* Intervall mit der Auswertung der Stammfunktion an dessen Randpunkten. Wir haben uns bisher hauptsächlich mit der Einschränkung des n -dimensionalen Integrals auf *offene* Teilmengen beschäftigt, und

§ 23. DIFFERENTIALFORMEN UND DER STOKESSCHE SATZ

insbesondere über die Ränder allgemeiner abgeschlossener Mengen noch nicht viel gesagt. Ausserdem sind Integrale über “nieder-dimensionale” Teilmengen (z.B. entartete Quader) stets Null, was der intuitiven Aussage zu widersprechen scheint, das Integral berechne Volumina (orientiert oder nicht).

Ein wichtiger Hinweis zu den Hintergründen und der Auflösung dieser doch eher technischen Spannungen ist die folgende eher prinzipielle Beobachtung:

(c) Für eine Funktion $f : \mathbb{R} \rightarrow V$ (wie immer ist in diesem Zusammenhang V ein vollständiger normierter Vektorraum) ist ihre Ableitung wiederum eine Abbildung $f' : \mathbb{R} \rightarrow V$, so dass man in sinnvoller Weise auch umgekehrt nach der “Stammfunktion” $F : \mathbb{R} \rightarrow V$ von f fragen kann. In höheren Dimensionen ist hingegen das Differential einer Funktion $\mathbb{R}^n \rightarrow \mathbb{R}$ an jedem Punkt eine lineare Abbildung und hat als Funktion $\mathbb{R}^n \rightarrow \text{Hom}(\mathbb{R}^n, \mathbb{R})$ i.A. mehr Komponenten als f selbst.

Als erstes führen wir nun diejenige Klasse von Funktionen auf dem \mathbb{R}^n ein, die an den n -dimensionalen Integralsätzen partizipieren werden. Zur weiteren Motivation verweisen wir neben dem Auftreten der Determinante (vor dem Absolutbetrag...) in der Transformationsformel vor allem auf die Interpretation von alternierenden Multilinearformen als orientierte k -dimensionale Volumina (siehe Ende von § 18 ab S. 132).

Definition 23.1. Es sei $U \subset \mathbb{R}^n$ offen, $k \in \{0, 1, \dots, n\}$. Eine (stetige, differenzierbare, etc.) Differentialform vom Grad k (oder einfach eine k -Form) ist eine (stetige, differenzierbare, etc.) Abbildung

$$\omega : U \rightarrow \wedge^k(\mathbb{R}^n)^* \quad (23.2)$$

· Eine Differentialform ω tut also nichts anderes, als “an jedem Punkt” $x_0 \in U$ in linearer und alternierender Weise k Vektoren $v_1, \dots, v_k \in \mathbb{R}^n$ eine Zahl

$$\omega(x_0)(v_1, \dots, v_k) \in \mathbb{R} \quad (23.3)$$

zuzuweisen. Diese Zahl stellt eine Sorte “ k -dimensionales Volumen” des von (v_1, \dots, v_k) aufgespannten Parallelotops dar, physikalisch üblicherweise interpretiert als “Flussdichte”, in der Elektrodynamik etwa als Magnetfeld

· Formal gesehen handelt es sich nämlich bei dem $(\mathbb{R}^n)^*$ auf der rechten Seite von (23.2) um den Kotangentialraum $T_{x_0}^*U$ von U an x_0 (vgl. die Diskussion auf S. 66), und bei den Vektoren in (23.3) um Tangentialvektoren in $T_{x_0}U$. Das Parallelotop ist also als *infinitesimal und orientiert* zu denken.

· In diesem Zusammenhang ist es üblich, die Basis von $(\mathbb{R}^n)^*$, welche dual zur Standardbasis (e_1, \dots, e_n) des $\mathbb{R}^n \supset U$ ist, als (dx^1, \dots, dx^n) zu schreiben. Für eine k -Form ω existieren dann gemäss (19.37) $\binom{n}{k}$ (stetige, differenzierbare, etc.) Funktionen $\omega_{i_1, i_2, \dots, i_k} : U \rightarrow \mathbb{R}$ mit

$$\omega(x) = \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} \omega_{i_1, \dots, i_k}(x) \cdot dx^{i_1} \wedge \dots \wedge dx^{i_k} \quad \forall x \in U \quad (23.4)$$

$\omega_{i_1, \dots, i_k}(x)$ ist also die Flussdichte durch die von $(e_{i_1}, \dots, e_{i_k})$ aufgespannte Hyperfläche.

· Das euklidische Volumen in der sog. Standardorientierung des \mathbb{R}^n wird von der n -Form

$$\alpha := dx^1 \wedge \cdots \wedge dx^n \quad (23.5)$$

gemessen. Ist $\mathbb{R}^{n-1} \hookrightarrow \mathbb{R}^n$ ein $n - 1$ -dimensionaler Unterraum mit einem ausgezeichneten Vektor³² $\mathbf{n} \in \mathbb{R}^n \setminus \mathbb{R}^{n-1}$, so induziert man durch

$$\alpha_{\mathbf{n}}(v_1, \dots, v_{n-1}) = \alpha(\mathbf{n}, v_1, \dots, v_{n-1}) \quad (23.6)$$

eine Volumenform auf \mathbb{R}^{n-1} . Speziell für die Hyperebenen $\mathbb{R}_i^{n-1} = \{x^i = 0\}$, $\mathbf{n} = e_i$,

$$\alpha_i := (-1)^{i-1} dx^1 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^n \quad (23.7)$$

wobei $\widehat{\cdot}$ “auslassen” bedeutet.

· Im Spezialfall $k = 0$ ist eine Differentialform nichts anderes als eine gewöhnliche reellwertige Funktion $f : U \rightarrow \mathbb{R}$ (und weist als “Volumen” Punkten $x \in U$ einfach den Wert $f(x)$ zu). Ist f stetig differenzierbar, so ist das Differential von f , welches man in diesem Zusammenhang als $df := Df : U \rightarrow (\mathbb{R}^n)^*$ schreibt, dann bereits ein erstes Beispiel für eine 1-Form. Die Entwicklung (23.4) vereinfacht sich dabei zu:

$$df = \sum_{i=1}^n \frac{\partial f}{\partial x^i} dx^i \quad (23.8)$$

(speziell ist für $f(x) = x^i$, $df = dx^i$, was die Notation so nützlich macht...)

· Ein Vorteil der Differentialformen gegenüber anderen Tensorfeldern ist, dass sich ihre Ableitungen wieder zu Differentialformen zusammensetzen lassen, in Fortsetzung von (23.8) auf $k > 0$ und ohne Einführung weiterer Strukturen. Im Folgenden nehmen wir an, alle auftretenden Differentialformen seien (häufig genug) stetig differenzierbar, und schreiben $\Omega^k(U)$ für den zugehörigen Vektorraum. (Beachte: $\Omega^k(U)$ ist unendlich-dimensional, die Entwicklung (23.4) ist punktweise endlich.)

Definition 23.2 (Cartan-Ableitung). Die Cartan- (oder äussere) Ableitung einer k -Form $\omega \in \Omega^k(U)$ ist die $k + 1$ -Form

$$d\left(\sum_{i_1 < \cdots < i_k} \omega_{i_1, \dots, i_k} \cdot dx^{i_1} \wedge \cdots \wedge dx^{i_k}\right) := \sum_{i_1 < \cdots < i_k} d(\omega_{i_1, \dots, i_k}) \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k} \quad (23.9)$$

wobei auf der rechten Seite das punktweise Dachprodukt (19.40) steht. (Es gelten auch die Rechenregeln (19.42). Die Formel (19.41) muss man glücklicherweise praktisch nie benutzen.) Speziell für $k = 1$:

$$d\left(\sum_{i=1}^n \omega_i dx^i\right) = \sum_{i=1}^n \sum_{j=1}^n \left(\frac{\partial \omega_i}{\partial x^j} dx^j\right) \wedge dx^i = \sum_{i_1 < i_2} \left(\frac{\partial \omega_{i_2}}{\partial x^{i_1}} - \frac{\partial \omega_{i_1}}{\partial x^{i_2}}\right) dx^{i_1} \wedge dx^{i_2} \quad (23.10)$$

Lemma 23.3. (i) Für $\omega_1 \in \Omega^{k_1}(U)$, $\omega_2 \in \Omega^{k_2}(U)$ gilt die Leibniz-Regel

$$d(\omega_1 \wedge \omega_2) = (d\omega_1) \wedge \omega_2 + (-1)^{k_1} \omega_1 \wedge (d\omega_2) \quad (23.11)$$

(ii) $d^2 = 0$, d.h. für alle k -Formen $\omega \in \Omega^k(U)$ ist

$$d(d(\omega)) = 0 \in \Omega^{k+2}(U) \quad (23.12)$$

³² \mathbf{n} ist nicht “normal” im Sinne einer euklidischen Metrik auf \mathbb{R}^n .

§ 23. DIFFERENTIALFORMEN UND DER STOKESSCHE SATZ

Beweis. (i) folgt durch Anwendung der gewöhnlichen Produktregel

$$d(f_1 \cdot f_2) = df_1 \cdot f_2 + f_1 \cdot df_2 \quad (23.13)$$

(für $k_1 = k_2 = 0$ gilt $\wedge = \cdot$) auf die Koeffizientenfunktionen:

$$\begin{aligned} \omega_1 \wedge \omega_2 &= \\ &= \sum_{i_1, \dots, i_{k_1}} \sum_{j_1, \dots, j_{k_2}} (\omega_1)_{i_1, \dots, i_{k_1}} \cdot (\omega_2)_{j_1, \dots, j_{k_2}} \cdot dx^{i_1} \wedge \dots \wedge dx^{i_{k_1}} \wedge dx^{j_1} \wedge \dots \wedge dx^{j_{k_2}} \end{aligned} \quad (23.14)$$

und die Rechenregeln (19.42).

• Auch (ii) zeigen wir zunächst für $k = 0$. Mit (23.8) und (23.10) erhalten wir

$$d(df) = \sum_{i_1 < i_2} \left(\frac{\partial^2 f}{\partial x^{i_1} \partial x^{i_2}} - \frac{\partial^2 f}{\partial x^{i_2} \partial x^{i_1}} \right) dx^{i_1} \wedge dx^{i_2} = 0 \quad (23.15)$$

wegen der Symmetrie der zweiten Ableitungen 10.7. Insbesondere ist auch $d(dx^i) = 0$, so dass die Aussage für $k > 0$ durch Anwenden von (i) auf (23.9) folgt. \square

Beispiel 23.4. Die Rechenregeln für die vertrauten Vektordifferentialoperatoren in \mathbb{R}^3 lassen sich recht elegant in die Sprache der Differentialformen umformulieren. Der entscheidende Hinweis ist die Ähnlichkeit zwischen (23.10) und der expliziten Formel (10.35) für die Rotation eines Vektorfeldes. Für die Details müssen wir aber noch einmal kurz in die lineare Algebra ausholen.

• Zunächst einmal erkennen wir die musikalischen Isomorphismen $\cdot^b : T_{x_0}U \rightarrow T_{x_0}^*U$ und $\cdot^\sharp : T_{x_0}^*U \rightarrow T_{x_0}U$ als Spezialfälle von (20.5) bezüglich dem euklidischen Standardprodukt (20.6) auf \mathbb{R}^3 .

• Allgemein ist das für $\omega_1, \omega_2 \in \wedge^k(\mathbb{R}^n)^*$ durch

$$\langle \omega_1, \omega_2 \rangle := \sum_{i_1 < \dots < i_k} (\omega_1)_{i_1, \dots, i_k} (\omega_2)_{i_1, \dots, i_k} \quad (23.16)$$

definierte innere Produkt symmetrisch und positiv definit, induziert also einen Isomorphismus

$$\wedge^k(\mathbb{R}^n)^* \cong (\wedge^k(\mathbb{R}^n)^*)^* \quad (23.17)$$

• Andererseits wird das Dachprodukt

$$\begin{aligned} \wedge^k(\mathbb{R}^n)^* \times \wedge^{n-k}(\mathbb{R}^n)^* &\rightarrow \wedge^n(\mathbb{R}^n)^* \\ (\omega, \tilde{\omega}) &\mapsto \omega \wedge \tilde{\omega} \end{aligned} \quad (23.18)$$

durch die Identifikation

$$\begin{aligned} \wedge^n(\mathbb{R}^n)^* &\rightarrow \mathbb{R} \\ \alpha &= dx^1 \wedge \dots \wedge dx^n \mapsto 1 \end{aligned} \quad (23.19)$$

zu einer Bilinearform

$$\wedge^k(\mathbb{R}^n)^* \times \wedge^{n-k}(\mathbb{R}^n)^* \rightarrow \mathbb{R} \quad (23.20)$$

In den Basen (19.37) ist explizit

$$\begin{aligned} & \left((dx^{i_1} \wedge \cdots \wedge dx^{i_k}) \wedge (dx^{j_1} \wedge \cdots \wedge dx^{j_{n-k}}) \right) / dx^1 \wedge \cdots \wedge dx^n \\ &= \begin{cases} 0 & \text{falls } \{i_1, \dots, i_k\} \cup \{j_1, \dots, j_{n-k}\} \neq \{1, \dots, n\} \\ \pm 1 & \text{je nach Parität der Permutation} \\ & (i_1, \dots, i_k, j_1, \dots, j_{n-k}) = (\sigma(1), \dots, \sigma(n)) \text{ sonst} \end{cases} \end{aligned} \quad (23.21)$$

Offenbar ist also (23.20) nicht-entartet, und liefert daher einen Isomorphismus

$$(\wedge^k(\mathbb{R}^n)^*)^* \cong \wedge^{n-k}(\mathbb{R}^n)^* \quad (23.22)$$

Durch Zusammenschalten mit (23.17) erhalten wir den sogenannten Hodge-Stern-Operator

$$\star : \wedge^k(\mathbb{R}^n)^* \xrightarrow{\cong} \wedge^{n-k}(\mathbb{R}^n)^* \quad (23.23)$$

Zwei Formeln dafür:

$$\omega_1 \wedge (\star \omega_2) = \langle \omega_1, \omega_2 \rangle \cdot dx^1 \wedge \cdots \wedge dx^n \quad (23.24)$$

und

$$\begin{aligned} \star(dx^{i_1} \wedge \cdots \wedge dx^{i_k}) &= \pm 1 \cdot dx^{j_1} \wedge \cdots \wedge dx^{j_{n-k}} \\ &\text{je nach Parität der Permutation } (i_1, \dots, i_k, j_1, \dots, j_{n-k}) = (\sigma(1), \dots, \sigma(n)) \end{aligned} \quad (23.25)$$

· Zurück im \mathbb{R}^3 überprüft man nun, dass die abstrakte Formel (10.34) nichts anderes ist als

$$\text{rot}(Y) = \underbrace{(\star(d \underbrace{Y^b}_{\in \wedge^1}))}_{\in \wedge^1}^{\in \wedge^2} \in TU \quad (23.26)$$

Für (10.33) kann man schreiben

$$\text{div } Y = \star(d \underbrace{\star \underbrace{Y^b}_{\in \wedge^1}}_{\in \wedge^3}) \in \wedge^0(\mathbb{R}^3)^* = \mathbb{R} \quad (23.27)$$

Ausserdem ist (per Definition!)

$$(\text{grad } f)^b = df \quad (23.28)$$

Aus diesen drei Formeln folgt mit 23.3 (ii)

$$\begin{aligned} \text{rot} \circ \text{grad } f &= (\star d(df))^{\sharp} = 0 \\ \text{div} \circ \text{rot } Y &= \star d(\star \star dY^b) = 0 \end{aligned} \quad (23.29)$$

wobei wir noch $\star(\star(\omega)) = (-1)^{k(n-k)}\omega$ benutzt haben wie man am leichtesten aus (23.25) folgert.

§ 23. DIFFERENTIALFORMEN UND DER STOKESSISCHE SATZ

Als nächstes identifizieren wir eine Klasse von Mengen, die in den höherdimensionalen Integralsätzen die Rolle der kompakten Intervalle mit ihren Randpunkten übernehmen. Dazu verallgemeinern wir die Definition 20.10.

Fakten/Definition 23.5. Der *untere Halbraum* des \mathbb{R}^n ist die Teilmenge

$$\mathbb{R}_-^n := \{x \in \mathbb{R}^n \mid x^n \leq 0\} \quad (23.30)$$

Es sei $M \subset \mathbb{R}^n$ nicht leer und $p \in M$. Für eine offene Teilmenge $U \subset \mathbb{R}^n$ mit $p \in U$ heisst eine Abbildung

$$\varphi : U \cap M \rightarrow \mathbb{R}^k \quad (23.31)$$

eine *k-dimensionale (glatte) Karte mit Rand von M um p* (kurz: Karte) falls

(i) φ injektiv und stetig ist, eine offene Teilmenge $V \subset \mathbb{R}^k$ existiert so, dass $\varphi(U \cap M) = V \cap \mathbb{R}_-^k$,³³ sowie

(ii) die Umkehrabbildung $\psi := \varphi^{-1} : V \cap \mathbb{R}_-^k \rightarrow \mathbb{R}^n$ (unendlich oft) differenzierbar ist³⁴, mit $\text{rang}(D\psi(y)) = k \forall y \in V \cap \mathbb{R}_-^k$.

· Eine (glatte) *k-dimensionale Untermannigfaltigkeit mit Rand* ist eine nicht-leere Teilmenge $M \subset \mathbb{R}^n$ mit der Eigenschaft, dass für jeden Punkt $p \in M$ eine (glatte) *k-dimensionale Karte mit Rand von M um p* existiert.

· Die Umkehrabbildung $\psi = \varphi^{-1}$ heisst auch lokale (reguläre) Parametrisierung von M um p .

· Man nennt $p \in M$ einen (geometrischen) Randpunkt, falls $\varphi(p)^k = 0$, m.a.W. falls $\varphi(p) \in \partial\mathbb{R}_-^k = \mathbb{R}_k^{k-1}$, dem Rand des unteren Halbraumes. Beachte aber, dass p im Allgemeinen kein Randpunkt im n -dimensionalen Sinne von (21.15) ist. (Tatsächlich ist für $k < n$ der offene Kern einer Untermannigfaltigkeit leer.)

