



UNIVERSITÄT  
HEIDELBERG  
**SEMPER  
APERTUS**

**SKRIPT ZUR VORLESUNG**  
**HÖHERE MATHEMATIK III**  
**FÜR DAS**  
**STUDIUM DER PHYSIK**

**gehalten an der**  
Universität Heidelberg  
**im**  
Wintersemester 2020/21  
**von**  
Johannes Walcher



# INHALTSVERZEICHNIS

<b>6</b>	<b>FUNKTIONENTHEORIE</b>	<b>123</b>
§ 15	Komplex differenzierbare Funktionen . . . . .	124
§ 16	Integration in $\mathbb{C}$ . . . . .	129
§ 17	Cauchy-Formeln . . . . .	134
§ 18	Residuensatz . . . . .	140
<b>7</b>	<b>SYMMETRIEN</b>	<b>149</b>
§ 19	Multi-lineare Algebra . . . . .	149
§ 20	Gruppen und ihre Darstellungen . . . . .	161
§ 21	Euklidische und unitäre Räume . . . . .	171
<b>8</b>	<b>INTEGRATIONSTHEORIE</b>	<b>181</b>
§ 22	Integration im $\mathbb{R}^n$ . . . . .	181
§ 23	Integralsätze . . . . .	205
§ 24	Vervollständigung . . . . .	229
§ 25	Eigenschaften des Lebesgue-Integrals . . . . .	237
<b>9</b>	<b>FUNKTIONALE</b>	<b>247</b>
§ 26	Hilbert-Räume . . . . .	247
§ 27	Lineare Operatoren . . . . .	258
§ 28	Sturm-Liouville-Theorie . . . . .	269
§ 29	Fourier-Transformation . . . . .	269

## INHALTSVERZEICHNIS

# KAPITEL 6

## FUNKTIONENTHEORIE

Diese Vorlesung ist die Fortsetzung der Höheren Mathematik II 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. Hierzu wiederholen wir in Kapitel 7 zunächst einige Ergebnisse der Linearen Algebra, auf die wir uns bereits im letzten Semester regelmässig, und manchmal etwas unreflektiert, berufen haben, und bauen sie an denjenigen Stellen aus, die in der Quantenmechanik eine ausgezeichnete Rolle spielen. Anschliessend und darauf aufbauend kehren wir in Kapitel 8 zur Analysis zurück, entwickeln die Integrationstheorie im  $\mathbb{R}^n$  und besprechen Differentialformen und Integralsätze als erster Schritt zur Umkehrung der mehr-dimensionalen Differentialrechnung aus §10. Im abschliessenden Kapitel 9 beginnen wir noch die Untersuchung von linearen Operatoren in Hilberträumen, und geben (hoffentlich) einen Einblick in die Fourier-Transformation sowie die Theorie partieller Differentialgleichungen. Dabei behalten die grundsätzlichen Überlegungen aus dem Sommer zum Verhältnis zu den theoretischen Physik-Vorlesungen natürlich weiter ihre Gültigkeit.

Wir beginnen das Wintersemester jedoch mit einem Kapitel zur Funktionentheorie, die wir im Sommer aufgrund der durch die Umstellung des Lehrformats bedingten zeitlichen Zwänge überspringen mussten. Hier geht es vordergründig einfach um den Begriff der Ableitung komplexwertiger Funktionen einer komplexen Veränderlichen und deren Umkehrung, in Analogie zu §9 und §14. Die Antwort ist jedoch höchst überraschend, und ohne echtes Analogon in der rein reellen Analysis. Es stellt sich nämlich heraus, dass für  $U \subset \mathbb{C}$  offen eine Funktion  $f : U \rightarrow \mathbb{C}$  lokal genau dann eine Stammfunktion besitzt (d.h., komplex integrierbar ist), wenn sie holomorph (nämlich, komplex differenzierbar) ist. Stetigkeit reicht nicht mehr aus! Eine globale Konsequenz aus dieser “Vereinheitlichung” von Differentiation und Integration ist, dass sich bestimmte komplexe Integrale in viel systematischerer Weise als etwa die Beispiele (14.29), (14.38) durch Studium der Ableitungen des Integranden auswerten lassen. In diesem Kapitel steht damit unser bisheriges Leitziel, dass wir nach Lösungen von physikalischen Problemen durch Approximation suchen, etwas zurück. Vielmehr bietet die Funktionentheorie (in eigentlich unerwarteter Weise) ein Hilfsmittel für *exakte* Aussagen über Strukturen in physikalischen Theorien. Ich hoffe, Sie werden davon in absehbarer Zeit profitieren!

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 3 (Stetigkeit), 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/wise2021/hoemaphys/](http://www.mathi.uni-heidelberg.de/~walcher/teaching/wise2021/hoemaphys/)

## § 15 Komplex differenzierbare Funktionen

Wir diskutieren in diesem § den Begriff der Differenzierbarkeit für komplexwertige Funktionen  $f : U \rightarrow \mathbb{C}$  einer komplexen Veränderlichen  $z \in U \subset \mathbb{C}$ . Da per Definition  $\mathbb{C} \cong \mathbb{R}^2$  als zwei-dimensionaler reeller Vektorraum (sogar mit ausgezeichnete Basis, und die metrische Struktur/Norm stimmt auch), mag es scheinen, dass es sich hier allenfalls um einen Spezialfall der mehr-dimensionalen Differentiation handeln kann, wenn auch vielleicht einer mit etwas physikalischer Relevanz, da  $2 < 3$  vom Ende des §10. Es könnte kaum ein Schein mehr trügen.

Benutzt man nämlich zur Differentiation von Funktionen einer komplexen Veränderlichen nicht nur die reell-lineare Struktur des  $\mathbb{R}^2$ , sondern ausserdem die Körperstruktur von  $\mathbb{C}$ , so erhält man erstaunlicher Weise einen Ableitungsbegriff, der sich einerseits praktisch aller Eigenschaften der reellen Ableitung erfreut (natürlich nicht derjenigen, die von der Anordnung abhängen), andererseits aber die Funktionen derartig einschränkt, dass viele der technischen Komplikationen (wie Stetigkeit der Ableitung etc.) gleichsam von selbst verschwinden.

Spätestens in der Quantenmechanik lernt man, dass komplexe Funktionswerte in der Physik eine enorm wichtige Rolle spielen. Komplexe Definitionsbereiche treten hingegen nach wie vor nicht direkt in den Bildmengen physikalischer Theorien auf, sie bieten aber ein unverzichtbares praktisches Hilfsmittel. Wir werden Beispiele solcher “Lösung reeller Probleme über den Umweg ins Komplexe” im §18 vorstellen.

**Definition 15.1.** Sei  $U \subset \mathbb{C}$  eine offene Menge. Eine Funktion  $f : U \rightarrow \mathbb{C}$  heisst komplex differenzierbar in  $z_0 \in U$ , falls der Differenzenquotient

$$\frac{f(z) - f(z_0)}{z - z_0} \quad (15.1)$$

sich stetig in  $z_0$  fortsetzen lässt. In diesem Fall heisst der Wert der Fortsetzung in  $z_0$  die Ableitung von  $f$  in  $z_0$ . Schreibweise:  $f'(z_0) \in \mathbb{C}$ .

- Eine Funktion heisst holomorph in  $z_0$ , wenn sie in einer offenen Umgebung von  $z_0$  komplex differenzierbar ist, und holomorph auf  $U$ , wenn sie dort überall komplex differenzierbar ist.

*Bemerkungen.* · Wie bereits angedeutet ist diese Definition formal identisch zu 9.1, und erlaubt eine offensichtliche Verallgemeinerung zu normierten *komplexen* Vektorräumen als Wertebereich.

- Es gelten die gewohnten Produkt- und Kettenregeln. Polynome sind überall komplex differenzierbar, und rationale Funktionen ausserhalb der Nullstellen des Nenners. Exponential- und trigonometrische Funktionen behalten ebenfalls ihre gewohnten Ableitungen.