· Liegt eine Untermannigfaltigkeit erst einmal vor, so gibt es stets sehr viele Karten um jeden Punkt $p \in M$: Ist $\varphi : U \cap M \rightarrow V \cap \mathbb{R}_-^k$ eine Karte und $\Phi : V \cap \mathbb{R}_-^k \rightarrow \tilde{V} \cap \mathbb{R}_-^k$ ein Diffeomorphismus, so ist $\tilde{\varphi} := \Phi \circ \varphi : U \cap M \rightarrow \tilde{V} \cap \mathbb{R}_-^k$ eine weitere. Mit Hilfe der Sätze 12.1 bzw. 12.3 kann man umgekehrt zeigen, dass für zwei Karten (U_1, φ_1) , (U_2, φ_2) der “Kartenwechsel”

$$\varphi_2 \circ \varphi_1^{-1} : \varphi_1(U_1 \cap U_2 \cap M) \rightarrow \varphi_2(U_1 \cap U_2 \cap M) \quad (23.32)$$

ein Diffeomorphismus von Teilmengen des \mathbb{R}^k der Form $(\text{offene Menge}) \cap \mathbb{R}_-^k$ ist. Nun sind Diffeomorphismen insbesondere in beiden Richtungen stetig, und ein Punkt $x \in \mathbb{R}_-^k$ liegt in \mathbb{R}_k^{k-1} genau dann, wenn jede Umgebung von x sowohl Punkte in \mathbb{R}_-^k als auch Punkte ausserhalb von \mathbb{R}_-^k enthält. Daraus folgt, dass der Kartenwechsel (23.32) $\varphi_1(U_1 \cap U_2 \cap M) \cap \mathbb{R}_k^{k-1}$ bijektiv auf $\varphi_2(U_1 \cap U_2 \cap M) \cap \mathbb{R}_k^{k-1}$ abbildet. Insbesondere hängt die Definition eines Randpunktes nicht von der gewählten Karte ab. Ausserdem folgt daraus, dass auf dem Rand die Jacobi-Matrix des Kartenwechsels

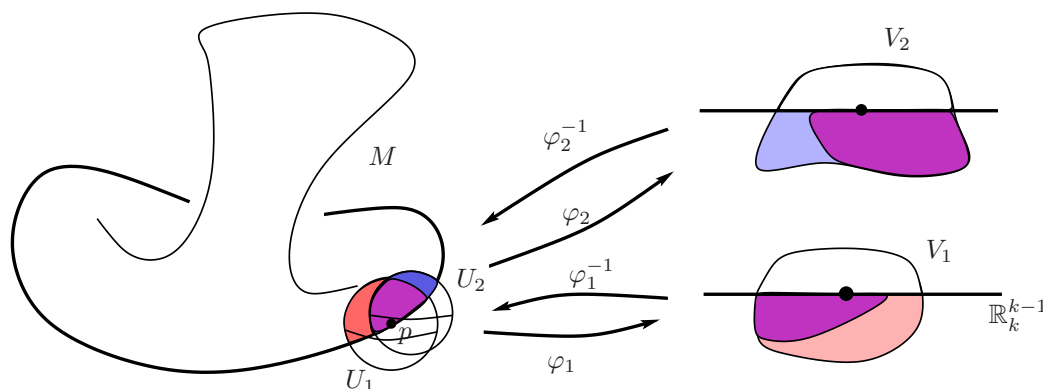
³³Für p im geometrischen Inneren erlauben wir hier auch $V \subset \mathbb{R}_-^k$. “Karte mit Rand” bedeutet also nicht, dass der Rand nicht leer ist, sorry. Das komische daran ist nur, dass ein linearer Unterraum $\mathbb{R}^k \subset \mathbb{R}^n$ sich als “Untermannigfaltigkeit mit (leerem) Rand nicht mit einer einzigen Karte überdecken lässt, als gewöhnliche Untermannigfaltigkeit hingegen schon.

³⁴Nach unserer Vereinbarung aus § 10 bedeutet dies, dass ψ die Einschränkung einer differenzierbaren Abbildung in einer offenen Umgebung von $V \cap \mathbb{R}_-^k$ ist.

die Block-Gestalt

$$D(\varphi_2 \circ \varphi_1^{-1})|_{\varphi_1(U_1 \cap U_2 \cap M) \cap \mathbb{R}_k^{k-1}} = \begin{pmatrix} \left(\frac{\partial(\varphi_2 \circ \varphi_1^{-1})^j}{\partial x^i} \right)_{i,j=1,\dots,k-1} & * \\ 0 & \frac{\partial(\varphi_2 \circ \varphi_1^{-1})^k}{\partial x^k} \end{pmatrix} \quad (23.33)$$

hat. (“Die k -te Komponente von φ_2 ist für $x^k = 0$ stets 0 und hängt daher nicht von (x^1, \dots, x^{k-1}) ab.”)



· Da $D(\varphi_2 \circ \varphi_1^{-1})$ überall Rang k hat, gilt dann auf dem Rand

$$\text{rang} \left(\frac{\partial(\varphi_2 \circ \varphi_1^{-1})^j}{\partial x^i} \right)_{i,j=1,\dots,k-1} \Big|_{\mathbb{R}_k^{k-1}} = k - 1 \quad \text{und} \quad \frac{\partial(\varphi_2 \circ \varphi_1^{-1})^k}{\partial x^k} \Big|_{\mathbb{R}_k^{k-1}} \neq 0 \quad (23.34)$$

Aus diesen Betrachtungen folgt (evtl. mit weiteren Zwischenschritten), dass die Menge aller Randpunkte einer glatten k -dimensionalen Untermannigfaltigkeit mit Rand (entweder leer oder) eine $k - 1$ -dimensionale Untermannigfaltigkeit mit leerem Rand ist. Wir schreiben ∂M für diesen Rand (auch wenn er nicht gleich dem topologischen Rand ist!). Ausserdem ist M ohne den geometrischen Rand eine k -dimensionale Untermannigfaltigkeit mit leerem Rand.

· Beispiele von Untermannigfaltigkeiten sind wie in 20.10 Graphen von differenzierbaren Abbildungen $f : V \cap \mathbb{R}_-^k \rightarrow \mathbb{R}^{n-k}$ sowie für eine glatte Abbildung $F : U \rightarrow \mathbb{R}^{n-k+1}$ ($U \subset \mathbb{R}^n$ offen) mit $\text{rang}(DF(p)) = n - k + 1$ die “Subniveaumengen” $F^{-1}(\{x = (x^k, 0, \dots, 0) \in \mathbb{R}^{n-k+1} \mid x^k \leq 0\})$. (Wir denken uns hier die Komponenten von $x \in \mathbb{R}^{n-k+1}$ als $(x^k, x^{k+1}, \dots, x^n)$ durchnummeriert.)

· Für jeden Punkt $p \in M$ können wir den Tangentialraum von M in p wie in (20.44) definieren. (Unsere lokalen Parametrisierungen stehen ja an den Randpunkten ein wenig in den oberen Halbraum über.) ($\psi = \varphi^{-1}$)

$$T_p M = D\psi(\varphi(p))(\mathbb{R}^k) = \{D\psi(\varphi(p))(w) \mid w \in \mathbb{R}^k\} \quad (23.35)$$

Ist $p \in \partial M$, so teilt der Tangentialraum an den Rand,

$$T_p \partial M = D\psi(\varphi(p))(\mathbb{R}_k^{k-1}) \quad (23.36)$$

den Tangentialraum an dem Rand in zwei Halbräume von “inneren” und “äusseren” Normalenvektoren, Richtungsvektoren von Kurven, die “nach M hinein” bzw.

§ 23. DIFFERENTIALFORMEN UND DER STOKESSCHE SATZ

“aus M hinaus” laufen. Weil diese Charakterisierung ebenfalls nicht von der Karte abhängt, gilt in (23.34) tatsächlich

$$\frac{\partial(\varphi_2 \circ \varphi_1^{-1})^k}{\partial x^k} \Big|_{\mathbb{R}_k^{k-1}} > 0 \quad (23.37)$$

Beispiele: Die n -dimensionale Kugel

$$K^n = \{x \in \mathbb{R}^n \mid \langle x, x \rangle \leq 1\} \quad (23.38)$$

ist eine glatte Untermannigfaltigkeit mit Rand S^{n-1} .

· Die Halbkugel

$$\{x \in \mathbb{R}^n \mid \langle x, x \rangle = 1, x^n \geq 0\} \quad (23.39)$$

ist eine $n - 1$ -dimensionale Untermannigfaltigkeit mit Rand S^{n-2} .

Das Wesen der Differentialgeometrie besteht darin, über Untermannigfaltigkeiten Aussagen zu machen, die “von der Wahl der Karten unabhängig ist”. (Noch allgemeiner, und insbesondere im Zusammenhang mit der Relativitätstheorie, interessiert man sich für Aussagen, die auch “von der Wahl der Einbettung in einen umgebenden \mathbb{R}^n ” unabhängig sind.) Dennoch oder gerade deshalb ist es sinnvoll, für spezielle Zwecke die Mannigfaltigkeit mit geeignet gewählten Karten, einem sog. *Atlas*, zu überdecken. Unser Ziel ist nun die Integration einer Differentialform $\omega \in \Omega^k(\mathfrak{U})$ über eine k -dimensionale Untermannigfaltigkeit mit Rand $M \subset \mathfrak{U}$, mit $\mathfrak{U} \subset \mathbb{R}^n$ offen. Wir arbeiten dazu im Folgenden mit der euklidischen Norm $\|\cdot\|_2$ auf \mathbb{R}^n , und machen noch zwei weitere Annahmen.

Zerlegung der Eins

· Aus (23.31) erhält man durch Verkleinern von U für jedes $p \in M$ eine Karte der Form

$$\varphi_p : B_{2r_p}(p) \rightarrow V_p \cap \mathbb{R}_-^k \quad (23.40)$$

für einen geeignet gewählten Radius $r_p > 0$ so, dass noch $B_{2r_p}(p) \subset \mathfrak{U}$.

· Nun überdecken die offenen Mengen

$$\{B_{r_p}(p) \mid p \in M\} \quad (23.41)$$

trivialerweise ganz M (d.h. jedes $p \in M$ ist in einer dieser Mengen enthalten).

Nehmen wir dann an, dass M kompakt ist, so folgt,³⁵ dass bereits endlich viele der $B_{r_p}(p)$ M vollständig überdecken. M.a.W. existieren endlich viele Punkte $p_\iota \in M$ und Radien $r_\iota > 0$, $\iota = 1, \dots, \varsigma$ so, dass

$$M \subset \bigcup_{\iota=1}^{\varsigma} B_{r_\iota}(p_\iota) \quad (23.42)$$

· Es sei nun $h : [0, \infty) \rightarrow \mathbb{R}$ eine glatte Funktion mit der Eigenschaft, dass

(i) $1 \geq h(t) \geq 0 \forall t \in [0, \infty)$

³⁵Dies ist eine Konsequenz des Satzes von Heine-Borel, vgl. auch den Begriff der “Überdeckungskompaktheit”, den wir in den Übungen kennengelernt haben.

- (ii) $h(t) = 1$ für $t \in [0, 1)$, und
- (iii) $\text{supp}(h) \subset [0, 2)$.

Für eine mögliche Konstruktion von h sei $u(t) := \exp(-1/t)$ für $t > 0$, $u(t) = 0$ für $t \leq 0$. Dann ist u unendlich oft differenzierbar mit $u(0) = 0$, $u(t) > 0$ für $t > 0$. Daher ist $v(t) := u(t-1) \cdot u(3/2-t)$ ebenfalls glatt mit erfüllt $v \geq 0$ und $\text{supp}(v) = [1, 3/2]$. Setzt man dann

$$h(t) := 1 - \frac{\int_0^t v(s) ds}{\int_0^{3/2} v(s) ds}, \tag{23.43}$$

so erfüllt h die gewünschten Eigenschaften.

Bemerkung: Es ist kein Zufall, dass in dieser Konstruktion u nicht überall analytisch ist, nämlich um 0 herum keine Potenzreihendarstellung besitzt: Die obigen Bedingungen implizieren, dass die Nullstellen von h Häufungspunkte besitzen, was für eine nicht-triviale analytische Funktion auf einer zusammenhängenden Menge nicht möglich ist, siehe § 5 und § 11.

· Damit ist für jedes $\iota = 1, \dots, \varsigma$ die Funktion (glatter Hut)

$$\begin{aligned} \hat{h}_\iota &: \mathbb{R}^n \rightarrow \mathbb{R}, \\ \hat{h}_\iota(y) &:= h\left(\frac{\|y - p_\iota\|_2}{r_\iota}\right) \end{aligned} \tag{23.44}$$

unendlich oft differenzierbar (die euklidische Norm ist es ja zumindestens ausserhalb von 0, dort ist aber ja h konstant...), und erfüllt

- (i) $1 \geq \hat{h}_\iota \geq 0$
- (ii) $\hat{h}_\iota(y) = 1$ für $y \in B_{r_\iota}(p_\iota)$
- (iii) $\text{supp}(\hat{h}_\iota) \subset B_{2r_\iota}(p_\iota)$.

Wegen $\hat{h}_\iota > 0$ auf $B_{r_\iota}(p_\iota)$ verschwindet $\sum_{\iota=1}^{\varsigma} \hat{h}_\iota$ nirgends auf $\mathfrak{V} := \bigcup_{\iota=1}^{\varsigma} B_{r_\iota}(p_\iota) \subset \mathfrak{U}$

Daher gilt für die auf \mathfrak{V} definierten Funktionen

$$h_\iota := \frac{\hat{h}_\iota}{\sum_{\iota=1}^{\varsigma} \hat{h}_\iota} \tag{23.45}$$

$\sum_\iota h_\iota = 1$, und $\text{supp}(h_\iota) \cap M \subset B_{2r_\iota}(p_\iota)$ ist kompakt. Man nennt so eine Sammlung von Funktionen eine “der Überdeckung (23.42) untergeordnete Teilung der Eins”.

· Der Rand einer kompakten Untermannigfaltigkeit mit Rand ist abgeschlossen und daher wieder kompakt, und durch Einschränkung erhalten wir aus $(B_{r_\iota}(p_\iota))$ und (h_ι) einen endlichen Atlas mit zugehöriger Teilung der Eins auch auf dem Rand von M .

Rückzug

Wir nehmen nun weiter an, der endliche Atlas

$$\mathfrak{A} = \{(B_{2r_\iota}(p_\iota), \varphi_\iota) \mid \iota = 1, \dots, \varsigma\} \tag{23.46}$$

könne so gewählt werden, dass für alle Kartenwechsel

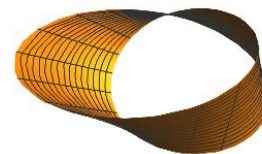
$$\det D(\varphi_\iota \circ \varphi_\kappa^{-1})(x) > 0 \quad \forall x \in \varphi_\kappa(B_{2r_\iota}(p_\iota) \cap B_{2r_\kappa}(p_\kappa)) \cap M \tag{23.47}$$

§ 23. DIFFERENTIALFORMEN UND DER STOKESSCHE SATZ

Man nennt Untermannigfaltigkeiten, für die eine solche Wahl möglich ist, **orientierbar**. Nach einer solchen Wahl zerfallen die Karten von M in zwei Orientierungsklassen, je nach dem Vorzeichen der Funktionaldeterminante, welche sich durch Umkehren einer Koordinate, $x_{k-1} \mapsto -x_{k-1}$ ineinander überführen lassen. (Stimmt nicht ganz, bei nicht zusammenhängenden Kartengebieten, und bei $k = 1 \dots$), Dabei ist es nicht möglich zu sagen, welche der beiden Orientierungen positiv bzw. negativ ist. Wichtig ist allerdings, dass man wegen (23.33) und (23.37) durch Einschränken eines orientierten Atlases auf den Rand von M eine Orientierung auf ∂M induziert (aus $\det(D\varphi_\iota \circ \varphi_\kappa^{-1}) > 0$ und (23.37) folgt $\det\left(\frac{\partial(\varphi_\iota \circ \varphi_\kappa^{-1})^j}{\partial x^i}\right)_{i,j=1,\dots,k-1} > 0$ am Rand). Wir schreiben für unsere Untermannigfaltigkeit fortan M^\uparrow , um uns daran zu erinnern, dass wir eine Orientierung gewählt haben, und für ihren Rand mit kanonisch induzierter Orientierung ∂M^\uparrow .

Beispiel: Der Möbius-Streifen, Bild der Abbildung

$$[0, 2\pi] \times [-1/2, 1/2] \ni (\vartheta, s) \mapsto \begin{pmatrix} \cos \vartheta \left(1 - s \sin \frac{\vartheta}{2}\right) \\ \sin \vartheta \left(1 - s \sin \frac{\vartheta}{2}\right) \\ s \cos \frac{\vartheta}{2} \end{pmatrix} \quad (23.48)$$



ist eine nicht-orientierbare kompakte 2-dimensionale Untermannigfaltigkeit des \mathbb{R}^3 mit Rand.

· Es sei nun $\omega \in \Omega^l(\mathfrak{U})$ eine l -Form auf \mathfrak{U} (gleich interessieren uns vor allem $l = k$ und $k-1$). Für jede Karte $\varphi : U \cap M^\uparrow \xrightarrow{\cong} V \cap \mathbb{R}_-^k$ von M^\uparrow erhalten wir dann durch *Rückzug* eine l -Form auf $V \cap \mathbb{R}_-^k$ wie folgt: Für $x \in V \cap \mathbb{R}_-^k$ und $v_1, \dots, v_l \in T_x(V) \cong \mathbb{R}^k$ sind (mit $\psi = \varphi^{-1}$) $D\psi(x)(v_i) \in T_{\psi(x)}M^\uparrow \subset T_{\psi(x)}\mathbb{R}^n$ und wir setzen

$$(\psi^*(\omega)(x))(v_1, \dots, v_l) := \omega(\psi(x))(D\psi(x)(v_1), \dots, D\psi(x)(v_l)) \quad (23.49)$$

(Dies ist klarerweise eine alternierende l -Form auf T_xV , die Differenzierbarkeit folgt aus Kettenregeln, d.h. $\psi^*(\omega) \in \Omega^l(V \cap \mathbb{R}_-^k)$.)

Beispiel: Für eine 0-Form, d.h. eine Funktion $f : \mathfrak{U} \rightarrow \mathbb{R}$ ist $\psi^*(f) = f \circ \psi$. Für ihr Differential gilt ($v \in T_xV$):

$$\begin{aligned} (\psi^*(df)(x))(v) &= \sum_{j=1}^n \frac{\partial f}{\partial y^j}(\psi(x)) \overbrace{\sum_{i=1}^k \frac{\partial \psi^j}{\partial x^i}(x) v^i}^{(D\psi(x)(v))^j} \\ &= \sum_{i=1}^k \left(\sum_{j=1}^n \frac{\partial f}{\partial y^j}(\psi(x)) \frac{\partial \psi^j}{\partial x^i}(x) \right) v^i \\ \text{(Kettenregel)} \quad &= \frac{\partial(f \circ \psi)}{\partial x^i}(x) v^i \\ &= d(\psi^*(f))(v) \end{aligned} \quad (23.50)$$

was man unschwer auf $l > 1$ verallgemeinert:

$$\psi^*(d\omega) = d(\psi^*(\omega)) \quad (23.51)$$

Hierbei operiert das d auf der rechten Seite im \mathbb{R}^k , das auf der linken Seite im \mathbb{R}^n , es sind aber nur die Ableitungen in den Tangentialrichtungen von M^\uparrow relevant. Diese Verträglichkeit mit dem Rückzug ist eine weitere wichtige Eigenschaft der äusseren Ableitung.

· Ausserdem ist der Rückzug offenbar linear und verträglich mit dem Dachprodukt. Insbesondere gilt

$$\psi^*(\omega) = \psi^*\left(\sum_{\iota} h_{\iota}\omega\right) = \sum_{\iota} \psi^*(h_{\iota} \cdot \omega) \quad (23.52)$$

für jede Karte $(U, \varphi = \psi^{-1})$ und jede l -Form $\omega \in \Omega^l(\mathfrak{U})$.

· Betrachten wir speziell für $k = l$ eine der Karten $(B_{2r_{\iota}}(p_{\iota}), \varphi_{\iota})$ in unserem orientierten Atlas (23.46), so ist mit der Teilung der Eins (23.45) (und $\psi_{\iota} = \varphi_{\iota}^{-1}$)

$$\psi_{\iota}^*(h_{\iota} \cdot \omega) \in \Omega_c^k(V_{\iota} \cap \mathbb{R}_{-}^k) \quad (23.53)$$

eine k -Form mit kompaktem Träger. Wegen $\dim \wedge^k(\mathbb{R}^k)^* = 1$ existiert eine Funktion $\omega_{\iota} \in \mathcal{C}_c^{\infty}(V_{\iota} \cap \mathbb{R}_{-}^k)$, so dass

$$\psi_{\iota}^*(h_{\iota} \cdot \omega) = \omega_{\iota}(x) \cdot dx^1 \wedge \cdots \wedge dx^k =: \omega_{\iota}(x) \cdot \alpha \quad (23.54)$$

Mit anderen Worten ist

$$\omega_{\iota}(x) = h_{\iota}(\psi_{\iota}(x)) \cdot \omega(\psi_{\iota}(x))(D\psi_{\iota}(x)(e_1), \dots, D\psi_{\iota}(x)(e_k)) \quad (23.55)$$

wobei (e_1, \dots, e_k) die Standardbasis des \mathbb{R}^k ist. Wir definieren dann das *Integral* der k -Form $\omega \in \Omega^k(\mathfrak{U})$ über die orientierte kompakte k -dimensionale Untermannigfaltigkeit mit Rand M^\uparrow durch die Formel:

$$\int_{M^\uparrow} \omega := \sum_{\iota} \int_{\mathbb{R}_{-}^k} \omega_{\iota}(x) dx = \sum_{\iota} \int_{\mathbb{R}^k} \chi_{\mathbb{R}_{-}^k}(x) \omega_{\iota}(x) dx \quad (23.56)$$

wobei auf der rechten Seite das gewöhnliche Integral für stetige Funktionen mit kompaktem Träger im \mathbb{R}^k steht. Diese Definition ist sinnvoll wegen des folgenden Lemmas:

Lemma 23.6. *Der Wert des Integrals (23.56) ist unabhängig von dem orientierten Atlas und der Teilung der Eins.*

Beweis. Es sei $\tilde{\mathfrak{A}} = \{(\tilde{U}_{\kappa}, \tilde{\varphi}_{\kappa} = \tilde{\psi}_{\kappa}^{-1})\}$ ein weiterer endlicher Atlas von M^\uparrow mit $\tilde{U}_{\kappa} \subset \mathfrak{U}$ und der gleichen Orientierung wie \mathfrak{A} , d.h.

$$\det D(\varphi_{\iota} \circ \tilde{\varphi}_{\kappa}^{-1}) > 0 \quad \text{auf } \tilde{\varphi}_{\kappa}(\tilde{U}_{\kappa} \cap B_{2r_{\iota}}(p_{\iota}) \cap M^\uparrow) \quad \forall \iota, \kappa, \quad (23.57)$$

und (\tilde{h}_{κ}) eine untergeordnete Teilung der Eins, d.h.

(i) $1 \geq \tilde{h}_{\kappa} \geq 0$

(ii) $\text{supp}(\tilde{h}_{\kappa}) \subset \tilde{U}_{\kappa}$ ist kompakt.

(iii) $\sum \tilde{h}_{\kappa} = 1$ in einer Umgebung von M^\uparrow .

Dann gilt wegen der Linearität des Integrals

$$\begin{aligned} \sum_{\iota} \int_{\mathbb{R}_{-}^k} \omega_{\iota}(x) dx &= \sum_{\iota} \int_{\mathbb{R}_{-}^k} \sum_{\kappa} \tilde{h}_{\kappa}(\psi_{\iota}(x)) \cdot \omega_{\iota}(x) dx \\ &= \sum_{\iota} \sum_{\kappa} \int_{\varphi_{\iota}(B_{2r_{\iota}}(p_{\iota}) \cap \tilde{U}_{\kappa} \cap M^\uparrow)} \tilde{h}_{\kappa}(\psi_{\iota}(x)) \cdot \omega_{\iota}(x) dx \end{aligned}$$

§ 23. DIFFERENTIALFORMEN UND DER STOKESSCHE SATZ

Durch die Substitution $x = (\varphi_l \circ \tilde{\varphi}_\kappa^{-1})(\tilde{x})$ (d.h. $\psi_l(x) = \tilde{\psi}_\kappa(\tilde{x}) \in M^\uparrow$) wird daraus mit Hilfe der Transformationsformel 22.10 und Gl. (23.57) einerseits

$$= \sum_\kappa \sum_l \int_{\tilde{\varphi}_\kappa(\tilde{U}_\kappa \cap B_{2r_l}(p_l) \cap M^\uparrow)} \tilde{h}_\kappa(\tilde{\psi}_\kappa(\tilde{x})) \cdot \omega_l(x) \cdot \det D(\varphi_l \circ \tilde{\varphi}_\kappa^{-1})(\tilde{x}) d\tilde{x} \quad (23.58)$$

Wegen

$$D\psi_l(x) = D(\tilde{\psi}_\kappa \circ \tilde{\varphi}_\kappa \circ \varphi_l^{-1})(x) = D(\tilde{\psi}_\kappa)(\tilde{x}) \cdot D(\tilde{\varphi}_\kappa \circ \varphi_l^{-1})(x) \quad (23.59)$$

folgt aus (23.55) andererseits

$$\begin{aligned} \tilde{h}_\kappa(\tilde{\psi}_\kappa(\tilde{x})) \cdot \omega_l(x) &= \tilde{h}_\kappa(\tilde{\psi}_\kappa(\tilde{x})) \cdot h_l(\psi_l(x)) \cdot \omega(\psi_l(x))(D\psi_l(x)(e_1), \dots, D\psi_l(x)(e_k)) \\ &= h_l \cdot \underbrace{\tilde{h}_\kappa \cdot \omega(\tilde{\psi}_\kappa(\tilde{x}))(D\tilde{\psi}_\kappa(\tilde{x})(e_1), \dots, D\tilde{\psi}_\kappa(\tilde{x})(e_k))}_{=\tilde{\omega}_\kappa(\tilde{x})} \cdot \det D(\tilde{\varphi}_\kappa \circ \varphi_l^{-1})(x) \end{aligned} \quad (23.60)$$

wobei wir die Transformation (19.45) einer Volumenform unter einem Endomorphismus des \mathbb{R}^k benutzt haben.³⁶

• Nun heben sich wegen der Kettenregel, $D(\tilde{\varphi}_\kappa \circ \varphi_l^{-1}) = (D(\varphi_l \circ \tilde{\varphi}_\kappa^{-1}))^{-1}$, die Determinanten gegenseitig auf, so dass zurück in (23.58)

$$\begin{aligned} \sum_l \int_{\mathbb{R}_-^k} \omega_l(x) dx &= \sum_\kappa \sum_l \int_{\tilde{\varphi}_\kappa(\tilde{U}_\kappa \cap B_{2r_l}(p_l) \cap M^\uparrow)} h_l(\tilde{\psi}_\kappa(\tilde{x})) \cdot \tilde{\omega}_\kappa(\tilde{x}) d\tilde{x} \\ &= \sum_\kappa \int_{\mathbb{R}_-^k} \underbrace{\sum_l h_l(\tilde{x})}_{=1} \cdot \tilde{\omega}_\kappa(\tilde{x}) d\tilde{x} = \sum_\kappa \int_{\mathbb{R}_-^k} \tilde{\omega}_\kappa(\tilde{x}) d\tilde{x} \end{aligned} \quad (23.62)$$

wie behauptet. □

• Wir schreiben ab jetzt auch

$$\int_{\mathbb{R}_-^k} \psi_l^*(h_l \cdot \omega) \quad \text{für} \quad \int_{\mathbb{R}_-^k} \omega_l(x) dx \quad (23.63)$$

³⁶Um hier noch einmal explizit zu rechnen sei (A_i^j) die Matrixdarstellung von $D(\tilde{\varphi}_\kappa \circ \varphi_l^{-1})(x)$ (d.h. $D(\tilde{\varphi}_\kappa \circ \varphi_l^{-1})(x)(e_i) = \sum_j e_j A_i^j$). Dann ist mit (23.59)

$$\omega(D\psi_l(e_1), \dots, D\psi_l(e_k)) = \sum_{j_1, \dots, j_k=1}^k \omega(D\tilde{\psi}_\kappa(e_{j_1})A_1^{j_1}, \dots, D\tilde{\psi}_\kappa(e_{j_k})A_k^{j_k})$$

Da ω alternierend ist, steht auf der rechten Seite Null ausser falls $j_i = \sigma(i)$ für eine Permutation $\sigma \in S_k$, und zwar wegen der Multilinearität genau

$$= \omega(D\tilde{\psi}_\kappa(e_1), \dots, D\tilde{\psi}_\kappa(e_k)) \cdot \underbrace{\sum_{\sigma \in S_k} \text{sgn}(\sigma) \prod_{j=1}^k A_{\sigma^{-1}(j)}^j}_{=\det A} \quad (23.61)$$

nach der Leibniz-Formel für die Entwicklung der Determinante.

· Für $n = 1$ lässt sich nun (etwas umständlich) das bestimmte Integral

$$\int_a^b f(s)ds = \int_{[a,b]} f(x)dx \quad (23.64)$$

wiedererkennen als das Integral der 1-Form $f(s)ds$ auf dem offenen Intervall $I \supset [a, b]$ über die kompakte Untermannigfaltigkeit mit Rand $[a, b]$ in der durch die Anordnung von \mathbb{R} gegebenen Orientierung und der Standardparametrisierung $s = x$. Eine Umparametrisierung des Intervalls ist eine bijektive Abbildung $\tau : J \rightarrow I$, welche je nachdem $\tau' \geq 0$ orientierungserhaltend oder -umkehrend ist. Eine Stammfunktion ist eine 0-Form F mit $dF(s) = f(s)ds$, die in der auf dem Rand induzierten Orientierung zu $F(a) - F(b)$ integriert wird. Für eine Differentialform ω vom Grad k ist eine "Stammform" eine $k - 1$ -Form η mit $d\eta = \omega$, so dass die Verallgemeinerung der einen Richtung des Hauptsatzes (23.1) lautet:

Theorem 23.7 (Stokesscher Satz). *Es sei $\mathfrak{U} \subset \mathbb{R}^n$ offen und $M^\uparrow \subset \mathfrak{U}$ eine orientierte kompakte k -dimensionale Untermannigfaltigkeit mit Rand. Dann gilt für jede $k - 1$ -Form $\eta \in \Omega^{k-1}(\mathfrak{U})$*

$$\int_{M^\uparrow} d\eta = \int_{\partial M^\uparrow} \eta \quad (23.65)$$

Beweis. Es sei $\{(B_{2r_\iota}(p_\iota), \varphi_\iota = \psi_\iota^{-1})\}$ ein orientierter Atlas der obigen Art und (h_ι) eine untergeordnete Teilung der Eins. Dann erhält man durch geeignetes Einschränken Gleiches auf ∂M^\uparrow . (Wobei hier die p_ι nicht auf ∂M^\uparrow liegen müssen, und es auch mal sein kann, dass $\text{supp}(h_\iota) \cap \partial M^\uparrow = \emptyset$.) Da ∂M^\uparrow leeren Rand hat, lautet die Behauptung zunächst

$$\sum_\iota \int_{\mathbb{R}_-^k} \psi_\iota^*(h_\iota \cdot d\eta) = \sum_\iota \int_{\mathbb{R}_-^{k-1}} \psi_\iota^*(h_\iota \cdot \eta) \quad (23.66)$$

Wegen der Produktregel 23.3 und der Verträglichkeit (23.51) gilt auf der linken Seite

$$\begin{aligned} \sum_\iota \int_{\mathbb{R}_-^k} \psi_\iota^*(h_\iota \cdot d\eta) &= \sum_\iota \int_{\mathbb{R}_-^k} \psi_\iota^*(d(h_\iota \cdot \eta)) - \sum_\iota \int_{\mathbb{R}_-^k} \psi_\iota^*(dh_\iota \wedge \eta) \\ &= \sum_\iota \int_{\mathbb{R}_-^k} d(\psi_\iota^*(h_\iota \cdot \eta)) - \underbrace{\int_{M^\uparrow} d\left(\underbrace{\sum_\iota h_\iota}_{=1}\right) \wedge \eta}_{=0} \end{aligned} \quad (23.67)$$

Es genügt daher zu zeigen, dass in jeder Karte

$$\int_{\mathbb{R}_-^k} d(\psi_\iota^*(h_\iota \cdot \eta)) = \int_{\mathbb{R}_-^{k-1}} \psi_\iota^*(h_\iota \cdot \eta) \quad (23.68)$$

Für die $k - 1$ -Formen existieren Funktionen $\eta_\iota^i \in \mathcal{C}_c^\infty(V_\iota \cap \mathbb{R}_-^k)$ so dass wir mit $\alpha_i = (-1)^{i-1} dx^1 \wedge \dots \wedge \widehat{dx^i} \wedge \dots \wedge dx^k$ in $V_\iota \cap \mathbb{R}_-^k$ schreiben können:

$$\psi_\iota^*(h_\iota \cdot \eta) = \sum_{i=1}^k \eta_\iota^i \alpha_i \quad (23.69)$$

§ 23. DIFFERENTIALFORMEN UND DER STOKESSCHE SATZ

Es gilt dann

$$d(\psi_l^*(h_l \cdot \eta)) = \sum_{i=1}^k (d\eta_l^i) \wedge \alpha_i = \sum_{i=1}^k \frac{\partial \eta_l^i}{\partial x^i} \cdot \alpha \quad (23.70)$$

so dass mit Hilfe von Prop. 21.6

$$\begin{aligned} \int_{\mathbb{R}_-^k} d(\psi_l^*(h_l \cdot \eta)) &= \sum_{i < k} \int_{\mathbb{R}_-^{k-1}} \overbrace{\left(\int_{-\infty}^{\infty} \frac{\partial \eta_l^i}{\partial x^i}(x^1, \dots, x^k) dx^i \right)}{=0, \text{ da } \text{supp}(\eta_l^i) \text{ kompakt}} d^{k-1}x \\ &\quad + \int_{\mathbb{R}_-^{k-1}} \left(\int_{-\infty}^0 \frac{\partial \eta_l^k}{\partial x^k} dx^k \right) d^{k-1}x \quad (23.71) \\ &= \int_{\mathbb{R}_k^{k-1}} \eta_l^k(x^1, \dots, x^{k-1}, 0) d^{k-1}x \end{aligned}$$

nach dem ein-dimensionalen Hauptsatz, angewendet auf $\eta_l^k(x^1, \dots, x^{k-1}, \cdot) \in \mathcal{C}_c^\infty((-\infty, 0])$. Andererseits ist wegen $x^k = 0 = \text{konst.}$ auf \mathbb{R}_k^{k-1} , $\alpha_i|_{\mathbb{R}_k^{k-1}} = 0$ ausser für $i = k$. Daher gilt auch

$$\int_{\mathbb{R}_k^{k-1}} \psi_l^*(h_l \cdot \eta) = \int_{\mathbb{R}_k^{k-1}} \eta_l^k(x^1, \dots, x^{k-1}, 0) d^{k-1}x \quad (23.72)$$

Zusammen ist dies die Behauptung. □

Für die andere Richtung des Hauptsatzes (Konstruktion einer Stammform durch Integration) benötigen wir eine Annahme an die Topologie des Definitionsbereichs. Wie in der Funktionentheorie § 15 betrachten wir wenigstens Anfangs als natürliche Verallgemeinerung von offenen Intervallen offene und konvexe Mengen $U \subset \mathbb{R}^n$, d.h. $\forall a, b \in U$ liegt auch $\{a + t(b - a) \mid t \in [0, 1]\} \subset U$. Ausserdem ist klar, dass wegen $d^2 = 0$ (Lemma 23.3) für $\omega \in \Omega^k(U)$ höchstens dann ein $\eta \in \Omega^{k-1}(U)$ mit $d\eta = \omega$ existieren kann, wenn $d\omega = 0$. Zuvor noch ein algebraisches Resultat:

Lemma/Definition 23.8. Die Kontraktion (oder Paarung) von alternierenden Multilinearformen mit Vektoren ist die bilineare Abbildung

$$\begin{aligned} \iota_v : \mathbb{R}^n \times \wedge^k(\mathbb{R}^n)^* &\rightarrow \wedge^{k-1}(\mathbb{R}^n)^* \\ (\iota_v(\omega))(v_1, \dots, v_{k-1}) &:= \omega(v, v_1, \dots, v_{k-1}) \end{aligned} \quad (23.73)$$

(vgl. (23.6)). In den Standardbasen (e_1, \dots, e_n) und $\{dx^{i_1} \wedge \dots \wedge dx^{i_k} \mid i_1 < \dots < i_k\}$ ist

$$\iota_{e_i}(dx^{i_1} \wedge \dots \wedge dx^{i_k}) = \begin{cases} 0 & \text{falls } i \notin \{i_1, \dots, i_k\} \\ (-1)^r dx^{i_1} \wedge \dots \wedge \widehat{dx^{i_r}} \wedge \dots \wedge dx^{i_k} & \text{für } i = i_r \end{cases} \quad (23.74)$$

Es gilt für alle $\omega \in \wedge^k(\mathbb{R}^n)^*$:

$$\iota_{e_i}(dx^j \wedge \omega) + dx^j \wedge \iota_{e_i}(\omega) = \delta_i^j \omega = \begin{cases} \omega & j = i \\ 0 & \text{sonst} \end{cases} \quad (23.75)$$

sowie

$$\sum_{i=1}^n dx^i \wedge \iota_{e_i} \omega = k \cdot \omega \quad (23.76)$$

Beweis. Übungsaufgabe 3.3. Wichtig ist noch die offensichtlichen punktweise Verallgemeinerung dieser Formeln für Vektorfelder und Differentialformen. \square

Theorem 23.9 (Poincaré-Lemma). *Es sei $U \subset \mathbb{R}^n$ offen und konvex. Für $k > 0$ existiert für jedes $\omega \in \Omega^k(U)$ mit $d\omega = 0$ ein $\eta \in \Omega^{k-1}(U)$ so, dass $d\eta = \omega$.*

Bemerkungen. · Eine Differentialform $\omega \in \Omega^k(U)$ heisst *geschlossen*, falls $d\omega = 0$ und *exakt* falls $\omega = d\eta$ für ein $\eta \in \Omega^{k-1}(U)$. $d^2 = 0$ bedeutet, dass jede exakte Form geschlossen ist; das Poincaré-Lemma besagt, dass auf konvexen Mengen jede geschlossene Form exakt ist.

· In dieser Terminologie ist die Stammform η durch $d\eta = \omega$ eindeutig bis auf die Addition einer geschlossenen $k - 1$ -Form bestimmt.

· Wie bei den Sätzen der Funktionentheorie lässt sich die Voraussetzung an den Definitionsbereich noch wesentlich abschwächen. Der Beweis unten geht ohne Weiteres für “sternförmige” U durch.

· Dem Beispiel 23.4 folgend konstruiert das Poincaré-Lemma auf geeigneten Teilmengen des \mathbb{R}^3 aus Vektorfeldern E mit $\text{rot } E = 0$ ein Potential ϕ mit $E = -\text{grad } \phi$ bzw. aus einem Vektorfeld B mit $\text{div } B = 0$ ein Vektorpotential A mit $B = \text{rot } A$.

Beweis von Thm. 23.9. Wir wählen einen beliebigen Punkt $x_0 \in U$, und definieren eine lineare Abbildung $I_{x_0} : \Omega^k(U) \rightarrow \Omega^{k-1}(U)$ durch die Formel

$$(I_{x_0}(\omega))(x)(v_1, \dots, v_{k-1}) := \int_0^1 t^{k-1} \sum_{i=1}^n (x^i - x_0^i) \omega(x_0 + t(x - x_0))(e_i, v_1, \dots, v_{k-1}) dt \quad (23.77)$$

für $\omega \in \Omega^k(U)$ (noch nicht notwendig geschlossen), $x \in U$ und $v_1, \dots, v_{k-1} \in \mathbb{R}^n$. (Dies ist wohldefiniert, da offensichtlich multi-linear und alternierend in v_1, \dots, v_{k-1} , und glatt in x wegen der Glattheit von ω und des Vertauschungssatzes 22.9). Im Sinne des Integrals vektorwertiger Funktionen ist dies auch

$$I_{x_0}(\omega)(x) = \int_0^1 t^{k-1} \iota_{\sum_i (x-x_0)^i e_i} (\omega(x_0 + t(x-x_0))) dt \quad (23.78)$$

Beh.: Es gilt für alle $\omega \in \Omega^k(U)$

$$d(I_{x_0}(\omega)) + I_{x_0}(d\omega) = \omega \quad (23.79)$$

Ist insbesondere $d\omega = 0$, so erfüllt $\eta := I_{x_0}(\omega)$ die Bedingung $d\eta = \omega$.