- Insbesondere ist eine komplex differenzierbare Funktion stetig. Es stellt sich heraus, dass die Ableitung einer holomorphen Funktion automatisch stetig (und sogar

## § 15. KOMPLEX DIFFERENZIERBARE FUNKTIONEN

komplex differenzierbar) ist, weshalb der Begriff der “stetigen komplexen Differenzierbarkeit” nicht separat eingeführt wird.

· Als ein Beispiel für die zusätzlichen Einschränkungen in der komplexen Ebene betrachten wir die naive Fortsetzung<sup>1</sup> der differenzierbaren, aber nicht stetig differenzierbaren Funktion aus Beispiel (9.5) auf komplexe Argumente, d.h.

$$\mathbb{C} \setminus \{0\} \ni z \mapsto \tilde{f}(z) = z^2 \sin \frac{1}{z} \quad (15.2)$$

Solange  $z \neq 0$  folgt aus den obigen Regeln, dass die Ableitung von  $\tilde{f}$  nach  $z$  formal genauso aussieht wie die ihrer Einschränkung auf reelle Argumente, also  $\tilde{f}'(z) = 2z \sin \frac{1}{z} - \cos \frac{1}{z}$ . (Man schreibt dann sinnvollerweise  $f(z)$  auch für komplexe Argumente.) Da aber für  $z = iy$  rein imaginär  $\sin \frac{1}{iy} = i \sinh \frac{1}{y} \rightarrow \pm i\infty$  für  $y \rightarrow 0_{\pm}$  hat  $f$  noch nicht mal einen Grenzwert bei  $z_0 = 0$ , kann also dort auch nicht komplex differenzierbar sein. (Dieses Argument zeigt noch nicht die (wahre) Aussage, dass sich  $f$  überhaupt nicht als in 0 differenzierbare Funktion von der reellen Achse weg fortsetzen lässt.)

· Es gibt auch ganz vertraute komplexe Funktionen, die aus viel einfacheren Gründen nirgends komplex differenzierbar sind. Wichtigstes Beispiel hierfür ist die komplexe Konjugation,  $z \mapsto \bar{z}$ . Mit  $z - z_0 = r e^{i\varphi}$  für  $\varphi \in [0, 2\pi)$  hat der Differenzenquotient

$$\frac{\bar{z} - \bar{z}_0}{z - z_0} = e^{-2i\varphi} \quad (15.3)$$

offensichtlich keinen Grenzwert für  $z \rightarrow z_0$ . (Im Unterschied zum vorherigen Beispiel hat die Einschränkung von  $z \mapsto \bar{z}$  auf die reelle Achse, nämlich die Identität  $x \mapsto x$  eine andere, komplex differenzierbare Fortsetzung ins Komplexe, nämlich  $z \mapsto z$ .)

· Durch Umschreiben von (15.1) als

$$\lim_{z \rightarrow z_0} \frac{|f(z) - f(z_0) - f'(z_0)(z - z_0)|}{|z - z_0|} = 0 \quad (15.4)$$

und Vergleich mit 10.7 sieht man aber auch, dass man die komplexe Differenzierbarkeit als Spezialfall des zwei-dimensionalen reellen Begriffs auffassen kann: Multiplikation mit  $f'(z_0)$  ist eine lineare Abbildung  $\mathbb{R}^2 \rightarrow \mathbb{R}^2$ , und der Grenzwert  $z \rightarrow z_0$  in  $\mathbb{C}$  ist der gleiche wie  $(x, y) \rightarrow (x_0, y_0)$  in  $\mathbb{R}^2$  (wobei natürlich  $z = x + iy$ ,  $z_0 = x_0 + iy_0$ ).

· Zur Untersuchung des genauen Zusammenhangs zwischen den beiden Begriffen schreibt man  $f(z) = u(x, y) + iv(x, y)$  mit reellwertigen Funktionen  $u, v$  von  $(x, y) \in U \subset \mathbb{R}^2$ , und nutzt aus, dass falls der Grenzwert von (15.1) existiert, er (per Definition) unabhängig davon ist, aus welcher (“reellen”) Richtung man sich  $z_0$  nähert. Entlang der reellen Achse, d.h.  $z = z_0 + t$  für  $t \in \mathbb{R}$  findet man

$$f'(z_0) = \lim_{t \rightarrow 0} \frac{f(x_0 + t, y_0) - f(x_0, y_0)}{t} = \frac{\partial u}{\partial x}(x_0, y_0) + i \frac{\partial v}{\partial x}(x_0, y_0) \quad (15.5)$$

<sup>1</sup>Zur Erinnerung an S. 59: Eine Fortsetzung einer Abbildung  $F : X \rightarrow Y$  auf eine Obermenge  $\tilde{X} \supset X$  ist eine Abbildung  $\tilde{F} : \tilde{X} \rightarrow Y$  mit  $\tilde{F}|_X = F$

Entlang der imaginären Achse ist  $z = z_0 + it$  für  $t \in \mathbb{R}$  und

$$f'(z_0) = \lim_{t \rightarrow 0} \frac{f(x_0, y_0 + t) - f(x_0, y_0)}{it} = -i \frac{\partial u}{\partial y}(x_0, y_0) + \frac{\partial v}{\partial x}(x_0, y_0) \quad (15.6)$$

Durch Koeffizientenvergleich ergibt sich:

**Lemma 15.2.** *Eine Funktion  $f : U \rightarrow \mathbb{C}$  ist genau dann in  $z_0 \in U$  komplex differenzierbar, wenn die zugehörige  $\mathbb{R}^2$ -wertige Funktion in  $(x_0, y_0)$  (total) differenzierbar ist und die partiellen Ableitungen die Cauchy-Riemann-Gleichungen erfüllen:*

$$\frac{\partial u}{\partial x}(x_0, y_0) = \frac{\partial v}{\partial y}(x_0, y_0) \quad \text{und} \quad \frac{\partial u}{\partial y}(x_0, y_0) = -\frac{\partial v}{\partial x}(x_0, y_0) \quad (15.7)$$

In diesem Fall ist die komplexe Ableitung gleichwertig durch (15.5) oder (15.6) gegeben.

*Beweis.* Es bliebe allenfalls noch nachzureichen, dass eine lineare Abbildung mit einer Matrix-Darstellung der Form

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix} \in \text{Mat}_{2 \times 2}(\mathbb{R}) \quad (15.8)$$

sich als Multiplikation von  $\mathbb{C}$  mit  $a + ib$  identifizieren lässt. □

### Geometrische Interpretation

Wir nehmen (15.7) als Gelegenheit zur Schärfung unserer vektoranalytischen Anschauung. Dazu betrachten wir in der  $(u, v)$ -Ebene die Bilder der  $x$ - und  $y$ -Koordinatenrichtungen. Es ist

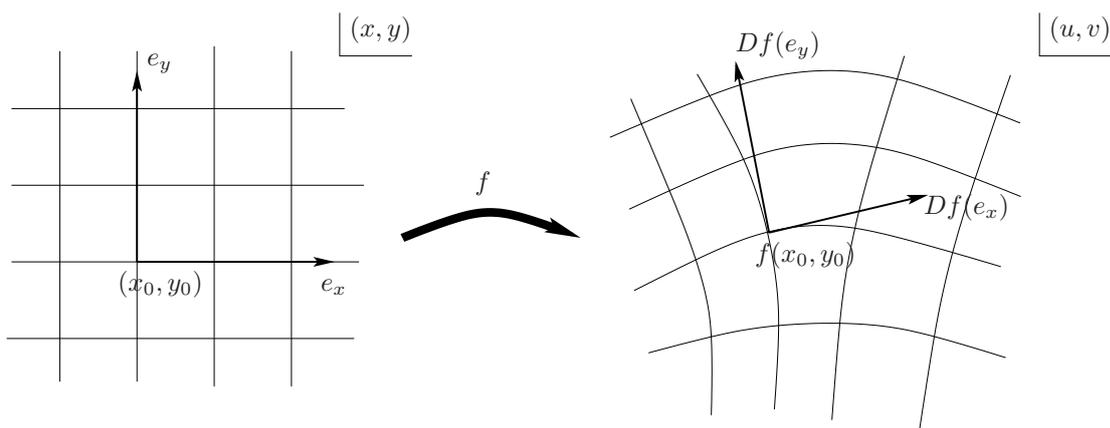
$$Df(e_x) = (\partial_x u, \partial_x v)^T, \quad Df(e_y) = (\partial_y u, \partial_y v)^T \quad (15.9)$$