§ 23. DIFFERENTIALFORMEN UND DER STOKESSCHE SATZ

Bew.: Der Einfachheit halber transferieren wir die Behauptung nach $x_0 = 0$, und benutzen die Linearität in ω , um uns auf solche der Form $\omega(x) = f(x) \cdot dx^{i_1, \dots, i_k}$ zurückzuziehen, wobei wir noch $dx^{i_1, \dots, i_k} = dx^{i_1} \wedge \dots \wedge dx^{i_k}$ abkürzen. Dann aber rechnen wir unter Vertauschen von Integration und Differentiation:

$$\begin{aligned}
 & d\left(\int_0^1 t^{k-1} \sum_{i=1}^n x^i \iota_{e_i} \omega(tx) dt\right) + \int_0^1 t^k \sum_{i=1}^n x^i \iota_{e_i} (d\omega)(tx) dt = \\
 & \int_0^1 \left[t^{k-1} f(tx) \sum_{i=1}^n dx^i \wedge \iota_{e_i} (dx^{i_1, \dots, i_k}) dt + t^k \sum_{i,j=1}^n x^i \frac{\partial f}{\partial x^j}(tx) dx^j \wedge \iota_{e_i} (dx^{i_1, \dots, i_k}) \right. \\
 & \qquad \qquad \qquad \left. + t^k \sum_{i,j=1}^n x^i \iota_{e_i} \left(\frac{\partial f}{\partial x^j}(tx) dx^j \wedge dx^{i_1, \dots, i_k} \right) \right] dt \\
 & \text{(mit Lemma 23.8)} = \int_0^1 \left[t^{k-1} f(tx) \cdot k \cdot dx^{i_1, \dots, i_k} + t^k \underbrace{\sum_{i=1}^n x^i \frac{\partial f}{\partial x^i}(tx) \cdot dx^{i_1, \dots, i_k}}_{= \frac{df(tx)}{dt}} \right] dt \\
 & = \int_0^1 \frac{d}{dt} (t^k f(tx)) dt \cdot dx^{i_1, \dots, i_k} = f(x) \cdot dx^{i_1, \dots, i_k}
 \end{aligned} \tag{23.80}$$

wegen $k > 0$ und dem gewöhnlichen Hauptsatz in einer Dimension. \square

Fazit: Die Theorie der Differentialformen liefert eine einigermaßen zufriedenstellende höher-dimensionale Verallgemeinerung des Hauptsatzes der Differential- und Integralrechnung. Bei Bedarf kann er durch verschiedene Grenzprozesse noch vervollständigt werden.

Zusätzlich zum Einschub von 20.9 als Wiedereinstimmung auf die Eigenwerttheorie machen wir noch

Zwei Schlussbemerkungen

1. Es mag seltsam vorkommen, dass wir Integration zwar (teilweise) durch “Volumenmessung” motiviert haben, den Stokesschen Satz aber nur für *orientierte* Untermannigfaltigkeiten mit Rand formuliert und bewiesen haben. Sicher hat doch auch der Möbius-Streifen einen Flächeninhalt, nur vielleicht keinen orientierten! Zu seiner Definition betrachten wir zunächst wieder die lineare Situation.

• Wir bemerken, dass das euklidische Volumen von k -dimensionalen Parallelotopen durch einige wenige geometrische Eigenschaften festgelegt ist: Für jeweils k Vektoren $b_1, \dots, b_k \in \mathbb{R}^n$ soll gelten:

(i) $\text{vol}^{(k)}(b_1, \dots, b_k)$ ist invariant unter “Scherungen”, d.h. Addition des Vielfachen eines Vektors zu einem anderen.

(ii) $\text{vol}^{(k)}(b_1, \dots, b_k) = \prod_{i=1}^k \|b_i\|_2$, falls $\langle b_i, b_j \rangle = 0$ für $i \neq j$.

Denn wie wir aus § 20 wissen (sollten), lassen sich durch Operationen der Art (i) die (b_1, \dots, b_k) in einen Satz von k orthogonalen Vektoren $(\tilde{b}_1, \dots, \tilde{b}_k)$ überführen (Gram-Schmidt-Orthogonalisierungsverfahren), für die wir das Volumen durch die

Vorschrift (ii) berechnen können. Sind speziell die (b_1, \dots, b_k) linear abhängig, so ist eines der $\tilde{b}_i = 0$ und $\text{vol}^{(k)}(b_1, \dots, b_k) = 0$.

· Die Operationen unter (i) werden durch eine Matrix $R = (R_i^j)$ mit Determinante 1 dargestellt, d.h.

$$\tilde{b}_j = \sum_i R_j^i b_i \quad (23.81)$$

mit $\det(R_j^i) = 1$, und

$$\text{vol}^{(k)}(b_1, \dots, b_k) = \prod_{j=1}^k \|\tilde{b}_j\|_2 = \sqrt{\det(\text{diag}(\langle \tilde{b}_j, \tilde{b}_j \rangle))} = \sqrt{\det g} \quad (23.82)$$

wobei wir die Multiplikativität der Determinante benutzt haben, $\det R = 1$, und die sog. Gramsche Matrix gegeben ist durch

$$g = (g_{ij})_{i,j=1,\dots,k} = (\langle b_i, b_j \rangle)_{i,j=1,\dots,k} \quad (23.83)$$

Das Volumen (23.82) erfüllt nun tatsächlich die obigen Bedingungen, ist aber im Gegensatz zu (23.3) nie negativ und auch insbesondere invariant unter beliebigen Vertauschungen der b_i .

· Sind die (b_i) linear unabhängig, d.h. $\text{span}(b_1, \dots, b_k)$ ist ein k -dimensionaler Unterraum von \mathbb{R}^n , so ist $\det g \neq 0$ (man spricht dann von der induzierten Metrik). Unter Basiswechsel

$$\tilde{b}_j = \sum_i R_j^i b_i \quad (23.84)$$

transformiert sich die Gramsche Determinante gemäss

$$\sqrt{\det \tilde{g}} = |\det R| \cdot \sqrt{\det g} \quad (23.85)$$

· Ist nun M eine k -dimensionale Untermannigfaltigkeit des \mathbb{R}^n (mit oder ohne Rand), und $(U, \varphi = \psi^{-1})$ eine Karte, so hat die "induzierte euklidische Metrik" in der Basis $(D\psi(x)(e_i))_{i=1,\dots,k}$ von $T_{\psi(x)}M$ die Darstellung

$$g_{ij}(x) = \langle D\psi(x)(e_i), D\psi(x)(e_j) \rangle = \sum_{r=1}^n B_i^r(x) B_j^r(x) \quad (23.86)$$

wobei B_i^r die Jacobi-Matrix von $D\psi(x)$ in den Standardbasen ist. (Vgl. auch die Transformation des Laplace-Operators, die wir in der HöMa 2 diskutiert haben)

· Ist M kompakt, so überdeckt man wie bei Differentialformen M mit einem endlichen Atlas $\{(U_\iota, \varphi_\iota)\}$, einer untergeordneten Teilung der Eins (h_ι) , und definiert

$$\text{vol}^{(k)}(M) := \sum_\iota \int_{V_\iota} h_\iota(\psi_\iota(x)) \cdot \sqrt{\det g(x)} dx \quad (23.87)$$

· Wegen (23.85) ist diese Definition wieder unabhängig von der Wahl der Karten, diesmal aber auch unabhängig von Orientierungen: Unter einer Substitution $x =$

$\varphi \circ \tilde{\varphi}^{-1}(\tilde{x})$ ist

$$\begin{aligned} \int_V h \cdot \tilde{h} \cdot \sqrt{\det g(x)} dx &= \int_{\tilde{V}} h \cdot \tilde{h} \cdot \sqrt{\det g(x)} \cdot |\det D(\varphi \circ \tilde{\varphi}^{-1})(\tilde{x})| d\tilde{x} \\ &= \int_{\tilde{V}} h \cdot \tilde{h} \sqrt{\det \tilde{g}(\tilde{x})} d\tilde{x} \end{aligned} \quad (23.88)$$

Denn der Basiswechsel in $T_{\psi(x)}M = T_{\tilde{\psi}(\tilde{x})}M$ ist bestimmt durch

$$D\tilde{\psi}(\tilde{x}) = D\psi(x) \cdot D(\varphi \circ \tilde{\varphi}^{-1})(\tilde{x}) \quad (23.89)$$

oder anders gesagt gilt:

$$\tilde{g}_{ij}(\tilde{x}) = \sum_{r,s=1}^k \frac{\partial \varphi^r}{\partial \tilde{x}^i}(\tilde{x}) \frac{\partial \varphi^s}{\partial \tilde{x}^j}(\tilde{x}) \cdot g_{rs}(x) \quad (23.90)$$

Beispiel: Für den Möbius-Streifen in der Parametrisierung (23.48) ist

$$B = \begin{pmatrix} -\sin \vartheta \left(1 - s \sin \frac{\vartheta}{2}\right) - \frac{s}{2} \cos \vartheta \cos \frac{\vartheta}{2} & -\cos \vartheta \sin \frac{\vartheta}{2} \\ \cos \vartheta \left(1 - s \sin \frac{\vartheta}{2}\right) - \frac{s}{2} \sin \vartheta \cos \frac{\vartheta}{2} & -\sin \vartheta \sin \frac{\vartheta}{2} \\ -\frac{s}{2} \sin \frac{\vartheta}{2} & \cos \frac{\vartheta}{2} \end{pmatrix} \quad (23.91)$$

$$g = \begin{pmatrix} 1 + \frac{3s^2}{4} - \frac{s^2}{2} \cos \vartheta - 2s \sin \frac{\vartheta}{2} & 0 \\ 0 & 1 \end{pmatrix} \quad (23.92)$$

und der Flächeninhalt

$$\int_0^{2\pi} \int_{-1/2}^{1/2} \sqrt{1 + \frac{3s^2}{4} - \frac{s^2}{2} \cos \vartheta - 2s \sin \frac{\vartheta}{2}} ds d\vartheta \approx 6,353271 \dots \quad (23.93)$$

• Der Laplace-Operator auf einer Untermannigfaltigkeit

$$\Delta = \frac{1}{\sqrt{\det g}} \sum_{i,j=1}^k \frac{\partial}{\partial x^i} \left(g^{ij} \sqrt{\det g} \frac{\partial}{\partial x^j} \right) \quad (23.94)$$

2. Nach unserer Definition (23.35) sind die Tangentialräume an verschiedenen Punkten einer k -dimensionalen Untermannigfaltigkeit $M \subset \mathbb{R}^n$ wohlunterschiedene k -dimensionale Unterräume des umgebenden \mathbb{R}^n . Im Allgemeinen macht es also keinen Sinn, für $p_1 \neq p_2$ Vektoren $v_1 \in T_{p_1}M$ und $v_2 \in T_{p_2}M$ zu vergleichen oder gar zu addieren.

• Dies ist besonders dann wichtig in Erinnerung zu behalten, wenn p_1 und p_2 in einer gemeinsamen Kartenumgebung $(U, \varphi = \psi^{-1})$ liegen. Dann existieren nämlich Vektoren $w_1, w_2 \in \mathbb{R}^k$ mit $v_i = D\psi(\varphi(p_i))(w_i)$, und man könnte dazu verleitet werden zu sagen, dass “ $v_1 = v_2$ falls $w_1 = w_2$ ”. Eine solche Aussage wäre aber deshalb unsinnig, weil sie von der Wahl der Karte abhinge: Ist $(\tilde{U}, \tilde{\varphi} = \tilde{\psi}^{-1})$ eine

andere Karte mit $p_1, p_2 \in \tilde{U}$ so gäbe es Vektoren \tilde{w}_1, \tilde{w}_2 mit $v_i = D\tilde{\psi}(\tilde{\varphi}(p_i))(\tilde{w}_i)$. Es ist

$$\tilde{w}_i = D(\tilde{\varphi} \circ \varphi^{-1})(\varphi(p_i))(w_i) \quad (23.95)$$

und aus $w_1 = w_2$ folgt noch lange nicht $\tilde{w}_1 = \tilde{w}_2$, da die Differentiale der Kartenwechsel im Allgemeinen nicht konstant sind.

· Unter Verwendung der euklidischen Metrik lassen sich aber Tangentialvektoren für infinitesimal benachbarte Punkte vergleichen, und dann durch Integration auch für weiter entfernte. Für eine Kurve $\gamma : [t_1, t_2] \rightarrow M$ mit $\gamma(t_i) = p_i$ heisst eine Schar von Tangentialvektoren $v(t) \in T_{\gamma(t)}M$ *parallel entlang* γ , falls sie in erster Ordnung nur in den Richtungen orthogonal zu M variiert, d.h.

$$\mathbb{R}^n \ni \dot{v}(t) \perp T_{\gamma(t)}M \subset \mathbb{R}^n \quad \forall t \in [t_1, t_2] \quad (23.96)$$

Für vorgegebenes γ ist dies eine gewöhnliche Differentialgleichung, deren Lösung mit $v(t_1) = v_1$ den “Paralleltransport von $v_1 \in T_{p_1}M$ entlang γ nach p_2 ” definiert. (Das Ergebnis hängt allerdings i.A. von γ ab!)

· Hier und auch schon im Punkt zuvor benutzt man die Tatsache, dass der umgebende \mathbb{R}^n mit einer kanonischen Basis ausgerüstet ist, was uns erlaubt, Tangentialvektoren an den \mathbb{R}^n in verschiedenen Punkten in kanonischer Weise zu vergleichen. Mit Hilfe der Gramschen Matrix lässt sich der Begriff aber auch intrinsisch formulieren.

· Physikalisch sind Tangentialräume als “Laboratorien von wohlunterschiedenen Beobachtern” relevant, der Paralleltransport (oder auch mal andere sog. Zusammenhänge) zum “Datenvergleich”. Mehr zu all dem in der Differentialgeometrie.

KAPITEL 9

FUNKTIONALE

In diesem letzten Kapitel stellen wir (doch noch) einige Begriffe und Resultate aus der Theorie unendlich-dimensionaler normierter Vektorräume zusammen, wie sie insbesondere für ein konzeptionell befriedigendes Verständnis der Quantenmechanik mathematisch sinnvoll sind. Einige unserer Betrachtungen könnten auch als Einstieg in die auch klassisch relevante Theorie der partiellen Differentialgleichungen oder der Integraltransformationen dienen, die wir aber aus Zeitgründen ausklammern müssen. Der Blick auf die Quantenmechanik begründet dann auch die Wahl der Sturm-Liouville-Theorie als beispielhafte Anwendung der allgemeinen Untersuchungen.

§ 24 Hilbert-Räume

In der klassischen Physik stellte man sich Zustände physikalischer Systeme in idealisierter Form als Punkte in metrischen Räumen mit weiteren (hauptsächlich linearen) Strukturen dar. Wie wir in der Höheren Mathematik II ausführlich diskutiert haben, dient dabei die lokale Topologie des metrischen Raumes (damit meint man im Wesentlichen das System seiner offenen Teilmengen) dazu, die durch die übrigen Strukturen berechneten reellen Grössen *approximativ* mit realen Gegebenheiten zu vergleichen: Die Theorie eines physikalischen Systems ist erfolgreich, wenn die gemessenen und gedachten/berechneten Werte (wie beispielsweise die Positionen in einem Bezugssystem und die Koordinaten bezüglich einer Basis eines Vektorraums) innerhalb hinreichend kleiner Umgebungen voneinander liegen. Prototyp ist natürlich die klassische Mechanik, in der man Positionsänderungen mit Lösungen von gewöhnlichen Differentialgleichungen vergleicht.

In der Quantenmechanik tritt über den klassischen Konfigurationsraum eines physikalischen Systems ein komplexer Vektorraum von möglichen Zuständen. Elemente dieses Vektorraums ordnen dabei als sog. “Wellenfunktionen” möglichem Realtext eine komplexe Zahl, die sog. Amplitude, zu. Die Theorie ist erfolgreich, wenn der Absolutbetrag dieser Amplitude mit der (richtig gedeuteten, und natürlich approximativen) *Wahrscheinlichkeit* für die betreffenden experimentellen Befunde übereinstimmt. Der leicht irreführende Name “Wellenfunktion” rührt daher, dass man sich in einer gewissen (mathematisch unpräzisen) Näherung manchmal den im physikalischen Raum verteilten Realtext als Basis des Zustandsraum vorstellen kann. Tatsächlich schliesst allerdings die “Inkommensurabilität” verschiedener Observablen (Stichwort Unschärferelation) die Auszeichnung einer (selbst approximativen) Basis aus. Dies hat insbesondere zur Folge, dass die Approximation im Ort nicht die einzig physikalisch sinnvolle ist.

Zur rein mathematischen Motivation erinnern wir an dieser Stelle an die Rolle von Vollständigkeit und Positivität für die Eigenwerttheorie im endlich-dimensionalen Fall: Zwar ist diese Theorie zunächst algebraischer Natur. Für die allgemeine Existenz von Eigenwerten haben wir aber den Fundamentalsatz der Algebra gebraucht, welchen wir wiederum als Folge der Vollständigkeit von \mathbb{R} (bzw. \mathbb{C}) hergeleitet hatten. Diagonalisierbarkeit war eine Folge der Selbstadjungiertheit bezüglich einem inneren Produkt, und der Zusammenhang zu Symmetriegruppen entstand über Differentiation bzw. die Exponentialreihe. Vergleiche auch das Variationsprinzip 20.9.

Eine sparsame Kombination dieser Konzepte ergibt sich aus dem bereits in 4.5 festgehaltenen Zusammenhang zwischen innerem Produkt und Norm:

Lemma 24.1. *Es sei V ein komplexer Vektorraum.*

(i) *Ist $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{C}$ ein hermitesches inneres Produkt, so definiert*

$$\|\cdot\| : V \rightarrow \mathbb{R}, \quad v \mapsto \|v\| := \sqrt{\langle v, v \rangle} \quad (24.1)$$

eine Norm auf V .

(ii) *Inneres Produkt und Norm erfreuen sich verschiedener komplexer Verallgemeinerungen von vertrauten geometrischen Eigenschaften. Insbesondere gilt die Cauchy-Schwarz-Ungleichung:*

$$|\langle v, w \rangle| \leq \|v\| \cdot \|w\| \quad \forall v, w \in V \quad (24.2)$$

(iii) *Eine Norm $\|\cdot\| : V \rightarrow \mathbb{R}$ stammt genau dann von einem inneren Produkt, falls sie die Parallelogrammgleichung*

$$\|v + w\|^2 + \|v - w\|^2 = 2(\|v\|^2 + \|w\|^2) \quad (24.3)$$

erfüllt.

Beweis. Die Bedingungen an die Form $\langle \cdot, \cdot \rangle$ sind (vgl. 20.1): $\langle v, \mu w_1 + w_2 \rangle = \mu \langle v, w_1 \rangle + \langle v, w_2 \rangle$, $\langle v, w \rangle = \overline{\langle w, v \rangle}$ und $\langle v, v \rangle > 0$ für $v \neq 0$. Zu zeigen ist für (i), dass $\|\lambda v\| = |\lambda| \cdot \|v\|$, $\|v\| = 0 \Rightarrow v = 0$ und $\|v + w\| \leq \|v\| + \|w\|$. Die beiden ersten Aussagen folgen direkt aus der Sesquilinearität bzw. der Positivität des inneren Produkts. Für die Dreiecksungleichung zeigen wir nach Ausmultiplizieren

$$\|v + w\|^2 = \|v\|^2 + \underbrace{\langle v, w \rangle + \langle w, v \rangle}_{=2 \operatorname{Re}(\langle v, w \rangle)} + \|w\|^2 \quad (24.4)$$

zunächst die C.-S.-Ungleichung:

· Für $v \neq 0$ (für $v = 0$ ist die Aussage trivial) ist die Zerlegung

$$w = \underbrace{\langle v, w \rangle \frac{v}{\|v\|^2}}_{w_{\parallel}} + \underbrace{w - \langle v, w \rangle \frac{v}{\|v\|^2}}_{w_{\perp}} \quad (24.5)$$

orthogonal bezüglich $V = \mathbb{C} \cdot v \oplus v^{\perp}$, d.h. $\langle v, w_{\perp} \rangle = 0$ und $\langle w_{\perp}, w_{\parallel} \rangle = 0$. Es folgt

$$\|w\|^2 = \|w_{\parallel}\|^2 + \underbrace{\|w_{\perp}\|^2}_{\geq 0} \geq \frac{|\langle v, w \rangle|^2}{\|v\|^2} \quad (24.6)$$

§ 24. HILBERT-RÄUME

und daraus die Behauptung.

· Damit ist die rechte Seite von (24.4) beschränkt durch

$$\|v + w\|^2 \leq (\|v\| + \|w\|)^2 \quad (24.7)$$

und daraus folgt die Dreiecksungleichung. Ebenso folgt die Gültigkeit der Parallelogrammgleichung für gegebenes inneres Produkt aus (24.4) + $\|v - w\|^2$.

· Für die umgekehrte Richtung sei $\|\cdot\| : V \rightarrow \mathbb{R}$ eine Norm, welche die Parallelogrammgleichung erfüllt.

Beh.:

$$h(v, w) := \frac{1}{4}(\|v + w\|^2 - \|v - w\|^2) + \frac{i}{4}(\|v - iw\|^2 - \|v + iw\|^2) \quad (24.8)$$

definiert eine positiv definite symmetrische Sesquilinearform mit $h(v, v) = \|v\|^2$.

Bew.: Wir schreiben $h(v, w) = r(v, w) + ir(v, -iw)$ mit

$$r(v, w) := \frac{1}{4}(\|v + w\|^2 - \|v - w\|^2) \quad (24.9)$$

Offenbar ist r eine stetige reellwertige Funktion auf $V \times V$ mit $r(v, w) = r(w, v)$ $\forall v, w \in V$ und $r(v, -w) = -r(v, w)$, insbesondere $r(v, 0) = 0$.

· Aus der Homogenität der Norm folgt, dass $r(\lambda v, \lambda w) = |\lambda|^2 r(v, w)$ für $\lambda \in \mathbb{C}$. Insbesondere gilt

$$r(v, -iw) = r(iv, w) = r(w, iv) = -r(w, -iv) \quad (24.10)$$

und daraus folgt $\overline{h(v, w)} = h(w, v)$ sowie $r(v, -iv) = -r(v, -iv) = 0$, d.h. $h(v, v) = r(v, v) = \|v\|^2$.

· Aus der Gültigkeit von (24.3) folgt für $v, w_1, w_2 \in V$:

$$\begin{aligned} \|v + w_1 + w_2\|^2 &= -\|v + w_1 - w_2\|^2 + 2\|v + w_1\|^2 + 2\|w_2\|^2 \\ \|v - w_1 - w_2\|^2 &= -\|v + w_1 - w_2\|^2 + 2\|v - w_1\|^2 + 2\|w_2\|^2 \end{aligned} \quad (24.11)$$

und daraus

$$\begin{aligned} 4(r(v, w_1 + w_2) - r(v, w_1) - r(v, w_2)) &= \|v + w_1 + w_2\|^2 - \|v - w_1 - w_2\|^2 \\ &\quad - \|v + w_1\|^2 + \|v - w_1\|^2 - \|v + w_2\|^2 + \|v - w_2\|^2 \\ &= \|v + w_1\|^2 + \|v - w_1\|^2 - 2\|w_1\|^2 - \|v + w_2\|^2 - \|v - w_2\|^2 + 2\|w_2\|^2 \\ &= 2\|v\|^2 - 2\|v\|^2 = 0 \end{aligned} \quad (24.12)$$

r ist also additiv im zweiten Argument (und wegen der Symmetrie daher auch im ersten). Daraus folgt rekursiv, dass $r(v, pw) = p \cdot r(v, w)$, wegen $r(v, -w) = -r(v, w)$ sogar $\forall p \in \mathbb{Z}$, d.h. auch $r(v, \frac{1}{q}w) = \frac{1}{q}r(v, w) \forall q \in \mathbb{N}$. Es gilt also $r(v, \alpha w) = r(v, \alpha w) \forall \alpha \in \mathbb{Q}$, was sich wegen der Stetigkeit auf $\alpha \in \mathbb{R}$ fortsetzt.

· Zuletzt folgt rein algebraisch aus der Definition, dass $h(v, iw) = ih(v, w)$, und daher ist h komplex (sesqui-)linear. \square

Jeder (auch unendlich-dimensionale) unitärer Raum ist also in kanonischer Weise normiert, und wir können in sinnvoller Weise nach der Konvergenz von Cauchy-Folgen fragen.

Definition 24.2. Ein Hilbertraum ist ein unitärer Raum $(\mathcal{H}, \langle \cdot, \cdot \rangle)$, welcher mit der durch die Norm (24.1) induzierten Abstandsfunktion ein vollständiger metrischer Raum ist.

Beispiel 24.3. ³⁷ Jeder endlich-dimensionale unitäre Raum $(V, \langle \cdot, \cdot \rangle)$ ist ein Hilbertraum. Wir bemerken, dass man die Vollständigkeit hier auf zwei leicht verschiedene Weisen auf die von \mathbb{R} bzw. \mathbb{C} zurückführen kann: Entweder benutzt man die eher analytische Aussage, dass alle Normen auf einem endlich-dimensionalen reellen oder komplexen Vektorraum äquivalent zueinander sind, insbesondere zur max-Norm auf \mathbb{R}^n , s. 4.18. Oder man nutzt zunächst die eher algebraische Tatsache aus, dass V als unitärer Raum isomorph zu $\mathbb{C}^{n=\dim V}$ mit dem inneren Standardprodukt

$$\langle z, w \rangle = \sum_{i=1}^n \bar{z}^i w^i \tag{24.13}$$

ist (dessen Vollständigkeit dann allerdings noch wie oben zu zeigen wäre). Zur Erinnerung: Diese Aussage ist äquivalent zur Existenz einer Orthonormalbasis (b_1, \dots, b_n) von V , d.h. $\langle b_i, b_j \rangle = \delta_{ij}$, die man sich am bequemsten durch Induktion nach der Dimension konstruiert (vgl. 20.2): Für $v \in V$ mit $v \neq 0$ ist

$$v^\perp := \text{Ker}(\langle v, \cdot \rangle) \tag{24.14}$$

ein unitärer Raum mit $\dim v^\perp = \dim V - 1$, von dem man eine ONB durch $\hat{v} = \frac{v}{\|v\|}$ zu einer ONB von V ergänzen kann.

- Durch Verletzung der Parallelogrammgleichung zeigt man, dass auf \mathbb{C}^n andere Normen wie etwa max oder Manhattan nicht von einem inneren Produkt herkommen, auch wenn sie die gleiche Topologie induzieren.
- Ein kanonisches unendlich-dimensionales Beispiel erhält man durch den (formal präzisierbaren) Limes $n \rightarrow \infty$ in (24.13): Für eine abzählbare Menge \mathbb{L} (z.B. \mathbb{N} oder \mathbb{Z}) ist der Raum der quadratsummierbaren Folgen

$$\ell^2(\mathbb{L}) := \left\{ (a_k)_{k \in \mathbb{L}} \subset \mathbb{C} \mid \sum_{k \in \mathbb{L}} |a_k|^2 < \infty \right\} \tag{24.15}$$

(Zur Erinnerung: Die Konvergenz auf der rechten Seite ist gleichbedeutend damit, dass die Menge der endlichen Teilsummen

$$\sum_k' |a_k|^2 \tag{24.16}$$

beschränkt ist.) mit dem inneren Produkt

$$\langle (a_k), (b_k) \rangle = \sum_{k \in \mathbb{L}} \bar{a}_k b_k \tag{24.17}$$

ein Hilbertraum: Die Wohldefiniertheit, dass nämlich Summen von quadratsummierbaren Folgen wieder quadratsummierbar sind, sowie die Konvergenz von (24.17)

³⁷Wir schreiben, unabhängig vom Vektorraum, im Folgenden immer $\langle \cdot, \cdot \rangle$ für das innere Produkt und $\|\cdot\|$ für die Norm.

§ 24. HILBERT-RÄUME

für $(a_k), (b_k) \in \ell^2(\mathbb{L})$, folgt aus der C.S.U. bzw. Dreiecksungleichung in endlicher Dimension zusammen mit dem Majorantenkriterium für Reihen. Sesquilinearität, Symmetrie und Positivität sind sowieso klar. Zur Vollständigkeit sei

$$((a_k^{(n)})_{k \in \mathbb{L}})_{n \in \mathbb{N}} \subset \ell^2(\mathbb{L}) \tag{24.18}$$

eine Cauchy-Folge (von Folgen). Dann folgt aus $\sum_{k \in \mathbb{L}} |a_k^{(n)} - a_k^{(m)}|^2 < \epsilon$ für $n, m > N$ gross genug die gleiche Ungleichung für jedes k , d.h. für jedes $k \in \mathbb{L}$ ist $(a_k^{(n)})_{n \in \mathbb{N}}$ eine Cauchy-Folge von komplexen Zahlen. Sei

$$a_k := \lim_{n \rightarrow \infty} a_k^{(n)} \tag{24.19}$$

Für $n, m > N$ gross genug ist dann jede endliche Teilsumme

$$\sum_k' |a_k^{(n)} - a_k^{(m)}|^2 \leq \sum_{k \in \mathbb{L}} |a_k^{(n)} - a_k^{(m)}|^2 < \epsilon \tag{24.20}$$

und mit $m \rightarrow \infty$ folgt

$$\sum_k' |a_k^{(n)} - a_k|^2 < \epsilon \tag{24.21}$$

für jede endliche Teilsumme. Aus dem Kriterium 5.7 folgt

$$\sum_{k \in \mathbb{L}} |a_k^{(n)} - a_k|^2 \leq \epsilon \tag{24.22}$$

für $n \geq N$ und daraus $(a_k)_{k \in \mathbb{L}} \in \ell^2(\mathbb{L})$ (genauer: $(a_k^{(n)} - a_k)_{k \in \mathbb{L}} \in \ell^2(\mathbb{L})$ und $\ell^2(\mathbb{L})$ ist ein Vektorraum), sowie $(a_k)_{k \in \mathbb{L}} = \lim_{n \rightarrow \infty} (a_k^{(n)})$ in der ℓ^2 -Norm.

· Ein “Unbeispiel” erhält man durch eine andere Limesbildung in (24.13): Der Raum der endlichen Folgen, d.h. $\mathcal{F}^{\text{fin}}(\mathbb{L}, \mathbb{C})$ mit dem gleichen inneren Produkt ist zwar ein unitärer Raum, aber nicht vollständig, d.h. kein Hilbertraum. Dies gibt uns auch Gelegenheit darauf hinzuweisen, dass auf $\mathcal{F}^{\text{fin}}(\mathbb{N}, \mathbb{C})$ die ℓ^1 -Norm, d.h.

$$\sum_{k \in \mathbb{L}} |a_k| \tag{24.23}$$

nicht äquivalent zu (24.17) ist: Die Folgenfolge $a_k^{(n)} = \frac{1}{k}$ für $k \leq n$, 0 sonst, ist in der ℓ^1 -Norm keine Cauchy-Folge, in der ℓ^2 -Norm aber schon. (Sie konvergiert aber nicht in $\mathcal{F}^{\text{fin}}(\mathbb{N})$, sondern erst in dessen Vervollständigung $\ell^2(\mathbb{N})$.)

· Der Fall $\mathbb{L} = \mathbb{Z}^n$ spielt in der Festkörperphysik als Kristallgitter eine grosse Rolle und kann auch mal als Diskretisierung des physikalischen Raumes \mathbb{R}^n benutzt werden.

· Die natürliche kontinuierliche Version von $\mathcal{F}^{\text{fin}}(\mathbb{Z}^n, \mathbb{C})$ ist allerdings

$$\mathcal{C}_c(\mathbb{R}^n, \mathbb{C}) = \{f : \mathbb{R}^n \rightarrow \mathbb{C} \mid f \text{ stetig und } \text{supp}(f) \text{ kompakt}\} \tag{24.24}$$

mit dem inneren Produkt

$$\langle f, g \rangle = \int_{\mathbb{R}^n} \bar{f}(x)g(x) dx \tag{24.25}$$

(s. Aufgabe 5.3 für die Positivität). Dieser unitäre Raum ist ebenfalls nicht vollständig, wobei zusätzlich zum Verhalten bei ∞ auch lokale Divergenzen eine Rolle spielen, s. unten.

· Zunächst noch die Bemerkung, dass wir in unendlicher Dimension das obige Gram-Schmidt-Verfahren (damit meint man allgemein die Konstruktion eines Orthonormalsystems (b_1, b_2, \dots) aus einem abzählbaren System (v_1, v_2, \dots) von linear unabhängigen Vektoren durch

$$\begin{aligned} b_1 &:= \frac{v_1}{\|v_1\|} \\ b_2 &:= \frac{b_2 - \langle b_1, v_2 \rangle b_1}{\|b_2 - \langle b_1, v_2 \rangle b_1\|} \\ &\vdots \end{aligned} \quad (24.26)$$

nicht zur Konstruktion einer “echten” Vektorraumbasis (orthogonal oder sonstig) benutzen können. Beispielsweise ist in $\ell^2(\mathbb{N})$ das System $E = (e_i)_{i=1,2,\dots}$ mit

$$e_i = (0, \dots, 0, \underset{\substack{\uparrow \\ i\text{-te Stelle}}}{1}, 0, \dots, 0, \dots) \quad (24.27)$$

orthonormal, d.h. $\langle e_i, e_j \rangle = \delta_{ij}$, aber keine Basis von $\ell^2(\mathbb{N})$ (der nicht von abzählbarer Dimension ist³⁸), und E kann auch nicht zu einer ONB ergänzt werden, denn

$$\begin{aligned} E^\perp &= \{(a_k)_{k \in \mathbb{N}} \mid \langle e_i, (a_k)_{k \in \mathbb{N}} \rangle = 0 \ \forall i\} \\ &= \{(a_k)_{k \in \mathbb{N}} \mid a_i = 0 \ \forall i\} \\ &= \{0\} \end{aligned} \quad (24.28)$$

(Es gibt also gar keine zu allen e_i orthogonale Vektoren!) Stattdessen kann aber jede Folge $(a_k) \in \ell^2(\mathbb{N})$ beliebig genau (in der ℓ^2 -Norm) durch endliche Linearkombinationen der e_i approximiert werden: Die Folgen

$$a_k^{(n)} = \begin{cases} a_k & k \leq n \\ 0 & k > n \end{cases} \quad (24.29)$$

erfüllen $(a_k^{(n)})_k \in \mathcal{F}^{\text{fin}}(\mathbb{N}, \mathbb{C}) \subset \ell^2(\mathbb{N})$ und

$$\lim_{n \rightarrow \infty} \|(a_k^{(n)} - a_k)_{k \in \mathbb{N}}\| = 0 \quad (24.30)$$

Definition 24.4. Für einen metrischen Raum X heisst eine Teilmenge $Q \subset X$ *dicht*, falls $\overline{Q} = X$ (der topologische Abschluss)

· Für einen Hilbertraum $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ heisst eine Teilmenge $(b_\iota)_{\iota \in I}$ (I irgendeine Menge) *Hilbertbasis*, falls $\langle b_\iota, b_\kappa \rangle = \delta_{\iota\kappa}$ und $\text{span}_{\mathbb{C}}(b_\iota)$ (d.i. die Menge aller endlichen komplexen Linearkombinationen) dicht in \mathcal{H} ist.

· Ein Hilbertraum heisst *separabel*, falls er eine endliche oder abzählbare Hilbertbasis besitzt.

³⁸ E ist aber eine Basis von $\mathcal{F}^{\text{fin}}(\mathbb{N}, \mathbb{C})$!

Prinzipien 24.5. · In der Physik spielen nur separable Hilberträume eine Rolle: Ist $(b_k)_{k=1,2,\dots}$ eine Hilbertbasis von \mathcal{H} , so folgt aus der Dichtheit von \mathbb{Q} in \mathbb{R} sowie der Tatsache, dass abzählbare Vereinigungen von abzählbaren Mengen wieder abzählbar sind, dass die abzählbare Menge

$$\text{span}_{\mathbb{Q}(i)}\{b_k \mid k = 1, 2, \dots\} = \left\{ \sum_{k=1}^{\infty} a_k b_k \mid a_k \in \mathbb{Q} + i\mathbb{Q}, \text{ fast alle } 0 \right\} \quad (24.31)$$

dicht in \mathcal{H} liegt—Jeder “physikalische” Zustand in \mathcal{H} lässt sich rational (d.h. durch endlich viele Schritte beliebig genau) in $\text{span}_{\mathbb{Q}}(b_k)$ approximieren.

· Umgekehrt gilt: Existiert überhaupt eine abzählbare dichte Teilmenge (zunächst nicht notwendig ein Unterraum), so folgt durch Anwenden des Gram-Schmidt-Verfahrens, dass \mathcal{H} separabel ist.

· Man zeigt leicht, dass ein separabler Hilbertraum entweder endlich-dimensional oder isomorph zu $\ell^2(\mathbb{N})$ ist: Ist (b_k) eine Hilbertbasis von \mathcal{H} , so ist

$$\begin{aligned} i : \ell^2(\mathbb{N}) &\rightarrow \mathcal{H} \\ (a_k)_{k \in \mathbb{N}} &\mapsto \sum_{k=1}^{\infty} a_k b_k \end{aligned} \quad (24.32)$$

wohldefiniert, linear, bijektiv und normerhaltend, d.h. $\|(a_k)\| = \|i(a_k)\|$, wobei die erste Norm die in $\ell^2(\mathbb{N})$ ist, die zweite in \mathcal{H} . Mit der Parallelogrammgleichung folgt, dass i auch das innere Produkt erhält, d.h. ein unitärer Isomorphismus ist.

· Man nennt Hilbertbasen oft auch vollständige Orthonormalsysteme und schreibt sie nach Dirac als

$$\{|k\rangle \mid k \in \mathbb{N}\} \quad \text{bzw.} \quad \{|\text{Realtex}\rangle\} \quad (24.33)$$

Schrödinger’sche Wellenfunktionen

Für die übliche Quantenmechanik zunächst relevant ist der Hilbertraum der quadratintegrierbaren komplexwertigen Funktionen auf dem \mathbb{R}^n (speziell $n = 1, 2, 3$, allgemeiner: Funktionen auf einer Untermannigfaltigkeit des \mathbb{R}^n). Was ist damit gemeint?

· Die Idee zur Abhilfe der Unvollständigkeit von $\mathcal{C}_c(\mathbb{R}^n, \mathbb{C})$ (unter Berücksichtigung der Separabilität!) ist es, einfach “alle” Funktionen $f : \mathbb{R}^n \rightarrow \mathbb{C}$ als Zustände zuzulassen, für die

$$(\|f\|_2)^2 = \int_{\mathbb{R}^n} |f(x)|^2 dx < \infty \quad (24.34)$$

(quadratintegrierbare/normierbare Wellenfunktionen).

· Im Unterschied zum Übergang von $\mathcal{F}^{\text{fin}}(\mathbb{Z}^n, \mathbb{C})$ zu $\ell^2(\mathbb{Z}^n)$ müssen wir dabei aber berücksichtigen, dass wir gar nicht wirklich alle Funktionen auf (Quadrat-)Integrierbarkeit untersuchen können. Dies ist etwas klarer im masstheoretischen Zugang zum Lebesgue-Integral (s. (22.28) für ein Beispiel einer Funktion, bei der wir nicht entscheiden können, was für ein Integral sie hat.), kann aber auch im Daniell-Prozess unschwer erledigt werden.