Als Konsequenz der Cauchy-Riemann-Gleichungen gilt bezüglich dem euklidischen inneren Produkt in Definitions- und Wertebereich

$$\langle Df(e_x), Df(e_y) \rangle = \langle e_x, e_y \rangle = 0, \quad \|Df(e_x)\| = \|Df(e_y)\| \quad (15.10)$$

und daraus folgt (im Fall  $f'(z_0) \neq 0$ ), dass das Differential den Winkel zwischen je zwei beliebigen Richtungen im Tangentialraum  $T_{(x_0, y_0)}U \cong \mathbb{R}^2$  erhält, Längen aber um einen gemeinsamen Faktor vergrößert oder verkleinert. Solche Abbildungen heißen “konform”. Vom komplexen Standpunkt ist dies aber nichts anderes als die in §2 gegebene Interpretation der Multiplikation mit einer komplexen Zahl (hier:  $f'(z_0)$ ) als “Drehstreckung”. (Die Umkehrung konform  $\Rightarrow$  komplex differenzierbar gilt bis auf Orientierungswechsel/komplexe Konjugation.)

## § 15. KOMPLEX DIFFERENZIERBARE FUNKTIONEN



· Eine weitere von Ingenieuren und Physikern häufig benutzte Konsequenz aus den Cauchy-Riemann-Gleichungen ergibt sich unter der (nachträglich redundanten) Annahme, dass  $u$  und  $v$  mindestens zweimal stetig differenzierbar sind:

$$\begin{aligned} \partial_x^2 u &\stackrel{(15.7)}{=} \partial_x \partial_y v \stackrel{10.13}{=} \partial_y \partial_x v \stackrel{(15.7)}{=} -\partial_y^2 u \\ &\Rightarrow \partial_x^2 u + \partial_y^2 u = 0 \end{aligned} \quad (15.11)$$

Der Real- (und auch Imaginär-)Teil einer holomorphen Funktion löst also die zweidimensionale Laplace-Gleichung.

### Analytische Funktionen

Die aber vielleicht wichtigste Quelle holomorpher Funktionen sind die im §5 eingeführten Potenzreihen mit komplexen Koeffizienten. Es gilt nämlich

**Proposition 15.3.** *Eine Potenzreihe  $P(z) = \sum_{n=0}^{\infty} a_n z^n$  ist an jeder Stelle  $z_0 \in B_R(0)$  innerhalb ihrer Konvergenzscheibe komplex differenzierbar mit der (gliedweisen) Ableitung  $P'(z) = \sum_{n=0}^{\infty} a_n n z^{n-1}$*

*Beweis.* Wir hatten bereits im §8 gesehen, dass  $P(z)$  auf  $B_R(0)$  eine stetige Funktion darstellt. Der Nachweis der Differenzierbarkeit geht ganz ähnlich.

· Zunächst halten wir fest, dass die Reihe der gliedweisen Ableitungen  $\sum a_n n z^{n-1}$  wegen  $\lim_{n \rightarrow \infty} n^{1/n} = 1$  den gleichen Konvergenzradius hat wie  $P$  (siehe (5.45)) und daher ebenso wie  $P$  für jedes  $|z| < R$  absolut konvergiert. Insbesondere ist Umordnen erlaubt.

· Es sei  $r$  so, dass  $|z_0| < r < R$ . Für gegebenes  $\epsilon > 0$  sei dann  $N$  so gross, dass

$$\sum_{n=N+1}^{\infty} |a_n| n r^{n-1} < \frac{\epsilon}{3} \quad (15.12)$$

Dann ist für alle  $|z| < r$

$$\begin{aligned} & \left| \sum_{n=0}^{\infty} a_n z^n - \sum_{n=0}^{\infty} a_n z_0^n - \sum_{n=0}^{\infty} a_n n z_0^{n-1} (z - z_0) \right| \\ & \leq \left| \sum_{n=0}^N (a_n z^n - a_n z_0^n - a_n n z_0^{n-1} (z - z_0)) \right| \\ & \quad + \underbrace{\sum_{n=N+1}^{\infty} |a_n| \cdot \underbrace{|z^n - z_0^n|}_{\leq nr^{n-1}|z-z_0|} + \sum_{n=N+1}^{\infty} |a_n| nr^{n-1} |z - z_0|}_{< \frac{2}{3} \epsilon \cdot |z - z_0|} \end{aligned} \tag{15.13}$$

Da Polynome komplex differenzierbar sind, existiert ein  $\delta > 0$  so, dass die endliche Summe für  $|z - z_0| < \delta$  kleiner als  $\frac{\epsilon}{3} \cdot |z - z_0|$  ist. Zusammen ist alles kleiner als  $\epsilon \cdot |z - z_0|$ , falls  $z \in B_\delta(z_0) \cap B_r(0)$  d.h. es folgt die Behauptung.  $\square$

**Definition 15.4.** Eine Funktion  $f : U \rightarrow \mathbb{C}$  heisst (komplex) analytisch, falls es für jeden Punkt  $z_0 \in U$  eine Folge  $(a_n^{(z_0)})$  komplexer Zahlen und ein  $\delta > 0$  gibt so, dass  $B_\delta(z_0) \subset U$  und für alle  $z \in B_\delta(z_0)$

$$f(z) = \sum_{k=0}^{\infty} a_k^{(z_0)} (z - z_0)^k \tag{15.14}$$

- Insbesondere hat die Potenzreihe auf der rechten Seite positiven Konvergenzradius  $R^{(z_0)}$ . Es muss aber nicht  $B_{R^{(z_0)}}(z_0) \subset U$  gelten.
- Durch wiederholte Anwendung von 15.3 folgt:

**Korollar 15.5.** *Eine analytische Funktion ist in ihrem Definitionsbereich unendlich oft komplex differenzierbar, und es gilt die “Taylor-Formel”*

$$f^{(k)}(z_0) = k! a_k^{(z_0)} \tag{15.15}$$

**Bemerkungen/Beispiele 15.6.** · Umgekehrt ist eine durch eine Potenzreihe  $P(z) = \sum a_n z^n$  dargestellte Funktion analytisch: Für  $z_0 \in B_R(0)$  erhält man durch “Umentwickeln”<sup>2</sup> von  $P(z_0 + (z - z_0))$  eine Potenzreihendarstellung der Form (15.14), welche mindestens für  $z \in B_\delta(z_0) \subset B_R(0)$  auch wieder die gegebene Funktion darstellt.

- Beispiel: Geometrische Reihe zur Illustration der Idee der analytischen Fortsetzung.
- Beispiel:  $Q(z) = z^2$  als biholomorphe Abbildung  $\{z \mid \operatorname{Re}(z) > 0\} \rightarrow \mathbb{C}_- = \mathbb{C} \setminus \mathbb{R}_{\leq 0}$

<sup>2</sup>Dies ist nicht strenggenommen eine direkte Konsequenz aus dem Umordnungssatz von S. 48. Auf jeden Fall konvergiert aber

$$a_k^{(z_0)} = \frac{1}{k!} P^{(k)}(z_0) = \sum_{n=k}^{\infty} \binom{n}{k} a_n z_0^{n-k} \tag{15.16}$$

Die Aussage über Konvergenz der umentwickelten Reihe lässt sich am einfachsten mit den Methoden aus §17 zeigen.