· Physikalisch wichtiger ist es, sich zu überlegen, aus welchen Gründen (auch vernünftige) Funktionen nicht quadratintegrierbar sind.

Beispiel: $n = 1, \alpha, \beta \in \mathbb{R}$.

$$f(x) = \begin{cases} |x|^\alpha & 0 < |x| < 1 \\ |x|^\beta & |x| \geq 1 \\ \text{beliebig} & \text{für } x = 0 \end{cases} \quad (24.35)$$

Dann gilt:

$$\int_{\mathbb{R}} |f(x)|^2 dx < \infty \iff \beta < -\frac{1}{2} \text{ und } \alpha > -\frac{1}{2} \quad (24.36)$$

· Die meisten anderen Fälle kann man mit Majoranten auf (24.36) zurückführen. In der diskreten Version sind all Funktionen lokal beschränkt, ausserdem existieren keine Nullmengen.

Definition 24.6. · Eine Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ heisst messbar, falls für alle kompakten Mengen $K \subset \mathbb{R}^n$ und alle reellen Schranken $B > 0$ die Funktion

$$f_K^B := \max(\min(f, B \cdot \chi_K), -B \cdot \chi_K) \in \mathcal{L}^1(\mathbb{R}^n) \quad (24.37)$$

(f_K^B hat Träger in K und ist durch B beschränkt, $|f_K^B| \leq B \cdot \chi_K$ und $B \cdot \chi_K$ hat auf jeden Fall endliches Integral. (24.37) detektiert also die Schwierigkeiten, die durch zu viele Unstetigkeiten in f entstehen.)

· Eine Funktion $f : \mathbb{R}^n \rightarrow \mathbb{C}$ heisst messbar, falls ihre Real- und Imaginärteil messbar sind.

Lemma 24.7. · Die Menge $\mathcal{M}(\mathbb{R}^n, \mathbb{R})$ der reellen messbaren Funktionen ist ein Algebra-Verband, d.h. mit $f, g \in \mathcal{M}(\mathbb{R}^n, \mathbb{R})$ sind auch $f + g, \lambda f, f \cdot g, \max(f, g)$ und $\min(f, g)$ messbar.

· $\mathcal{M}(\mathbb{R}^n, \mathbb{C})$ ist eine \mathbb{C} -Algebra und abgeschlossen unter $f \mapsto \bar{f}$ und $f \mapsto |f|$.

Bemerkungen. Die wichtigste Neuerung ist die Abgeschlossenheit unter Multiplikation. Die Einschränkung auf messbare Funktionen wird dann sicherstellen, dass das innere Produkt von quadratintegrierbaren Funktionen wohldefiniert ist. Im Allgemeinen folgt aus $|f| \in \mathcal{M}(\mathbb{R}^n)$ nicht notwendig $f \in \mathcal{M}(\mathbb{R}^n)$. \square

Aus den Konvergenzsätzen für das Lebesgue-Integral folgert man noch:

Lemma 24.8. Ist $(f_k)_{k \in \mathbb{N}} \subset \mathcal{M}(\mathbb{R}^n)$ eine punktweise konvergente Folge von messbaren Funktionen, so ist $f(x) = \lim_{k \rightarrow \infty} f_k(x)$ ebenfalls messbar.

Beweisidee. Wende 22.4 auf die Folge $((f_k)_K^B)_{k \in \mathbb{N}}$ an. \square

Im Folgenden spielt auch noch die Tatsache eine Rolle, dass Nullfunktionen, e.s. die Funktionen, deren algebraischer Träger (ohne Abschluss) eine Nullmenge ist, integrierbar, also insbesondere messbar sind, und man daher eine Funktion auf einer Nullmenge abändern darf, ohne die Messbarkeit zu beeinflussen. Die analoge Aussage für integrierbare Funktionen hätte man verwenden können, um mit Hilfe von B. Levi den Limes einer monotonen Folge $(f_k) \subset \mathcal{L}^1(\mathbb{R}^n, \mathbb{R})$ mit beschränkten Integralen zu definieren: $f = \lim f_k$ hat Werte $\pm\infty$ höchstens auf einer Nullmenge N_f , d.h. mit $\tilde{f}(x) = 0$ für $x \in N_f$, $\tilde{f}(x) = f(x)$ für $x \notin N_f$ ist $\tilde{f} \in \mathcal{L}^1(\mathbb{R}^n, \mathbb{R})$ und $\int_{\mathbb{R}^n} \tilde{f} = \int_{\mathbb{R}^n} f = \lim \int_{\mathbb{R}^n} f_k$.

Das Analogon des Majorantenkriteriums für Reihen ist dann

§ 24. HILBERT-RÄUME

Lemma 24.9. Zu $f \in \mathcal{M}(\mathbb{R}^n, \mathbb{C})$ existiere ein $F \in \mathcal{L}^1(\mathbb{R}^n, \mathbb{R}) \subset \mathcal{M}(\mathbb{R}^n, \mathbb{C})$ mit $|f| \leq F$ fast überall. Dann ist $f \in \mathcal{L}^1(\mathbb{R}^n, \mathbb{C})$.

Beweis. Mit (22.4) □

Definition 24.10.

$$\mathcal{L}^2(\mathbb{R}^n) = \mathcal{L}^2(\mathbb{R}^n, \mathbb{C}) := \{f \in \mathcal{M}(\mathbb{R}^n, \mathbb{C}) \mid |f|^2 \in \mathcal{L}^1(\mathbb{R}^n)\}$$

Lemma 24.11. (i) $\mathcal{L}^2(\mathbb{R}^n)$ ist ein komplexer Vektorraum.
(ii) Für $f, g \in \mathcal{L}^2(\mathbb{R}^n)$ ist $\bar{f} \cdot g \in \mathcal{L}^1(\mathbb{R}^n, \mathbb{C})$ und die Abbildung

$$(f, g) \mapsto \langle f, g \rangle := \int_{\mathbb{R}^n} \bar{f}(x)g(x)dx \in \mathbb{C} \tag{24.38}$$

ist eine (konjugiert) symmetrische Sesquilinearform.

Beweis. Die Aussagen folgen mit

$$\begin{aligned} |f + g|^2 &\leq (|f|^2 + |g|^2)^2 \leq (2 \max(|f|, |g|))^2 \leq 4 \max(|f|^2, |g|^2) \\ \text{bzw. } |\bar{f} \cdot g| &\leq \max(|f|^2, |g|^2) \end{aligned} \tag{24.39}$$

ähnlich wie bei Reihen aus dem Majorantenkriterium, hier 24.9. (Hier wird benötigt, dass mit f, g auch $\bar{f} \cdot g$ messbar ist.) □

Das eben eingeführte $\langle \cdot, \cdot \rangle$ ist aber nicht positiv definit, sondern nur semi-definit, d.h. $\langle f, f \rangle \geq 0 \forall f$, aber aus $\langle f, f \rangle = 0$ folgt nicht, dass $f = 0$. Vielmehr gilt:

$$\langle f, g \rangle = 0 \quad \forall f \in \mathcal{L}^2(\mathbb{R}^n) \Leftrightarrow g \text{ ist eine Nullfunktion} \tag{24.40}$$

und der Raum der Nullfunktionen ist ein Unterraum von $\mathcal{L}^2(\mathbb{R}^n)$.

Definition 24.12.

$$L^2(\mathbb{R}^n) := \mathcal{L}^2(\mathbb{R}^n) / \text{Nullfunktionen}$$

im Sinne der Quotientenbildung (18.13). Streng genommen sind Elemente von $L^2(\mathbb{R}^n)$ also Äquivalenzklassen $[f]$ für $f \in \mathcal{L}^2(\mathbb{R}^n)$ mit $[f] = [\tilde{f}] \Leftrightarrow f - \tilde{f}$ ist eine Nullfunktion. Diese präzise Notation findet man aber nur äusserst selten; normalerweise sagt man $f \in L^2(\mathbb{R}^n)$ und denkt sich die Äquivalenz. (Dies liegt u.A. daran, dass die wichtigsten Wellenfunktionen stetig sind, und als solche ihre Äquivalenzklasse in ausgezeichneter Weise repräsentieren.)

Theorem 24.13. $(L^2(\mathbb{R}^n), \langle \cdot, \cdot \rangle)$ ist ein Hilbertraum

Beweis. Hier folgen die algebraischen Eigenschaften und die Positivität direkt aus der Wohldefiniertheit der von (24.38) induzierten Sesquilinearform. Es bleibt also nur die Vollständigkeit bzgl. $\|f\|_2 = \sqrt{\langle f, f \rangle}$ zu zeigen. (Satz von Fischer-Riesz)
· Es sei dazu $(f_k) \subset L^2(\mathbb{R}^n)$ eine Cauchy-Folge (von Äquivalenzklassen von quadratintegrierbaren messbaren Funktionen), und behaupte, dass (f_k) in $L^2(\mathbb{R}^n)$ konvergiert, dass also eine (Klasse von) $f \in L^2(\mathbb{R}^n)$ existiert so, dass $\lim_{k \rightarrow \infty} \|f - f_k\|_2 = 0$.

· Da eine Cauchy-Folge genau dann konvergiert, wenn mindestens eine Teilfolge im Raum konvergiert, können wir nach Übergang zu einer Teilfolge annehmen, dass

$$\|f_k - f_{k-1}\| \leq 2^{-k} \quad k \geq 2 \quad (24.41)$$

Wir beweisen dann wie bei $\ell^2(\mathbb{L})$ zunächst punktweise Konvergenz fast überall und dann Konvergenz in der L^2 -Norm.

· Wir betrachten (für gewählte Repräsentanten $f_k \in \mathcal{L}^2(\mathbb{R}^n)$) die Funktionenfolge

$$g_1 := |f_1|, \quad g_2 := |f_1| + |f_2 - f_1|, \dots, \quad g_k = g_{k-1} + |f_k - f_{k-1}|, \dots \quad (24.42)$$

Dann ist $(g_k) \subset \mathcal{L}^2(\mathbb{R}^n)$ punktweise monoton wachsend mit

$$\|g_k\| \leq \|f_1\| + \sum_{i=2}^k \|f_i - f_{i-1}\| \quad (24.43)$$

wegen der Dreiecksungleichung für die Norm und (24.41). Daher ist $\left(\int_{\mathbb{R}^n} g_k^2\right)$ beschränkt und mit B. Levi 22.3 folgt im Sinne der obigen Bemerkung, dass der punktweise fast überall gebildete Limes $g = \lim g_k$ eine Klasse in $L^2(\mathbb{R}^n)$ definiert. Für jedes x fast überall ist also $(g_k(x))$ eine (reelle) Cauchy-Folge, und daher auf $(f_k(x))$. (Es gilt $|f_k - f_{k-1}| = |g_k - g_{k-1}|$.) Es sei $f(x) = \lim f_k(x)$ die punktweise (fast überall wohldefinierte) Grenzfunktion.

Beh.: $f \in \mathcal{L}^2(\mathbb{R}^n)$ und $f_k \rightarrow f$ in der L^2 -Norm.

Bew.: Es gilt $f \in \mathcal{M}(\mathbb{R}^n, \mathbb{C})$ wegen 24.8 und

$$|f_k| \leq |f_1| + \sum_{i=2}^k |f_i - f_{i-1}| = g_k \leq g \quad (\text{fast überall}) \quad (24.44)$$

Daraus folgt $|f - f_k|^2 \leq 4g^2$ (fast überall) und wegen $f - f_k \in \mathcal{M}(\mathbb{R}^n, \mathbb{C})$ aus 24.9, dass $f - f_k \in \mathcal{L}^2(\mathbb{R}^n)$ und mit 22.4, dass $\lim_{k \rightarrow \infty} \|f - f_k\| = 0$. (Es macht hier keinen Unterschied, ob die Einschränkung “fast überall” als von k abhängig oder unabhängig aufgefasst wird, denn die Vereinigung von abzählbar vielen Nullmengen ist wieder eine Nullmenge.)

□

Die wichtige **Separabilität** von $L^2(\mathbb{R}^n)$ scheint a priori (etwa von $\ell^2(\mathbb{Z}^n)$ her gesehen) etwas überraschend, und wird erst durch die explizite Konstruktion von vollständigen Orthonormalsystemen wirklich klar, s. § 26. Etwas abstrakter kann man sie auch in die folgenden zwei Schritte zerlegen.

(1) $\mathcal{C}_c(\mathbb{R}^n, \mathbb{C})$ liegt dicht in $L^2(\mathbb{R}^n)$. (Stetige Funktionen mit kompaktem Träger sind automatisch quadratintegrierbar und die einzigen stetigen Repräsentanten ihrer Äquivalenzklasse, man kann also kanonisch $\mathcal{C}_c \subset L^2$ identifizieren.) Den Beweis lassen wir aus. (Idee: Ersetze f durch eine Folge beschränkter Funktionen mit kompaktem Träger, die sich gemäss der Definition 21.15 unter dem Integral durch Funktionen in \mathcal{C}_c approximieren lassen.)

§ 24. HILBERT-RÄUME

(2) Treppenfunktionen mit rationalen Werten auf rationalen Quadern sind abzählbar und liegen dicht in $\mathcal{C}_c(\mathbb{R}^n)$ in der L^2 -Norm, d.h. für jedes $f \in \mathcal{C}_c(\mathbb{R}^n)$ existiert $\forall \epsilon > 0$ eine solche rationale Treppenfunktion $\varphi \in \mathcal{T}_{\text{rat}}(\mathbb{R}^n) = \mathcal{T}(\mathbb{Q}^n, \mathbb{Q}(i))$ mit

$$\int_{\mathbb{R}^n} |f - \varphi|^2 < \epsilon \tag{24.45}$$

(was man leicht durch gleichmäßige Approximation zeigt).

· Hinter dieser Zerlegung steckt die folgende “Transitivität der Dichtheit”: Es sei X ein metrischer Raum (vollständig oder nicht) und $Q \subset X$ dicht. Ist dann $P \subset Q$ dicht (in der durch die Einschränkung der Abstandsfunktion gegebenen metrischen Struktur), so ist P ebenfalls dicht in X (Übungsaufgabe). Wir machen von dieser Idee noch verschiedentlich Gebrauch.

· Beispielsweise kann man die stetigen Funktionen auch durch (beliebig oft) differenzierbare ersetzen und statt der Kompaktheit des Trägers “schnellen Abfall bei Unendlich” verlangen. Der Vektorraum³⁹

$$\mathcal{S}(\mathbb{R}^n, \mathbb{C}) = \{f \in \mathcal{C}^\infty(\mathbb{R}^n, \mathbb{C}) \mid \|x^\alpha D^\beta f\|_\infty < \infty \forall \alpha, \beta \in (\mathbb{Z}_{\geq 0})^n\} \tag{24.47}$$

heißt “Schwartz-Raum des \mathbb{R}^n ” und ist Ausgangspunkt vieler beliebiger Untersuchungen.

· Ein vermutlich aus den Physik-Vorlesungen oder -Übungen schon vertrautes Beispiel eines vollständigen Orthonormalsystems sind die Funktionen

$$\chi_k(t) := e^{2\pi i k t} \quad \text{für } k \in \mathbb{Z} \tag{24.48}$$

im Hilbertraum der periodischen Funktionen

$$\begin{aligned} L^2_{\text{per.}}(\mathbb{R}) &:= \{f \in \mathcal{M}(\mathbb{R}, \mathbb{C}) \mid f(t+1) = f(t) \text{ f.ü.}, \int_0^1 |f(t)|^2 dt < \infty\} / \text{Nullfunktionen} \\ &\cong \{f : [0, 1) \rightarrow \mathbb{C} \mid f \text{ messbar und quadratintegrierbar}\} / \text{Nullfunktionen} \end{aligned} \tag{24.49}$$

mit dem inneren Produkt

$$\langle f, g \rangle = \int_0^1 \bar{f}(t)g(t) dt \tag{24.50}$$

Es gilt

$$\langle \chi_k, \chi_l \rangle = \int_0^1 e^{2\pi i(l-k)t} dt = \begin{cases} \frac{1}{2\pi i(l-k)} e^{2\pi i(l-k)t} \Big|_0^1 = 0 & k \neq l \\ \int_0^1 dt = 1 & k = l \end{cases} \tag{24.51}$$

und man zeigt, dass für alle $f \in L^2_{\text{per.}}(\mathbb{R})$ die Fourier-Reihe

$$t \mapsto \sum_{k=-\infty}^{\infty} a_k e^{2\pi i k t} \tag{24.52}$$

³⁹Hier sind α, β sogenannte Multi-Indizes, und

$$x^\alpha D^\beta f(x) := (x^1)^{\alpha_1} \dots (x^n)^{\alpha_n} \cdot (\partial_{x_1})^{\beta_1} \dots (\partial_{x_n})^{\beta_n} f(x) \tag{24.46}$$

in der $L^2_{\text{per.}}$ -Norm gegen f konvergiert, mit den Fourier-Koeffizienten

$$a_k = \langle \chi_k, f \rangle = \int_0^1 f(t) e^{-2\pi i k t} dt \in \mathbb{C} \quad (24.53)$$

Der durch (24.52) gestiftete Isomorphismus

$$L^2_{\text{per.}}(\mathbb{R}) \cong \ell^2(\mathbb{Z}) \quad (24.54)$$

impliziert insbesondere die *Parsevalsche Gleichung*

$$\sum_{k=-\infty}^{\infty} |a_k|^2 = \int_0^1 |f(t)|^2 dt \quad (24.55)$$

Für die historisch und praktisch interessante Frage der punktweisen Konvergenz von (24.52), den faszinierenden Zusammenhang mit der Funktionentheorie *etc. pp.* müssen wir jetzt auf die Lehrbücher (oder den nächsten Durchlauf dieser Vorlesung) verweisen. Ditto für die kontinuierliche Fourier-Transformation auf dem \mathbb{R}^n . Stattdessen beschäftigen wir uns wieder mit der Linearen Algebra.

§ 25 Lineare Operatoren

Hilberträume⁴⁰ bilden also das mathematische Substrat der Quantenmechanik. Da aber je zwei unendlich-dimensionale Hilberträume isomorph zueinander sind (vgl. 24.5), mag man sich fragen, wo denn die physikalische Information eigentlich versteckt ist. Die Antwort liegt bei den *linearen Abbildungen* auf und zwischen den fraglichen Hilberträumen, sowie ihrem Zusammenspiel.

· Bekanntlich ist eine lineare Abbildung eindeutig durch ihre Wirkung auf eine Basis bestimmt. Im endlich-dimensionalen Fall bedarf es dazu einer endlichen Menge an Information. Im unendlich-dimensionalen Fall ist es aber wie bereits betont im Allgemeinen noch nicht einmal möglich, eine Basis explizit anzugeben. Im Beisein einer Norm kann man sich aber schon mit weit weniger begnügen. Wir ergänzen dazu zunächst einige Definitionen und Aussagen aus § 4.

Definition 25.1. Ein *Banach-Raum* ist ein normierter Vektorraum $(V, \|\cdot\|)$ (über \mathbb{R} oder \mathbb{C}), der mit der durch die Norm induzierten Abstandsfunktion ein vollständiger metrischer Raum ist.

Beispiele: Endlich-dimensionale Vektorräume.

- Der Vektorraum der beschränkten stetigen Funktionen auf einem metrischen Raum mit der Supremumsnorm ist ein Banachraum, s. 8.2.
- Gemäss 24.1 ist ein Banach-Raum genau dann ein Hilbertraum, wenn die Norm die Parallelogrammgleichung erfüllt.
- Ist V ein Banachraum, und $U \subset V$ ein Unterraum, so induziert man durch Einschränkung der Norm von V eine Norm auf U . Mit dieser Norm ist U genau dann

⁴⁰Es interessieren wie gesagt nur separable Hilberträume und wir lassen dieses Adjektiv jetzt weg.