### Wichtige Ableitungen als Übungsaufgaben

Potenzreihen mit Argumenten aus endlich-dimensionalen Banach-Algebren sind innerhalb ihrer Konvergenzkugel ebenfalls gliedweise ableitbar. Allerdings gilt für die Ableitung im nicht-kommutativen Fall die etwas allgemeinere Formel (vgl. schon (9.3))

$$DP(x_0)(x - x_0) = \sum_{n=1}^{\infty} a_n \sum_{k=0}^{n-1} x_0^k (x - x_0) x_0^{n-k-1} \quad (15.17)$$

(Annahme:  $a_n \in \mathbb{C}$ ) Jedenfalls ist immer

$$D \exp(0) = \text{id}_{\mathcal{A}} \quad (15.18)$$

· Für  $x \in \text{Mat}_{n \times n}(\mathbb{R})$  mit  $\det x \neq 0$  gilt

$$D \det(x)(v) = \det x \operatorname{tr}(x^{-1}v) \quad (15.19)$$

· Quaternionische Ableitung

## § 16 Integration in $\mathbb{C}$

Nach der einigermaßen glatten Verallgemeinerung von §9 auf eine komplexe Veränderliche stellen wir jetzt die Frage nach der Umkehrung der komplexen Ableitung, d.h. der Existenz von Stammfunktionen. Anstatt allerdings wie bisher alles von Anfang an richtig hinzuschreiben, gehen wir in diesem § entdeckerkisch vor und tasten uns schrittweise an die endgültigen Definitionen heran.<sup>3</sup>

Zu Beginn müssen wir uns auf eine Klasse von Teilmengen von  $\mathbb{C}$  als Definitionsbereiche einigen, auf denen wir nach Stammfunktionen suchen sollen. Ein Blick zurück auf 14.3 zeigt, dass die für die Definition des bestimmten Integrals zentrale Eigenschaft von Intervallen ihre Konvexität ist: Für je zwei Punkte  $a, b \in I$  liegen auch alle Zwischenstellen  $(1-t)a + tb$ ,  $t \in [0, 1]$  in  $I$ . Eine sinnvolle komplexe Verallgemeinerung sind also offene und konvexe<sup>4</sup> Teilmengen  $U \subset \mathbb{C}$ , und mit dieser Erkenntnis starten wir sofort unseren ersten Versuch, ein "Integral von  $a$  nach  $b$  in  $U$ " zu definieren. Der Einfachheit halber beschränken wir uns auf einen (komplex) ein-dimensionalen Wertebereich. Wir nehmen also an,  $f : U \rightarrow \mathbb{C}$  sei stetig und setzen für  $n \in \mathbb{N}$ :

$$S_n(f; a, b) := \sum_{k=0}^{n-1} \left( \frac{b-a}{n} \right) \cdot f \left( a + \frac{k}{n}(b-a) \right) \quad (16.1)$$

So wie sich 15.1 von 9.1 nur im Namen der Variablen unterscheidet, ist diese Definition formal identisch zu (14.2). What could possibly go wrong?

Beh.: Es existiert  $\lim_{n \rightarrow \infty} S_n(f; a, b)$

<sup>3</sup>Wir sagen natürlich auch nicht absichtlich etwas Falsches!

<sup>4</sup>Zur Definition der Konvexität wird  $\mathbb{C}$  mit  $\mathbb{R}^2$  identifiziert:  $U \subset \mathbb{C}$  heisst konvex, falls  $\forall a, b \in U$ ,  $\{a + t(b-a) \mid t \in [0, 1]\} \subset U$ , s. S. 91.

Bew.: Wir betrachten auf dem Intervall  $[0, 1] \subset \mathbb{R}$  die stetige, komplexwertige Funktion

$$[0, 1] \ni t \mapsto \tilde{f}(t) := f(a + t(b - a)) = f((1 - t)a + tb) \in \mathbb{C} \quad (16.2)$$

Dann ist

$$S_n(f; a, b) = (b - a) \cdot S_n(\tilde{f}; 0, 1) \quad (16.3)$$

wo die rechte Seite nun echt mit (14.2) zusammentrifft, da  $0, 1 \in \mathbb{R}$ . Die Konvergenz folgt also aus der Stetigkeit von  $\tilde{f}$  und es gilt

$$\int_a^b f(w)dw := \lim_{n \rightarrow \infty} S_n(f; a, b) = (b - a) \int_0^1 \tilde{f}(t)dt \quad (16.4)$$

□

Beachte: Die Notation auf der linken Seite ist im Komplexen obsolet. In Anbetracht der endgültigen Definition 16.2 sollten wir schreiben

$$\int_{l(a,b)} f(w)dw \quad (16.5)$$

wobei  $l(a, b) = (t \mapsto a + t(b - a))$  für die “gerichtete Strecke von  $a$  nach  $b$ ” steht.

Beispiel: Angenommen,  $F : U \rightarrow \mathbb{C}$  ist eine holomorphe Funktion mit  $F'(z) = f(z) \forall z \in U$ . Dann folgt aus der Kettenregel, dass

$$[0, 1] \ni t \mapsto \tilde{F}(t) := F(a + t(b - a)) \in \mathbb{C} \quad (16.6)$$

(reell) differenzierbar ist mit der Ableitung

$$\dot{\tilde{F}}(t) = \frac{d}{dt} F(a + t(b - a)) = (b - a) \cdot F'(a + t(b - a)) = (b - a) \cdot \tilde{f}(t) \in \mathbb{C} \quad (16.7)$$

$\tilde{F}$  ist also auf  $[0, 1]$  eine Stammfunktion von  $(b - a)\tilde{f}$  (oder, wenn gewollt, fortgesetzt auf ein offenes Intervall  $(-\epsilon, 1 + \epsilon)$  für  $\epsilon > 0$ ). Nach dem Hauptsatz der (reellen) Differential- und Integralrechnung folgt daher

$$\int_a^b f(w)dw = \tilde{F}(1) - \tilde{F}(0) = F(b) - F(a) \quad (16.8)$$

Explizitere Beispiele sind Konstanten, Polynome, Potenzreihen, etc.

Können wir nun umgekehrt (16.4) benutzen, um für jede solche stetige Funktion  $f : U \rightarrow \mathbb{C}$  eine komplexe Stammfunktion zu konstruieren? NEIN! — In unserem Beweis des Hauptsatzes (14.16) haben wir nämlich *alle* Eigenschaften in 14.6 benutzt und man sieht:

- (i) (Orientierungswechsel) gilt auch für (16.4) ✓
- (ii) (Linearität und Integration von Konstanten) ✓
- (iii) (Standardabschätzung) (mit Norm = Absolutbetrag) ✓

ABER (iv) (Additivität im Integrationsbereich) gilt im Allgemeinen nicht! Der

§ 16. INTEGRATION IN  $\mathbb{C}$

Vorfaktor  $(b - a)$  in (16.4) verhindert, dass der Beweis von Seite 111 ohne weiteres durchgeht, ein einfaches Gegenbeispiel ist:  $f(z) = \bar{z}$ ,  $a = 0$ ,  $b = 1$ ,  $c = 1 + i$ .

$$\begin{aligned} \int_0^1 \bar{w}dw &= \int_0^1 tdt = \frac{1}{2} \\ \int_1^{1+i} \bar{w}dw &= i \int_0^1 \overline{(1+it)}dt = i \int_0^1 (1-it)dt = i + \frac{1}{2} \\ \int_0^{1+i} \bar{w}dw &= (1+i) \int_0^1 (1-i)tdt = 2 \cdot \int_0^1 tdt = 1 \\ &\Rightarrow \int_0^1 + \int_1^{1+i} \neq \int_0^{1+i} \text{!!!!} \end{aligned} \tag{16.9}$$

Da aber sonst alle Eigenschaften gelten, und da aus (16.8) umgekehrt folgt, dass die Additivität notwendig ist für die Existenz einer Stammfunktion, können wir folgern: Fazit: Eine stetige Funktion  $f : U \rightarrow \mathbb{C}$  auf einer offenen und konvexen Menge  $U \subset \mathbb{C}$  besitzt dann und nur dann eine komplexe Stammfunktion, d.h. eine holomorphe Funktion  $F : U \rightarrow \mathbb{C}$  mit  $F'(z) = f(z) \forall z \in U$  falls für alle  $a, b, c \in U$  gilt:

$$\int_a^b f(w)dw + \int_b^c f(w)dw + \int_c^a f(w)dw = 0 \tag{16.10}$$

Wie in  $\mathbb{R}$  ist eine solche Stammfunktion durch das unbestimmte Integral gegeben und bis auf eine additive Konstante eindeutig.

**Lemma 16.1.** *Sei  $U$  wie eben offen und konvex. Dann gilt die Aussage (16.10), falls  $f$  auf  $U$  holomorph ist!*

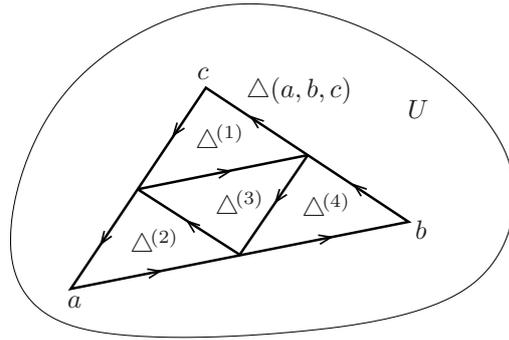
*Beweis.* Auf der Suche nach einem Ersatz für die gemeinsame Verfeinerung der Intervall-Zerlegungen aus 14.5 benutzen wir die Visualisierung von  $\mathbb{C}$  als zweidimensionale euklidische Ebene und daraus abgeleitete geometrische Sprechweisen. Für drei Punkte  $o, p, q \in \mathbb{C}$  bezeichne  $\blacktriangle(o, p, q)$  das (allenfalls entartete<sup>5</sup>) Dreieck mit diesen Punkten als Ecken und  $\triangle(o, p, q)$  seinen Rand. Beides sind abgeschlossene Teilmengen von  $\mathbb{C}$  im Sinne von 4.18. Der Durchmesser ist  $\text{diam } \blacktriangle(o, p, q) = \max\{|p - o|, |q - p|, |o - q|\}$  und der Umfang  $|\triangle(o, p, q)| = |p - o| + |q - p| + |o - q|$ . Für  $o, p, q \in U$  ist wegen der Konvexität von  $U$   $\blacktriangle(o, p, q) \subset U$ . Wir schreiben

$$\int_{\triangle(o,p,q)} f(w)dw = \int_o^p f(w)dw + \int_p^q f(w)dw + \int_q^o f(w)dw \tag{16.11}$$

für das "Integral über der Rand" in der angegebenen Orientierung. (Es gilt  $\int_{\triangle(o,p,q)} f(w)dw = \int_{\triangle(p,q,o)} f(w)dw = -\int_{\triangle(o,q,p)} f(w)dw$  gemäss Eigenschaft (i).)

<sup>5</sup>Sind  $a, b, c$  kollinear oder fallen zusammen, so gilt (16.10) auf jeden Fall.

Die Behauptung ist also, dass für je drei Punkte  $a, b, c \in U$ ,  $\int_{\Delta(a,b,c)} f(w)dw = 0$ , falls  $f$  holomorph ist. Wir zeigen dafür, dass für jedes  $\epsilon > 0$   $|\int_{\Delta(a,b,c)} f(w)dw| < \epsilon$ . Dazu zerteilen wir  $\blacktriangle(a, b, c)$  durch Halbieren der Seiten in vier kongruente Dreiecke, deren Ränder  $\Delta^{(1)}$ ,  $\Delta^{(2)}$ ,  $\Delta^{(3)}$  und  $\Delta^{(4)}$ , wir in einer Weise orientieren, dass



$$\int_{\Delta(a,b,c)} f(w)dw = \sum_{i=1}^4 \int_{\Delta^{(i)}} f(w)dw \quad (16.12)$$

Gemäss Behauptung ist jeder der vier Terme auf der rechten Seite ebenfalls null. Ohne dies zu wissen, suchen wir ein jeniges Dreieck aus, für das der Wert des Integrals (im Absolutbetrag) das Maximum von allen vieren ist, und bezeichnen es mit  $\Delta_1$ . Mit  $\Delta_0 := \Delta(a, b, c)$  gilt dann

$$\left| \int_{\Delta(a,b,c)} f(w)dw \right| = \left| \int_{\Delta_0} f(w)dw \right| \leq 4 \cdot \left| \int_{\Delta_1} f(w)dw \right| \quad (16.13)$$

Nun zerteilen wir analog  $\blacktriangle_1$ , suchen das Maximum der vier Integrale und bezeichnen das zugehörige Dreieck mit  $\blacktriangle_2$ . Durch Wiederholen des Verfahrens erhalten wir eine Folge  $(\blacktriangle_n)$  von Dreiecken mit Rändern  $(\Delta_n)$  über die

$$\left| \int_{\Delta(a,b,c)} f(w)dw \right| \leq 4^n \cdot \left| \int_{\Delta_n} f(w)dw \right| \quad (16.14)$$

Wegen der Kongruenz gilt  $\text{diam } \blacktriangle_n = 2^{-n} \text{diam } \blacktriangle_0$  und  $|\Delta_n| = 2^{-n} |\Delta_0|$ , bilden also Nullfolgen. Ausserdem ist  $\blacktriangle_{n+1} \subset \blacktriangle_n$  und nach 4.24 existiert daher genau ein Punkt  $z_0 \in \bigcap_{n=0}^{\infty} \blacktriangle_n$ .

· Da  $f$  nach Annahme in  $z_0$  komplex differenzierbar ist, existiert ein  $\delta > 0$  so, dass

$$\left| f(z) - f(z_0) - f'(z_0)(z - z_0) \right| \leq \epsilon \cdot |z - z_0| \quad (16.15)$$

falls  $|z - z_0| < \delta$ . Sei nun  $n$  so gross, dass  $\text{diam } \blacktriangle_n < \delta$ . Dann ist für  $w \in \Delta_n$   $|w - z_0| \leq \text{diam } \blacktriangle_n < \delta$ . Wegen der Eigenschaft (ii) des Integrals (16.4) gilt

$$\int_{\Delta_n} f(w)dw = \int_{\Delta_n} (f(w) - f(z_0) - f'(z_0)(w - z_0))dw \quad (16.16)$$

und mit Hilfe von Eigenschaft (iii) und (16.15) folgt

$$\begin{aligned} \left| \int_{\Delta_n} f(w)dw \right| &\leq \sup\{|f(w) - f(z_0) - f'(z_0)(w - z_0)| \mid w \in \Delta_n\} \cdot |\Delta_n| \\ &\leq \epsilon \cdot \text{diam } \blacktriangle_n \cdot |\Delta_n| \\ &\leq \epsilon \cdot \frac{\text{diam } \blacktriangle_0}{2^n} \cdot \frac{|\Delta_0|}{2^n} \end{aligned} \quad (16.17)$$

§ 16. INTEGRATION IN  $\mathbb{C}$

Mit (16.14) folgt

$$\left| \int_{\Delta(a,b,c)} f(w)dw \right| \leq \epsilon \cdot \text{diam } \blacktriangle(a,b,c) \cdot |\Delta(a,b,c)| \quad (16.18)$$

und daraus die Behauptung. □

Wir zeigen im nächsten §, dass falls  $f$  umgekehrt eine Stammfunktion besitzt, d.h. (16.10) gilt, dann  $f$  bereits holomorph ist. Zunächst aber flexibilisieren wir etwas Integral und Definitionsbereiche.

Die Motivation für die folgende Definition ist die bereits gemachte Beobachtung, dass falls eine Stammfunktion existiert, es zu ihrer Rekonstruktion nicht darauf ankommt, "wie genau wir von  $a$  nach  $b$  kommen", Hauptsache, wir kommen an (vgl. dazu die Warnungen auf Seite 113).