§ 25. LINEARE OPERATOREN

ein Banachraum, wenn er als Teilmenge des metrischen Raums V abgeschlossen ist. Der Abschluss \overline{U} von U ist der (eindeutige) kleinste abgeschlossene Unterraum von V mit $U \subset \overline{U}$.

· Umgekehrt lässt sich ein a priori nicht vollständiger normierter Vektorraum U in (bis auf Isomorphismus) eindeutiger Weise zu einem Banachraum vervollständigen: Man betrachtet dazu auf dem Vektorraum

$$U_0^{\mathbb{N}} := \{(u_k)_{k \in \mathbb{N}} \subset U \mid (u_k)_{k \in \mathbb{N}} \text{ ist Cauchy-Folge}\} \quad (25.1)$$

aller Cauchy-Folgen in U die Äquivalenzrelation

$$(u_k) \sim (u'_k) \iff \begin{array}{l} \text{Die Reissverschlussfolge } (u_1, u'_1, u_2, u'_2, \dots) \\ \text{ist eine Cauchy-Folge.} \end{array} \quad (25.2)$$

und definiert

$$\overline{U} = U_0^{\mathbb{N}} / \sim \quad (25.3)$$

mit der (wohl-definierten) Norm

$$\|[(u_k)]\| := \lim_{k \rightarrow \infty} \|u_k\| \quad (25.4)$$

(Äquivalent dazu führt man zunächst durch $\|(u_k)\| = \lim_{k \rightarrow \infty} \|u_k\|$ auf $U_0^{\mathbb{N}}$ eine sog. Halbnorm ein und teilt durch den Unterraum der Nullfolgen, auf dem $\|\cdot\| = 0$. Unsere Konstruktion von $L^2(\mathbb{R}^n)$ aus $\mathcal{C}_c(\mathbb{R}^n, \mathbb{C})$ ist eine Hilbertraum-Version dieser Ideen.)
Wir wissen ausserdem: Eine Abbildung $f : X \rightarrow Y$ zwischen metrischen Räumen $X \rightarrow Y$ ist genau dann stetig, wenn für jede Folge $(x_k) \subset X$, welche in X konvergiert, die Bildfolge $(f(x_k))$ in Y konvergiert.

Lemma/Definition 25.2. *Es seien U und W normierte Vektorräume (nicht notwendig vollständig). Dann ist eine lineare Abbildung $A : U \rightarrow W$ genau dann stetig, wenn sie beschränkt ist, das heisst,*

$$\|A\| := \sup\{\|A(u)\| \mid \|u\| = 1\} < \infty \quad (25.5)$$

Beweis. " \Rightarrow ": Ist A beschränkt, so gilt (leichte Übung...)

$$\|A(u_1) - A(u_2)\| = \|A(u_1 - u_2)\| \leq \|A\| \cdot \|u_1 - u_2\| \quad (25.6)$$

A ist also sogar Lipschitz-stetig, insbesondere stetig.

" \Leftarrow ": Angenommen, A ist nicht beschränkt. Dann gibt es eine Folge $(u_k) \subset U$ mit $\|u_k\| = 1$ aber $\|Au_k\| \geq k$. Dann konvergiert $(\tilde{u}_k) = (\frac{1}{k}u_k)$ gegen 0, aber $\|A(\tilde{u}_k)\|$ konvergiert nicht. A ist also auch nicht stetig. \square

Proposition 25.3 (B.L.T.). *Es sei V ein Banach-Raum, und U ein dichter Unterraum. Ist $A : U \rightarrow W$ eine beschränkte lineare Abbildung in einen Banachraum W , so existiert eine eindeutige Fortsetzung von A zu einer beschränkten linearen Abbildung $\overline{A} : V \rightarrow W$. Es gilt $\|\overline{A}\| = \|A\|$.*

Beweis. Für jedes $v \in V$ existiert eine Folge $(u_k) \subset U$ mit $\lim u_k = v$ (in V). (u_k) ist eine Cauchy-Folge und da A stetig ist, ist $(A(u_k))$ ebenfalls eine Cauchy-Folge. Da W vollständig ist, existiert

$$\overline{A}(v) := \lim_{k \rightarrow \infty} A(u_k) \tag{25.7}$$

Man sieht, dass diese Definition nicht von der approximierenden Folge abhängt. Für $u \in U$ geht offensichtlich $u_k = v \forall k$, d.h. $\overline{A}(u) = A(u)$. Linearität und die Aussage über die Norm von \overline{A} folgt aus der Stetigkeit von Addition, Skalarmultiplikation, und Norm. \square

Wie wir gleich sehen werden, kommen wir in der Physik mit *stetigen* linearen Abbildungen nicht aus. Man definiert daher allgemeiner:⁴¹

Proposition/Definition 25.4. *Ein linearer Operator auf einem Banachraum V ist ein (sinnvollerweise dichter) Unterraum $\mathcal{D}(A) \subset V$ zusammen mit einer linearen Abbildung $A : \mathcal{D}(A) \rightarrow W$.*

- Wir schreiben $\mathcal{B}(V, W)$ für den Vektorraum der beschränkten linearen Operatoren auf V , nach 25.3 oBdA mit $\mathcal{D}(A) = V \forall A \in \mathcal{B}(V, W)$.
- Mit der Operatornorm (25.5) ist $\mathcal{B}(V, W)$ ein vollständiger normierter Raum.
- Die Operatornorm ist submultiplikativ, d.h. im Falle $W = V$ ist $\mathcal{B}(V) := \mathcal{B}(V, V)$ eine Banach-Algebra.

Beweis. Die Argumente sind die gleichen wie im endlich-dimensionalen Fall, wofür wir auf 4.28 verweisen. Stattdessen bemerken wir noch, dass es auch unbeschränkte Operatoren gibt, die auf einem ganzen Banachraum definiert sind. Sie lassen sich aber (ähnlich wie überabzählbare Vektorraumbasen) nicht explizit vorzeigen. (Und spielen bei Hilberträumen auch keine Rolle.)

- Es sollte auch bemerkt werden, dass sich bei der Addition/Multiplikation von unbeschränkten Operatoren die Definitionsbereiche weiter ändern können. \square

Beispiele: Sei X ein metrischer Raum und $(\mathcal{C}^b(X), \|\cdot\|_\infty)$ der Banachraum der beschränkten stetigen Funktionen. Dann ist für jedes $f_0 \in \mathcal{C}^b(X)$ die Multiplikation mit f_0 ein beschränkter linearer Operator:

$$m_{f_0} : f \mapsto f_0 \cdot f \tag{25.8}$$

mit Norm

$$\|m_{f_0}\| = \|f_0\|_\infty \tag{25.9}$$

- Multiplikation mit einer unbeschränkten Funktion ist aber im Allgemeinen nicht beschränkt, so dass wir den Definitionsbereich einschränken müssen. Beispielsweise macht für $V = \ell^2(\mathbb{Z})$ die Vorschrift

$$\delta((a_k)_{k \in \mathbb{Z}}) := (ka_k)_{k \in \mathbb{Z}} \tag{25.10}$$

⁴¹Der Sprachgebrauch ist ähnlich zu den meromorphen Funktionen in 17.3, die ebenfalls nicht auf dem gesamten Definitionsbereich definiert sind.

§ 25. LINEARE OPERATOREN

genau dann Sinn, wenn $\delta((a_k)_{k \in \mathbb{Z}}) \in \ell^2(\mathbb{Z})$, d.h. wenn

$$(a_k)_{k \in \mathbb{Z}} \in \mathcal{D}(\delta) := \left\{ (a_k)_{k \in \mathbb{Z}} \mid \sum_{k=-\infty}^{\infty} |ka_k|^2 < \infty \right\} \quad (25.11)$$

Die Folge $a_k = \frac{1}{|k|+1}$ macht klar, dass $\mathcal{D}(\delta) \subsetneq \ell^2(\mathbb{Z})$. Da aber offensichtlich $\mathcal{D}(\delta) \supset \mathcal{F}^{\text{fin}}(\mathbb{Z}, \mathbb{C})$, und Letzteres dicht in $\ell^2(\mathbb{Z})$ liegt, ist auch $\mathcal{D}(\delta) \subset \ell^2(\mathbb{Z})$ dicht.

· Unter dem Fourier-Isomorphismus $i : L_{\text{per.}}^2(\mathbb{R}) \rightarrow \ell^2(\mathbb{Z})$, $i(e^{2\pi ikt}) = e_k$ (s. (24.54)),

$$(i^{-1} \circ \delta \circ i)(e^{2\pi ikt}) = ke^{2\pi ikt} = \frac{1}{2\pi i} \frac{d}{dt} e^{2\pi ikt} \quad (25.12)$$

wird δ zu einer *Ableitung*. Differentialoperatoren sind im Allgemeinen also auch nicht beschränkt.

· In der elementaren Quantenmechanik lernt man insbesondere die Orts- und Impulsoperatoren kennen, üblicherweise in ihrer Darstellung auf stetig differenzierbare Wellenfunktionen $\psi(x)$, $x \in \mathbb{R}$

$$(q\psi)(x) := x\psi(x), \quad (p\psi)(x) := \frac{\hbar}{i} \cdot \frac{d\psi}{dx}(x) \quad (25.13)$$

Diese Operatoren erfüllen (zumindest formal) die *kanonischen Vertauschungsrelationen*

$$[q, p] = q \circ p - p \circ q = i\hbar \cdot \text{id}, \quad (25.14)$$

(Stichwort: Heisenbergsche Unschärferelation), sind aber offenbar nicht beschränkt: Die rechten Seiten in (25.13) sind i.A. nicht quadratintegrierbar, selbst wenn ψ es ist. Der Schwartzraum (24.47) ist ein bequemer dichter Unterraum, auf dem beide Operatoren und alle ihre Produkte wohldefiniert sind.

· Die Notwendigkeit dieser Unbeschränktheit lässt sich etwa folgendermassen beleuchten:

Lemma 25.5. *Es sei \mathcal{A} eine normierte \mathbb{C} -Algebra mit Einselement und submultiplikativer Norm 4.29. Dann gilt (25.14) für keine $p, q \in \mathcal{A}$. Insbesondere gibt es keine endlich-dimensionalen Darstellungen der KVRn.*

(Es wird hier *nicht* gesagt, dass die Gleichung (25.14) keine Lösung hat, die Operatoren in (25.13) geben ja ein explizites Beispiel. Der Punkt ist nur, dass Operatoren, welche (25.14) erfüllen, nicht beschränkt sein können.)

Beweis. Angenommen, (25.14) gälte doch, und zwar oBdA mit $\hbar = -i$. Aus der Derivationseigenschaft des Kommutators folgt für $n = 1, 2, \dots$

$$\begin{aligned} [q, p^n] &= (qp)p^{n-1} - pqp^{n-1} + pqp^{n-1} - pp^{n-1}q \\ &= [q, p]p^{n-1} + p[q, p^{n-1}] = \dots \\ &= \sum_{i=0}^{n-1} p^i [q, p] p^{n-i} \\ &= np^{n-1} \end{aligned} \quad (25.15)$$

Daraus folgt einerseits induktiv, dass $p^n \neq 0 \forall n$, andererseits aus den Eigenschaften der Norm

$$\begin{aligned} n \cdot \|p^{n-1}\| &\leq 2\|q\| \cdot \|p^n\| \leq 2\|q\| \cdot \|p\| \cdot \underbrace{\|p^{n-1}\|}_{\neq 0} \\ \Rightarrow n &\leq 2\|q\| \cdot \|p\| \end{aligned} \tag{25.16}$$

für alle n , was unmöglich ist. □

Unser Ziel ist es nun, einige physikalisch bedeutsame Verallgemeinerungen von Aussagen aus Kap. 7 auf unendlich-dimensionale Hilberträume zu sammeln. Wir beginnen mit (20.5).

Theorem 25.6 (Darstellungssatz von Riesz). *Es sei $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ ein Hilbertraum. Dann ist der Raum der stetigen linearen Abbildungen (a.k.a. Funktionale)*

$$\mathcal{B}(\mathcal{H}, \mathbb{C}) = \{\lambda : \mathcal{H} \rightarrow \mathbb{C} \mid \|\lambda\| < \infty\} \tag{25.17}$$

mit der Operatornorm ein (separabler!) Hilbertraum. Die Abbildung

$$\mathcal{H} \ni v \mapsto (v^\flat : w \mapsto \langle v, w \rangle) \in \mathcal{B}(\mathcal{H}, \mathbb{C}) \tag{25.18}$$

ist ein Anti-Isomorphismus von Hilberträumen. Man schreibt (nicht konform mit 18.5) $\mathcal{H}^* \cong \mathcal{H}$ und nach Dirac (vgl. (24.33))

$$(|\text{blah}\rangle)^\flat = \langle \text{blah} | \tag{25.19}$$

Vor dem Beweis halten wir zunächst die Eigenschaften von orthogonalen Zerlegungen von Hilberträumen fest.

Proposition/Definition 25.7. *Für eine Teilmenge $M \subset \mathcal{H}$ ist*

$$M^\perp := \{v \in \mathcal{H} \mid \langle m, v \rangle = 0 \forall m \in M\} \tag{25.20}$$

ein abgeschlossener Unterraum von \mathcal{H} . Ist $M = U$ selbst ein abgeschlossener Unterraum, so ist

$$\mathcal{H} = U \oplus U^\perp \tag{25.21}$$

d.h. jedes $v \in \mathcal{H}$ hat eine eindeutige Zerlegung $v = v_\parallel + v_\perp$ mit $v_\parallel \in U$, $v_\perp \in U^\perp$.

Beweis. M^\perp ist klarerweise abgeschlossen unter den Vektorraumoperationen. Ist $(v_k) \subset M^\perp$ eine Folge, welche in \mathcal{H} gegen v konvergiert, dann gilt wegen der Stetigkeit des inneren Produkts (welche aus der C.S.U. folgt), $\forall m \in M$:

$$\langle v, m \rangle = \langle \lim_{k \rightarrow \infty} v_k, m \rangle = \lim_{k \rightarrow \infty} \langle v_k, m \rangle = 0 \tag{25.22}$$

also $v \in M^\perp$, d.h. M ist abgeschlossen.

· Da wegen der Nichtentartetheit des inneren Produkts offenbar $U \cap U^\perp = \{0\}$, müssen wir nur noch zeigen, dass $\mathcal{H} = U + U^\perp$. Für $v \in \mathcal{H}$ sei

$$d(v, U) = \inf\{\|v - u\| \mid u \in U\} \tag{25.23}$$

§ 25. LINEARE OPERATOREN

(der “Abstand” von v zu U) und $(u_k) \subset U$ eine Folge mit

$$d(v, U) = \lim_{k \rightarrow \infty} \|v - u_k\| \quad (25.24)$$

Dann gilt mit der Parallelogrammgleichung (24.3)

$$\begin{aligned} \|u_k - u_l\|^2 &= \|(u_k - v) - (u_l - v)\|^2 \\ &= 2(\|u_k - v\|^2 + \|u_l - v\|^2) - \underbrace{\|u_k + u_l - 2v\|^2}_{=4\|\frac{u_k+u_l}{2}-v\|^2 \geq 4d^2} \\ &\leq 2(\|u_k - v\|^2 + \|u_l - v\|^2) \end{aligned} \quad (25.25)$$

Mit $k, l \rightarrow \infty$ folgt, dass (u_k) eine Cauchy-Folge ist. Ihr Grenzwert $\lim u_k =: u_{\parallel}$ liegt dann wegen der Abgeschlossenheit in U , erfüllt

$$\|v - u_{\parallel}\| = d(v, U) \quad (25.26)$$

und hängt mit dieser Eigenschaft nicht von der Folge (u_k) ab (vorausgesetzt, sie erfüllt (25.24), Übung). Wir setzen $u_{\perp} = v - u_{\parallel}$ und behaupten $u_{\perp} \in U^{\perp}$. Dazu betrachten wir für $u \in U$ und $t \in \mathbb{R}$

$$\begin{aligned} \|u_{\perp}\|^2 &= d(v, U)^2 \leq \|v - u_{\parallel} + tu\|^2 \quad (\text{Def. von } d(v, U), -u_{\parallel} + tu \in U) \\ &= \|u_{\perp}\|^2 + 2t \operatorname{Re}\langle v - u_{\parallel}, u \rangle + t^2 \|u\|^2 \end{aligned} \quad (25.27)$$

Es folgt $\operatorname{Re}\langle u_{\perp}, u \rangle = 0$ (sonst könnten wir für kleines t mit geeignetem Vorzeichen die Ungleichung verletzen), und analog (vgl. Beweis von 24.1) $\operatorname{Im}\langle u_{\perp}, u \rangle = 0 \forall u \in U$, zusammen also $u_{\perp} \in U^{\perp}$. \square

Bemerkungen. Beachte, dass die Abgeschlossenheit hier wesentlich ist. Beispielsweise ist für eine Hilbertbasis (b_k) die endliche lineare Hülle $\operatorname{span}_{\mathbb{C}}(b_k)$ nicht abgeschlossen (sondern nur dicht in \mathcal{H}), aber $(b_k)^{\perp} = \{0\}$.

Beweis von 25.6. Es genügt offenbar zu zeigen, dass (25.18) eine bijektive normerhaltende konjugiert lineare Abbildung ist. (Denn dann gilt für die Norm auf $\mathcal{B}(\mathcal{H}, \mathbb{C})$ insbesondere die Parallelogrammgleichung.)

· Die Additivität in v ist klar, ebenso die Anti-Linearität. Die Injektivität folgt aus der Nichtentartetheit der Norm.

· Aus der Cauchy-Schwarz-Ungleichung folgt

$$|v^{\flat}(w)| = |\langle v, w \rangle| \leq \|v\| \cdot \|w\| \quad (25.28)$$

d.h. $\|v^{\flat}\| \leq \|v\|$, wegen $v^{\flat}(v) = \|v\|^2$ tatsächlich $\|v^{\flat}\| = \|v\|$. \cdot^{\flat} ist also normerhaltend.

· Zur Surjektivität sei $\lambda \in \mathcal{B}(\mathcal{H}, \mathbb{C})$, mit $\lambda \neq 0$. Wegen der Stetigkeit von λ ist

$$U := \operatorname{Ker}(\lambda) \quad (25.29)$$

ein abgeschlossener Unterraum von \mathcal{H} , und wegen $\lambda \neq 0$ ist $U \subsetneq \mathcal{H}$, d.h. in der Zerlegung $\mathcal{H} = U \oplus U^{\perp}$ ist $U^{\perp} \neq \{0\}$. Es existiert daher ein $v_0 \in U^{\perp} \setminus \{0\}$, und es ist auch $\lambda(v_0) \neq 0$. Für jedes $w \in \mathcal{H}$ ist

$$\tilde{w} := w - \frac{\lambda(w)}{\lambda(v_0)} v_0 \in \operatorname{Ker}(\lambda) = U \quad (25.30)$$

(Denn $\lambda(\tilde{w}) = \lambda(w) - \frac{\lambda(w)}{\lambda(v_0)}\lambda(v_0) = 0$.) Da $v_0 \perp U$ gilt also

$$0 = \langle v_0, \tilde{w} \rangle = \langle v_0, w \rangle - \lambda(w) \frac{\langle v_0, v_0 \rangle}{\lambda(v_0)} \quad (25.31)$$

d.h. $\forall w \in \mathcal{H}$

$$\lambda(w) = \left\langle \frac{\overline{\lambda(v_0)}}{\|v_0\|^2} v_0, w \right\rangle \quad (25.32)$$

Es ist also $\lambda^\sharp := \frac{\overline{\lambda(v_0)}}{\|v_0\|^2} v_0 \in \mathcal{H}$ ein Urbild von λ unter (25.18), $\lambda = (\lambda^\sharp)^\flat$. \square

Bemerkungen. Im endlich-dimensionalen Fall kann man aus der Injektivität und Dimensionsbetrachtungen direkt auf Surjektivität schliessen. Auf $\ell^2(\mathbb{N})$ ist beispielsweise die durch $e_k \mapsto 1 \ \forall k$ dicht definierte, aber unbeschränkte, Linearform nicht von der Form $\langle (a_k), \cdot \rangle$ für ein $(a_k) \in \ell^2(\mathbb{N})$.