**Definition 16.2.** Sei  $U \subset \mathbb{C}$  offen. Ein stetig differenzierbarer Weg in  $U$  von  $a$  nach  $b$ , beides Punkte in  $U$ , ist eine stetig differenzierbare Abbildung  $\gamma : [\alpha, \beta] \rightarrow U$  mit  $\gamma(\alpha) = a$ ,  $\gamma(\beta) = b$ . Für eine stetige Funktion  $f : U \rightarrow \mathbb{C}$  definieren wir dann das Integral von  $f$  über  $\gamma$  mit 14.3 als

$$\int_{\gamma} f(w)dw := \int_{\alpha}^{\beta} f(\gamma(t)) \cdot \dot{\gamma}(t) dt \quad (16.19)$$

$\uparrow$   
 hier ist die Multiplikation  
 in  $\mathbb{C}$  versteckt!

· Die Länge eines Weges ist das reellwertige Integral

$$l(\gamma) := \int_{\alpha}^{\beta} |\dot{\gamma}(t)| dt \quad (16.20)$$

· Ist  $\gamma_1 : [\alpha, \beta] \rightarrow U$  ein Weg von  $a$  nach  $b$  und  $\gamma_2 : [\beta, \gamma] \rightarrow U$  ein Weg von  $b$  nach  $c$  so definieren wir die Verkettung (oder Aneinanderreihung) von  $\gamma_1$  und  $\gamma_2$  durch

$$[\alpha, \gamma] \ni t \mapsto \begin{cases} \gamma_1(t) & t \in [\alpha, \beta] \\ \gamma_2(t) & t \in (\beta, \gamma] \end{cases} \quad (16.21)$$

Dies ist wegen  $\gamma_1(\beta) = b = \gamma_2(\beta)$  ein stetiger Weg von  $a$  nach  $c$ , aber in  $\beta$  nicht unbedingt differenzierbar, was aber zur Not auch repariert werden kann. Eine Verkettung von endlich vielen stetig differenzierbaren Wegen nennt man auch stückweise stetig differenzierbar. Integral und Länge sind als Summen über die Stücke definiert.

· Der *Orientierungswechsel* eines Wegs  $\gamma$  ist  $[\alpha, \beta] \ni t \mapsto \gamma(\beta + \alpha - t)$ . Unter Orientierungswechsel vertauschen Anfangs- und Endpunkt ihre Rolle und das Integral schnappt sich ein Vorzeichen. Die Länge bleibt sich gleich.

· Eine *Reparametrisierung* eines Wegs  $\gamma : [\alpha, \beta] \rightarrow U$  ist der Weg  $\tilde{\gamma} = \gamma \circ \tau$  für eine stetig differenzierbare Abbildung  $\tau : [\tilde{\alpha}, \tilde{\beta}] \rightarrow [\alpha, \beta]$  mit  $\tau(\tilde{\alpha}) = \alpha$ ,  $\tau(\tilde{\beta}) =$

---

<sup>6</sup>Das Alphabet hat hier kapitulieren müssen, die Bedeutung von  $\gamma$  sollte aber jeweils klar sein. Noch etwas allgemeiner muss der Definitionsbereich von  $\gamma_2$  nicht direkt an den von  $\gamma_1$  anschliessen, formal schreibe man etwa  $\gamma_2 : [\beta + \delta, \gamma + \delta] \rightarrow U$ , und ersetze  $\gamma_2(t)$  durch  $\gamma_2(t + \delta)$  in (16.21).

$\beta$ . Das Integral ändert sich nicht, die Länge bleibt jedoch nur unter *monotonen* Reparametrisierungen erhalten.

· Statt Wege sagt man auch Kurven oder Konturen, insbesondere dann, wenn man betonen will, dass Integrale nur von dem Bild  $\gamma([\alpha, \beta])$  (welches auch “Spur des Wegs” genannt wird) und der Orientierung, und nicht einer expliziten Parametrisierung abhängen. Ein Weg heisst geschlossen, wenn  $a = b$ .

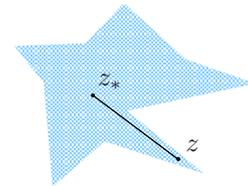
*Bemerkungen.* · Man prüfe nach, dass das Integral (16.19) die Regeln (i) und (ii) erfüllt und (iii) in der Form

$$\left| \int_{\gamma} f(w)dw \right| \leq \sup\{|f(w)| \mid w \in \gamma([\alpha, \beta])\} \cdot l(\gamma) \quad (16.22)$$

(Hier geht die Verträglichkeit der Multiplikation mit der euklidischen Norm = Absolutbetrag ein.)

· Bezüglich (iv) folgt wie in (16.8): Besitzt  $f$  auf  $U$  eine komplexe Stammfunktion  $F$ , so ist  $\int_{\gamma} f(w)dw = F(b) - F(a)$ . Dies ist gemäss 16.1 insbesondere dann der Fall, wenn  $f$  holomorph ist und  $U$  konvex. Im Allgemeinen aber *hängt das Integral tatsächlich vom Weg ab*.

· Etwas allgemeiner kann man aus 16.1 folgern: Ist  $U$  ein Sterngebiet, d.h. existiert ein “Sternmittelpunkt”  $z_* \in U$  so, dass  $\forall z \in U$  die Strecke  $l(z_*, z) \subset U$ , dann besitzt jede holomorphe Funktion auf  $U$  eine komplexe Stammfunktion.



· Man überlege sich so ausführlich wie möglich, dass die Definition 16.2 eine zufriedenstellende Verallgemeinerung des reellen Integrals 14.3 darstellt. Insbesondere treffen beide Definitionen zusammen, wenn  $a < b \in \mathbb{R}$  und  $f$  eine holomorphe Fortsetzung von einem offenen Intervall  $I \supset [a, b]$  auf eine offene Umgebung  $U \supset I$  hinaus in die komplexe Ebene besitzt. Man beachte in diesem Zusammenhang auch die formale Ähnlichkeit zwischen (16.19) und der “Substitutionsregel” (14.21).

## § 17 Cauchy-Formeln

**Beispiel 17.1.** In der Praxis werden nebst den geraden Strecken  $t \mapsto a + t(b - a)$  vor allem Kreise und Kreisbögen als Wege benötigt. Bereits in (2.16) hatten wir die offene Scheibe vom Radius  $R$  um  $z_0 \in \mathbb{C}$  definiert als  $B_R(z_0) = \{z \in \mathbb{C} \mid |z - z_0| < R\}$ , die abgeschlossene Scheibe durch  $\overline{B_R(z_0)} = \{z \in \mathbb{C} \mid |z - z_0| \leq R\}$ . Der Kreis

$$C_R(z_0) := \{z \in \mathbb{C} \mid |z - z_0| = R\} = \overline{B_R(z_0)} \setminus B_R(z_0) \quad (17.1)$$

ist der Rand dieser Scheiben. Wir schreiben jetzt für den geschlossenen Weg mit Spur  $C_R(z_0)$  in der positiven Orientierung (d.h. gegen den Uhrzeigersinn)  $C_R^+(z_0)$  und in der negativen Orientierung  $C_R^-(z_0)$ . Gebräuchliche Parametrisierungen sind

$$[0, 2\pi] \ni t \mapsto z_0 + Re^{\pm it} \quad (17.2)$$

und

$$[0, 1] \ni t \mapsto z_0 + Re^{\pm 2\pi it} \quad (17.3)$$

## §17. CAUCHY-FORMELN

Man greift auch häufig auf Bilder zurück, ohne Parametrisierungen explizit anzugeben. Die an die Definition 16.2 angepasste Verallgemeinerung des Hauptsatzkriteriums (16.10) lautet:

**Proposition 17.2.** Sei  $U \subset \mathbb{C}$  offen mit der Eigenschaft, dass für je zwei Punkte  $a, b \in U$  ein Weg von  $a$  nach  $b$  in  $U$  existiert. Dann besitzt eine stetige Funktion  $f : U \rightarrow \mathbb{C}$  genau dann eine komplexe Stammfunktion, wenn für jeden geschlossenen stückweise stetig differenzierbaren Weg  $\gamma$

$$\int_{\gamma} f(w)dw = 0 \quad (17.4)$$

In diesem Fall kann für jedes  $z_* \in U$  eine Stammfunktion von  $f$  rekonstruiert werden durch  $F(z) = \int_{\gamma(z_*, z)} f(w)dw$  für einen beliebigen Weg  $\gamma(z_*, z)$  von  $z_*$  nach  $z$  in  $U$  und es gilt für jeden (stückweise stetig differenzierbaren) Weg  $\gamma$  von  $a$  nach  $b$  in  $U$ :

$$\int_{\gamma} f(w)dw = F(b) - F(a) \quad (17.5)$$

*Beweis.* Siehe Lehrbücher. □

Im Unterschied zu 16.1 ist es aber *nicht richtig*, dass für beliebige nicht konvexe oder nicht-sterneförmige offene Mengen (17.4) für jede stetige Funktion gilt:

**Beispiel 17.3.** Das Integral der Funktion  $\mathbb{C} \setminus \{0\} \ni z \mapsto \frac{1}{z}$  über den Kreis vom Radius  $R > 0$  um den Nullpunkt ist

$$\begin{aligned} \int_{C_R^+(0)} \frac{1}{w} dw &= \int_0^{2\pi} \frac{1}{Re^{it}} \frac{d}{dt}(Re^{it}) dt = \int_0^{2\pi} \frac{1}{Re^{it}} Re^{it} i dt = \int_0^{2\pi} i dt \\ &= 2\pi i \end{aligned} \quad (17.6)$$

und insbesondere nicht Null. Man bemerke, dass (17.6) von  $R$  unabhängig ist, und zur Sicherheit rechne man noch einmal für beliebiges  $z_0 \in \mathbb{C}$ :

$$\int_{C_R^+(z_0)} \frac{1}{w - z_0} dw = 2\pi i \quad \text{und auch noch} \quad \int_{C_R^-(z_0)} \frac{1}{w - z_0} dw = -2\pi i \quad (17.7)$$

Fazit:  $z \mapsto \frac{1}{z}$  hat auf  $\mathbb{C} \setminus \{0\}$  keine Stammfunktion. Hingegen hatten Sie in Aufgabe 10.3 gesehen, dass auf der geschlitzten Ebene  $\mathbb{C} \setminus (-\infty, 0]$  der Hauptzweig des Logarithmus, definiert durch

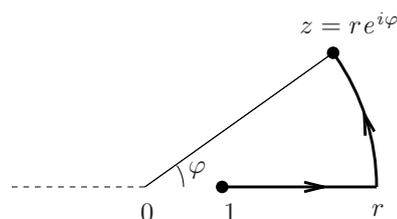
$$\text{Log}(z) = \ln r + i\varphi \quad (17.8)$$

für  $z = re^{i\varphi}$  mit  $r > 0$ ,  $\varphi \in (-\pi, \pi)$ , holomorph ist mit Ableitung  $\text{Log}'(z) = \frac{1}{z}$ . Auf dieser kleineren Menge also *existiert* eine Stammfunktion. In der Tat ist ja die geschlitzte Ebene sternförmig mit Sternmittelpunkt  $1 \in z$  und es gilt

$$\text{Log}(z) = \int_{l(1, z)} \frac{1}{w} dw \quad (17.9)$$

Statt auf der direkten Strecke  $l(1, z)$  kann auch entlang des rechts skizzierten Weges integriert werden. Das gerade Stück ist  $[1, r] \ni t \mapsto t$ , das gebogene  $[0, \varphi] \ni t \mapsto r e^{it}$ , und

$$\int_1^r \frac{1}{s} ds + \int_0^\varphi \frac{1}{r e^{it}} i r e^{it} dt = \ln r + i\varphi \quad (17.10)$$



- Es gehören zur Diskussion jetzt noch Antworten auf die folgenden zwei Fragen:
- (1) Ist es *notwendig*, dass  $f$  holomorph ist, damit sie auf konvexen Mengen eine Stammfunktion besitzt?
  - (2) Gibt es notwendige und hinreichende Charakterisierungen von Mengen  $U$  dafür, dass jede holomorphe Funktion auf  $U$  eine Stammfunktion besitzt?

Die Intention der Frage (1) ist *lokal*, d.h. man denkt an “kleine” Umgebungen eines gegebenen Punktes  $z_0 \in U$ , insbesondere Kreisscheiben  $B_R(z_0)$ . Die Antwort ist Ja, siehe 17.5.

Die Antwort auf die globale Frage (2) ist, dass genau dann jede auf einer offenen zusammenhängenden Menge  $U$  holomorphe Funktion eine Stammfunktion besitzt, wenn  $U$  “einfach zusammenhängend” ist. Anschaulich gesprochen bedeutet dies, dass  $U$  keine “Löcher im Inneren” hat oder, etwas präziser gesagt, dass sich jeder geschlossene Weg in  $U$  stetig auf einen Punkt zusammenziehen lässt. Eine vollständige Untersuchung würde uns hier zu weit führen. Ein Teil der Probleme rührt auch von der Mathematiker Wunsch her, möglichst allgemeine (auch nicht differenzierbare...) Wege zur Integration zuzulassen, und für solche ist es nicht so einfach zu zeigen, dass Inneres und Äusseres wohldefiniert sind.

In der Praxis werden stückweise stetig differenzierbare Wege genügen, und wir beschränken uns ausserdem auf solche, die nie stoppen und sich nicht selbst schneiden (sog. “einfache Kurven”). Dann lässt sich intuitiv bequem fassen die

Aussage: Sei  $\gamma : [\alpha, \beta] \rightarrow \mathbb{C}$  einfach geschlossen (d.h.  $\gamma(\alpha) = \gamma(\beta)$ , aber  $\gamma|_{[\alpha, \beta]}$  ist injektiv) und stückweise stetig differenzierbar mit  $\dot{\gamma} \neq 0$  auf allen Stücken. Sei  $U \subset \mathbb{C}$  eine offene Menge, die  $\gamma([\alpha, \beta])$  und ihr gesamtes Innere enthält. Dann gilt für jede holomorphe Funktion  $f : U \rightarrow \mathbb{C}$

$$\int_\gamma f(w) dw = 0 \quad (17.11)$$

Der nächste entscheidende Schritt ergibt sich aus der Verbindung von (17.11) mit dem Integral (17.6):

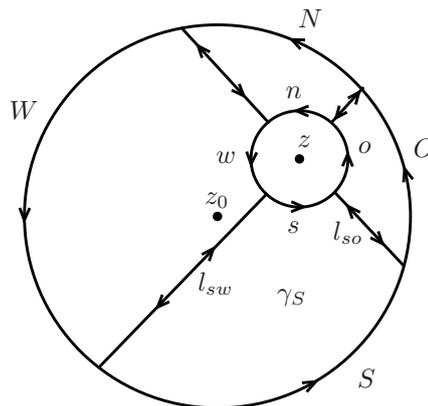
**Theorem 17.4.** Sei  $U \subset \mathbb{C}$  offen und  $f : U \rightarrow \mathbb{C}$  holomorph. Angenommen, für  $z_0 \in U$  und  $R > 0$  liegt  $\overline{B_R(z_0)} \subset U$ . Dann gilt für alle  $z \in B_R(z_0)$ :

$$f(z) = \frac{1}{2\pi i} \int_{C_R^+(z_0)} \frac{f(w)}{w - z} dw \quad (17.12)$$

§17. CAUCHY-FORMELN

*Beweis:* Beh.: Für  $r > 0$  mit  $B_r(z) \subset B_R(z_0)$  gilt

$$\int_{C_R^+(z_0)} \frac{f(w)}{w-z} dw = \int_{C_r^+(z)} \frac{f(w)}{w-z} dw \quad (17.13)$$



Bew.: Wir zerlegen wie skizziert

$$\begin{aligned} \int_{C_R^+(z_0)} &= \int_O + \int_N + \int_W + \int_S \\ \int_{C_r^+(z)} &= \int_o + \int_n + \int_w + \int_s \end{aligned} \quad (17.14)$$

und addieren und subtrahieren die geraden Stücke in die Nebenhimmelsrichtungen, sodass

$$\int_{C_R^+(z_0)} - \int_{C_r^+(z)} = \int_{\gamma_O} + \int_{\gamma_N} + \int_{\gamma_W} + \int_{\gamma_S} \quad (17.15)$$

wo z.B.

$$\int_{\gamma_S} = \int_S - \int_{l_{so}} - \int_s + \int_{l_{sw}} \quad (17.16)$$

Nun liegt jeder der vier Wege mit seinem Inneren in einer (anderen) offenen konvexen Menge  $\subset U$ , auf der der Integrand

$$w \mapsto \frac{f(w)}{w-z} \quad (17.17)$$

jeweils holomorph ist. Damit folgt aus (17.11), dass  $\int_{\gamma_S} = 0$  etc., und daraus dann (17.13).