· Ein weiteres erhellendes Beispiel ist die lineare Fortsetzung der Abbildung (Stichwort: Hilbert-Hotel)

$$T(e_k) := e_{k+1} \quad (25.33)$$

auf $\ell^2(\mathbb{N})$. T ist dicht definiert und stetig, die Fortsetzung auf ganz $\ell^2(\mathbb{N})$ ist

$$T((a_1, a_2, a_3, \dots)) = (0, a_1, a_2, \dots) \quad (25.34)$$

Man sieht leicht, dass T injektiv, aber nicht surjektiv ist.

· In Fortsetzung der Diskussion um 18.6 definiert man für zwei (“verschiedene”) Hilberträume $\mathcal{H}_1, \mathcal{H}_2$ ihre “äussere” oder orthogonale direkte Summe als die gewöhnliche direkte Summe mit dem inneren Produkt

$$\langle v_1 \oplus v_2, w_1 \oplus w_2 \rangle := \langle v_1, v_2 \rangle + \langle w_1, w_2 \rangle \quad (25.35)$$

(insbesondere ist also $\mathcal{H}_1 \perp \mathcal{H}_2$ in dieser Summe). Man prüft leicht, dass diese Definition Sinn macht, d.h. in einem Hilbertraum resultiert.

· Andererseits ist das algebraische Tensorprodukt von zwei Hilberträumen mit

$$\langle v_1 \otimes v_2, w_1 \otimes w_2 \rangle := \langle v_1, w_1 \rangle \cdot \langle v_2, w_2 \rangle \quad (25.36)$$

i.A. zwar ein unitärer Raum, aber nicht vollständig, wenn \mathcal{H}_1 und \mathcal{H}_2 beide unendlich-dimensional sind. (Man mache sich dies am Beispiel von $\mathcal{H}_1 = \mathcal{H}_2 = \ell^2(\mathbb{N})$ klar!) Das “richtige” Tensorprodukt von Hilberträumen erhält man durch Vervollständigung im Sinne von 25.3. Es gilt dann beispielsweise

$$L^2(\mathbb{R}^n) \overline{\otimes} L^2(\mathbb{R}^m) \cong L^2(\mathbb{R}^{n+m}) \quad (25.37)$$

(In der Praxis macht dies natürlich keinen grossen Unterschied: Sind $(b_k^{(1)})_{k \in \mathbb{N}}$, $(b_l^{(2)})_{l \in \mathbb{N}}$ Hilbertbasen von $\mathcal{H}_1, \mathcal{H}_2$, so ist $(b_k^{(1)} \otimes b_l^{(2)})_{k, l \in \mathbb{N}}$ eine Hilbertbasis von $\mathcal{H}_1 \overline{\otimes} \mathcal{H}_2$.)

§ 25. LINEARE OPERATOREN

· **Achtung:** Die direkte Summe bedeutet physikalisch die Zerlegung des Zustandsraums eines gegebenen Systems in experimentell (z.B. durch verschiedene Quantenzahlen) unterschiedene Gegebenheiten, während wie bereits betont das Tensorprodukt für die Zusammensetzung verschiedener Systeme zu einem grösseren⁴² zuständig ist.

· Natürlich kann man hier auch wieder symmetrische und anti-symmetrische Tensorprodukte und Multilinearformen in Verbindung setzen.

Physikalisch stellt 25.6 eine Korrespondenz zwischen Präparation (den Kets $|\psi\rangle \in \mathcal{H}$) und Registrierung (den Bras $\langle\varphi| \in \mathcal{H}^* \cong \mathcal{H}$) von quantenmechanischen Zuständen her, mit einer durch das Bra-Ket Produkt $\langle\varphi|\psi\rangle$ gegebenen “Übergangsamplitude” (die Wahrscheinlichkeit ist $\frac{|\langle\varphi|\psi\rangle|^2}{\langle\varphi|\varphi\rangle \cdot \langle\psi|\psi\rangle}$). Dabei steckt die eigentliche Information über den physikalischen Prozess in einer zwischengeschalteten unitären Transformation von \mathcal{H} , bzw., in der üblichen kontinuierlichen Situation, ihrer infinitesimal Erzeugenden. (Beispiel: der Hamiltonoperator $H = i\hbar \frac{d}{dt}U(t)$ als Ableitung der Zeitevolution) Am Höhepunkt längerer Überlegungen interpretiert man einen Ausdruck

$$\mathbb{E}_\psi(A) = \langle\psi, A\psi\rangle \quad \text{mit } \langle\psi, \psi\rangle = 1 \quad (25.38)$$

(üblicherweise als Sandwich $\langle\psi|A|\psi\rangle$ geschrieben) als “statistischen Erwartungswert der zum linearen Operator A assoziierten Observablen im (normierten) Zustand $|\psi\rangle$ ”. Damit dies sinnvoll ist, muss A gewisse Eigenschaften erfüllen, die wir in der endlich-dimensionalen Situation z.T. bereits im § 20 untersucht hatten, siehe 20.2 und 20.8 und vgl. auch das Auftauchen der Form (25.38) in 20.9. In den unendlichen Dimensionen eines allgemeinen separablen Hilbertraums ist Platz für einige weitere Neuheiten. Als Erstes müssen wir die in der Fussnote 5 gegebene Definition an die Beobachtung 25.5 anpassen, dass physikalisch relevante Operatoren nicht notwendig beschränkt sind.

Lemma/Definition 25.8. *Es sei \mathcal{H} ein Hilbertraum und $A : \mathcal{D}(A) \rightarrow \mathcal{H}$ ein dicht definierter linearer Operator, d.h. $\overline{\mathcal{D}(A)} = \mathcal{H}$.*

(i) *Für gegebenes $v \in \mathcal{H}$ existiert höchstens ein $v^A \in \mathcal{H}$ so, dass*

$$\langle v^A, w \rangle = \langle v, Aw \rangle \quad \forall w \in \mathcal{D}(A) \quad (25.39)$$

(ii) *Die Menge der $v \in \mathcal{H}$, für die v^A existiert, ist ein linearer Unterraum von \mathcal{H} und die Zuordnung $v \mapsto v^A$ ist linear in v . Man nennt*

$$\begin{aligned} \mathcal{D}(A^\dagger) &:= \{v \in \mathcal{H} \mid \exists v^A \in \mathcal{H} : \langle v^A, w \rangle = \langle v, Aw \rangle \forall w \in \mathcal{H}\} \\ A^\dagger(v) &:= \text{eben jenes } v^A \end{aligned} \quad (25.40)$$

den zu A adjungierten Operator.

(iii) *Ist $A \in \mathcal{B}(\mathcal{H})$ beschränkt, so ist $\mathcal{D}(A^\dagger) = \mathcal{H}$, und in diesem Fall gilt $(A^\dagger)^\dagger = A$ sowie $\|A^\dagger\| = \|A\|$.*

Beweis. (i) Erfüllen v^A und \tilde{v}^A beide (25.39), dann gilt $\langle v^A - \tilde{v}^A, w \rangle = 0 \forall w \in \mathcal{D}(A)$, d.h. $v^A - \tilde{v}^A \in \mathcal{D}(A)^\perp$. Aus $\overline{\mathcal{D}(A)} = \mathcal{H}$ folgt mit 25.7, dass $v^A - \tilde{v}^A = 0$.

⁴²Wie die Formeln zeigen, ist das Tensorprodukt von separablen Hilberträumen wieder separabel.

- (ii) Übungsaufgabe
 (iii) Für jedes feste $v \in \mathcal{H}$ ist die Abbildung

$$\mathcal{H} \ni w \mapsto \langle v, Aw \rangle \in \mathbb{C} \quad (25.41)$$

eine beschränkte Linearform, denn wegen der C.S.U. und $A \in \mathcal{B}(\mathcal{H})$ ist

$$|\langle v, Aw \rangle| \leq \|v\| \cdot \|A\| \cdot \|w\| \quad (25.42)$$

Wegen 25.6 existiert daher ein $v^A = A^\dagger(v)$ wie gefordert.

• Aus der für jeden stetigen linearen Operator gültigen Aussage, dass⁴³

$$\|A\| = \sup\{|\langle v, Aw \rangle| \mid \|v\| = \|w\| = 1\} \quad (25.45)$$

folgt zunächst

$$\|A^\dagger\| \leq \|A\| \quad (25.46)$$

A^\dagger ist also insbesondere beschränkt. Aus

$$\langle (A^\dagger)^\dagger v, w \rangle = \langle v, A^\dagger w \rangle = \overline{\langle A^\dagger w, v \rangle} = \overline{\langle w, Av \rangle} = \langle Av, w \rangle \quad (25.47)$$

für alle $v, w \in \mathcal{H}$ folgt dann $(A^\dagger)^\dagger = A$ und zuletzt durch Wiederholung von (25.46) $\|A\| = \|(A^\dagger)^\dagger\| \leq \|A^\dagger\|$. \square

Die Definition in Fussnote 6 bleibt im Wesen unverändert,

Definition 25.9. Ein linearer Operator U heisst unitär, falls $U^\dagger U = U U^\dagger = \text{id}_{\mathcal{H}}$.

man muss nur wieder darauf aufpassen, dass

$$\text{injektiv} \Leftrightarrow \text{surjektiv} \quad (25.48)$$

Beispiel. Für den in (25.33) definierten Operator T gilt

$$\langle (b_k), T(a_k) \rangle = 0 + \sum_{k=2}^{\infty} \bar{b}_k a_{k-1} = \sum_{k=1}^{\infty} \bar{b}_{k+1} a_k \quad (25.49)$$

d.h.

$$T^\dagger(b_1, b_2, b_3, \dots) = (b_2, b_3, b_4, \dots) \quad (25.50)$$

⁴³Einerseits ist

$$\sup\{|\langle v, Aw \rangle| \mid \|v\| = \|w\| = 1\} \leq \sup\{\|v\| \cdot \|Aw\| \mid \|v\| = \|w\| = 1\} = \|A\|, \quad (25.43)$$

andererseits (jedenfalls für $A \neq 0$)

$$\begin{aligned} \|A\| &= \sup\{\|Aw\| \mid \|w\| = 1, Aw \neq 0\} \\ &= \sup\left\{\left\langle \frac{Aw}{\|Aw\|}, Aw \right\rangle \mid \|w\| = 1, Aw \neq 0\right\} \\ &\leq \sup\{|\langle v, Aw \rangle| \mid \|v\| = \|w\| = 1\} \end{aligned} \quad (25.44)$$

§ 25. LINEARE OPERATOREN

Man sieht insbesondere $T^\dagger \circ T = \text{id}_{\ell^2(\mathbb{N})}$, aber $T \circ T^\dagger = \text{id}_{\ell^2(\mathbb{N})} - P_1$, wobei P_1 der sog. *Projektionsoperator* auf die erste Komponente ist

$$P_1((a_k)) = (a_1, 0, \dots) \quad (25.51)$$

TT^\dagger hat also nicht-trivialen Kern.

· Übungsaufgabe: U ist unitär $\Leftrightarrow \langle v, v \rangle = \langle Uv, Uv \rangle = \langle U^\dagger v, U^\dagger v \rangle \forall v$, d.h. U und U^\dagger sind beide *Isometrien*. (Beschränktheit ist dann automatisch.)

Bei den infinitesimal Erzeugenden/Observablen spiegelt sich (25.48) in einer wichtigen Konsequenz der Tatsache wider, dass für unbeschränkte Operatoren die Definitionsbereiche $\mathcal{D}(A)$ und $\mathcal{D}(A^\dagger)$ im Allgemeinen *verschieden* sind.

Definition 25.10. Ein linearer Operator A heisst symmetrisch (oder hermitesch), falls $\mathcal{D}(A) \subset \mathcal{D}(A^\dagger)$ und $A^\dagger|_{\mathcal{D}(A)} = A$.

Lemma 25.11. Ein linearer Operator A ist genau dann symmetrisch, wenn die assoziierte quadratische⁴⁴ Form, nämlich

$$\mathcal{D}(A) \ni v \mapsto q_A(v) := \langle v, Av \rangle \quad (25.52)$$

reell ist.

Beweis. · Ist A symmetrisch, so ist $v \in \mathcal{D}(A)$ insbesondere auch in $\mathcal{D}(A^\dagger)$ und es gilt

$$\overline{\langle v, Av \rangle} = \overline{\langle Av, v \rangle} = \langle v, Av \rangle \quad (25.53)$$

d.h. $q_A(v) \in \mathbb{R}$.

· Es sei umgekehrt q_A reell. Aus der Identität (vgl. (24.8))

$$\langle v, Aw \rangle = \frac{1}{4} (q_A(v+w) - q_A(v-w) + iq_A(v-iw) - iq_A(v+iw)) \quad (25.54)$$

$\forall v, w \in \mathcal{D}(A)$ folgt leicht $\langle w, Av \rangle = \overline{\langle v, Aw \rangle} = \langle Aw, v \rangle$, und daraus die Symmetrie von A . \square

Symmetrische Operatoren erfüllen also die physikalisch sinnvolle Bedingung, dass Erwartungswerte von Observablen reell sind. Symmetrie ist aber *nicht* hinreichend dafür, dass (i mal) der Operator in einer geeigneten Verallgemeinerung von 20.8 über die Exponentialfunktion eine Einparametergruppe unitärer Symmetrietransformationen von \mathcal{H} erzeugt.

Definition 25.12. Ein symmetrischer linearer Operator A heisst selbst-adjungiert, wenn $\mathcal{D}(A^\dagger) = \mathcal{D}(A)$.

Beispiel 25.13. · Multiplikation mit einer (vernünftigen, beispielsweise stetigen) reellen Funktion f definiert einen selbst-adjungierten Operator auf $L^2(\mathbb{R}^n)$. Definitionsbereich ist

$$\mathcal{D}(m_f) = \{ \psi \in L^2(\mathbb{R}^n) \mid f \cdot \psi \in L^2(\mathbb{R}^n) \} = \mathcal{D}(m_f^\dagger) \quad (25.55)$$

⁴⁴Man sollte wohl eher absolut quadratisch sagen, denn es gilt $q_A(\lambda v) = |\lambda|^2 q_A(v)$.

der sicherlich $\mathcal{C}_c(\mathbb{R}^n)$ enthält, und daher dicht ist.

· Der auf $\mathcal{D}(p) = \mathcal{C}_c^1(\mathbb{R})$ durch

$$p\psi(x) := \frac{\hbar}{i} \frac{d\psi}{dx}(x) \tag{25.56}$$

definierte Operator p ist symmetrisch: Für $\psi, \varphi \in \mathcal{C}_c^1(\mathbb{R})$ gilt

$$\begin{aligned} \int_{\mathbb{R}} \bar{\varphi}(x) \frac{\hbar}{i} \frac{d\psi}{dx}(x) dx &= - \int_{\mathbb{R}} \frac{\hbar}{i} \frac{d\bar{\varphi}}{dx}(x) \psi(x) dx \quad (\text{partielle Integration}) \\ &= \int_{\mathbb{R}} \overline{\frac{\hbar}{i} \frac{d\varphi}{dx}}(x) \psi(x) dx \end{aligned} \tag{25.57}$$

Zunächst ist zwar $\mathcal{D}(p^\dagger) \supsetneq \mathcal{D}(p)$ (beispielsweise geht die Rechnung (25.57) sicher auch durch, wenn φ nur stetig differenzierbar ist, aber nicht-kompakten Träger hat), man kann aber zeigen, dass es eine (eindeutige) selbst-adjungierte Fortsetzung gibt. Die Idee ist, dass bei Vergrößerung von $\mathcal{D}(p)$ der Definitionsbereich des adjungierten Operators kleiner wird. (Denn es sind mehr Bedingungen der Art (25.39) zu erfüllen.) Es ist aber nicht-trivial, dass man sich in der Mitte treffen kann.

· Die **formale Rechnung** für $t \in \mathbb{R}$,

$$\left(\exp\left(\frac{i}{\hbar}tp\right)\psi\right)(x) = \sum_{k=0}^{\infty} \frac{1}{k!} \left(t \frac{d}{dx}\right)^k \psi(x) = \psi(x) + t\psi'(x) + \frac{t^2}{2}\psi''(x) + \dots = \psi(x+t) \tag{25.58}$$

(Taylorreihe) weist $\frac{i}{\hbar}p$ als infinitesimalen Erzeuger der Translationen

$$(T_t\psi)(x) = \psi(x+t) \tag{25.59}$$

aus. (Etwas formaler argumentiert man mit der Differentialgleichung (s. (20.27))

$$\frac{d}{dt}(T_t(\psi)(x)) = \frac{d}{dx}(T_t(\psi))(x) = \frac{i}{\hbar}p(T_t(\psi))(x) \tag{25.60}$$

T_t ist klarerweise(?) auf ganz $L^2(\mathbb{R})$ fortsetzbar und unitär.

· Selbst-adjungiertheit ist notwendig und hinreichend für die Erzeugung von unitären Transformationen über die Exponentialabbildung, was wir hier allerdings nicht beweisen können.

· Ein erhellendes Unbeispiel ist der Impulsoperator auf der *Halbgeraden* $[0, \infty)$. Wäre er selbst-adjungiert, so würde er nach dem eben Zitierten unitäre Transformationen erzeugen. Auf der Halbgeraden ist aber die Translation T_t nach links ($t > 0$) nicht injektiv, nach rechts $t < 0$ nicht surjektiv:

$$T_2(\chi_{[0,1]}) = 0 \tag{25.61}$$

Die diskrete Version war gerade (25.33).

· Die Untersuchung der Symmetrie von $\frac{\hbar}{i} \frac{d}{dx}$ auf Funktionen auf $[0, \infty)$ mittels partieller Integration liefert den Randterm

$$\frac{\hbar}{i} \bar{\varphi}(0)\psi(0) \tag{25.62}$$

§ 26. STURM-LIOUVILLE-THEORIE

und erzwingt entweder $\psi(0) = 0 \forall \psi \in \mathcal{D}(p)$ oder $\varphi(0) = 0 \forall \varphi \in \mathcal{D}(p^\dagger)$. Der jeweils andere Definitionsbereich enthält Funktionen mit beliebigen Werten bei $x = 0$, so dass immer $\mathcal{D}(p) \neq \mathcal{D}(p^\dagger)$.⁴⁵

Es wären hier folgende Themen geplant: Definition des Spektrums eines Operators, Unterscheidung zwischen diskretem und kontinuierlichem Spektrum. Problem: In den üblichen Formulierungen benötigt man hier Maßtheorie. Man müsste auch ausführlicher die Fortsetzung von Operatoren und die Unterscheidung zwischen schwacher und starker Konvergenz diskutieren. Als Beispiel: Selbstadjungierte Operatoren aus positiven symmetrischen Formen.

§ 26 Sturm-Liouville-Theorie

Sturm-Liouville-Randwertaufgaben sind in der mathematischen Physik allgegenwärtig. Man sollte sie hier als Beispiel der Spektraltheorie ausführlich behandeln. Problem: Vollständigkeit der Eigenfunktionen.

§ 27 Fourier-Transformation

Anknüpfung an vorherigen § über harmonischen Oszillator (Hermite/Laguerre-Polynome), dann Übergang zu partiellen Differentialgleichungen und höher-dimensionalen Randwertaufgaben, Greensche Funktionen. Wohin mit Distributionen?

⁴⁵Dies ist natürlich formal nicht ganz korrekt, denn unsere "Funktionen" nehmen ja an Punkten keine echten Werte an...