2. Schritt: Aus (17.7) folgt, dass für jedes  $r > 0$  wie oben

$$f(z) = f(z) \cdot \underbrace{\frac{1}{2\pi i} \int_{C_r^+(z)} \frac{1}{w-z} dw}_{=1} = \frac{1}{2\pi i} \int_{C_r^+(z)} \frac{f(z)}{w-z} dw \quad (17.18)$$

sodass zusammen mit (17.13) für jedes solches  $r$

$$\frac{1}{2\pi i} \int_{C_R^+(z_0)} \frac{f(w)}{w-z} dw - f(z) = \frac{1}{2\pi i} \int_{C_r^+(z)} \frac{f(w) - f(z)}{w-z} dw \quad (17.19)$$

Zum Beweis der Behauptung reicht es daher aus, wenn wir für jedes  $\epsilon > 0$  ein  $r > 0$  angeben, für das wir zeigen können, dass der Betrag der rechten Seite kleiner als  $\epsilon$  ist.

Bew.: Da  $f$  in  $z$  holomorph ist, ist der Differenzenquotient  $w \mapsto \frac{f(w)-f(z)}{w-z}$  per Definition 15.1 stetig in  $z$  fortsetzbar, und daher nach 7.6 auf der folgenkompakten Menge  $\overline{B_R(z_0)}$  beschränkt, sagen wir durch  $M > 0$ . Sei nun für  $\epsilon > 0$   $r > 0$  so klein, dass  $Mr < \epsilon$ . Dann folgt aus der Standardabschätzung (16.22) des Integrals

$$\left| \frac{1}{2\pi i} \int_{C_r^+(z)} \frac{f(w) - f(z)}{w-z} dw \right| \leq \frac{1}{2\pi} M \cdot \underbrace{l(C_r^+(z))}_{=2\pi r} < \epsilon \quad (17.20)$$

□

**Folgerungen**

Man prüft nach, dass der Integrand von (17.12), z.B. in der Standardparametrisierung  $w(t) = z_0 + Re^{it}$ ,  $\dot{w}(t) = Rie^{it}$ ,

$$f(z) = \frac{1}{2\pi i} \int_0^{2\pi} \frac{f(z_0 + Re^{it})}{z_0 + Re^{it} - z} \cdot Rie^{it} dt \tag{17.21}$$

die Voraussetzungen der Proposition 14.8 über parameterabhängige Integrale erfüllt. Insbesondere ist für festes  $t$  der Integrand in jedem  $z \in B_R(z_0)$  komplex differenzierbar und die Abbildung auf die Ableitung nach  $z$

$$(t, z) \mapsto \frac{d}{dz} \left( \frac{f(z_0 + Re^{it})}{z_0 + Re^{it} - z} \cdot Rie^{it} \right) = \frac{f(z_0 + Re^{it})}{(z_0 + Re^{it} - z)^2} \cdot Rie^{it} \tag{17.22}$$

ist stetig. Es folgt zunächst

$$f'(z) = \frac{1}{2\pi i} \int_{C_R^+(z_0)} \frac{f(w)}{(w - z)^2} dw, \tag{17.23}$$

und daraus durch vollständige Induktion, dass  $f$  unendlich oft komplex differenzierbar ist mit

$$f^{(n)}(z) = \frac{n!}{2\pi i} \int_{C_R^+(z_0)} \frac{f(w)}{(w - z)^{n+1}} dw, \tag{17.24}$$

Insbesondere folgt daraus wie angekündigt: Besitzt eine stetige Funktion  $f : U \rightarrow \mathbb{C}$  eine komplexe Stammfunktion, so ist  $f$  bereits holomorph. (Denn: Eine Stammfunktion  $F$  von  $f$  ist holomorph und daher wie eben gezeigt unendlich oft komplex differenzierbar. Insbesondere ist  $f = F'$  holomorph.)

Holomorphie ist also notwendig und hinreichend für die lokale Existenz einer Stammfunktion, und damit ist die Frage (1) auf Seite 136 beantwortet. Aus der folgenden Verbindung zu 15.4 wird klar, warum Kreise als Integrationswege eine besondere Rolle spielen.

**Proposition 17.5.** *Sei  $U \subset \mathbb{C}$  offen und  $f : U \rightarrow \mathbb{C}$  holomorph. Sei  $z_0 \in U$  und  $R > 0$  mit  $\overline{B_R(z_0)} \subset U$ . Dann besitzt  $f$  eine für alle  $z \in B_R(z_0)$  konvergente Potenzreihenentwicklung*

$$f(z) = \sum_{k=0}^{\infty} a_k (z - z_0)^k \tag{17.25}$$

mit den Koeffizienten

$$a_k = a_k^{(z_0)} = \frac{f^{(k)}(z_0)}{k!} = \frac{1}{2\pi i} \int_{C_R^+(z_0)} \frac{f(w)}{(w - z_0)^{k+1}} dw \tag{17.26}$$

*Bemerkungen.* · Hier ist  $R$  stets kleiner (und nie gleich) als der Konvergenzradius der Potenzreihe (17.25): Da  $U$  offen ist, existiert stets noch eine etwas grössere Kreisscheibe  $\overline{B_{R'}(z_0)} \subset U$  mit  $R' > R$ , in deren Inneren die Potenzreihe (17.25) immer noch konvergiert.

§17. CAUCHY-FORMELN

*Beweis.* Wie oben nutzen wir die Trennung zwischen  $f$  und  $z$  in der Integraldarstellung (17.12) aus. Für festes  $z \in B_R(z_0)$  ist

$$\left| \frac{z - z_0}{w - z_0} \right| = \frac{|z - z_0|}{R} =: \rho < 1 \quad (17.27)$$

konstant für alle  $w \in C_R(z_0)$ . Dort gilt daher (geometrische Reihe)

$$\left| \frac{1}{w - z} - \sum_{k=0}^n \frac{(z - z_0)^k}{(w - z_0)^{k+1}} \right| = \left| \frac{1}{w - z_0} \sum_{k=n+1}^{\infty} \frac{(z - z_0)^k}{(w - z_0)^k} \right| \leq \frac{1}{R} \cdot \frac{\rho^{n+1}}{1 - \rho} \quad (17.28)$$

Mit (17.12) und der Standardabschätzung des Integrals folgt

$$\begin{aligned} & \left| f(z) - \sum_{k=0}^n a_k (z - z_0)^k \right| \\ &= \left| \frac{1}{2\pi i} \int_{C_R^+(z_0)} \frac{f(w)}{w - z} dw - \sum_{k=0}^n \frac{1}{2\pi i} \int_{C_R^+(z_0)} \frac{f(w)}{(w - z_0)^{k+1}} dw \cdot (z - z_0)^k \right| \\ &= \left| \frac{1}{2\pi i} \int_{C_R^+(z_0)} f(w) \left( \frac{1}{w - z} - \sum_{k=0}^n \frac{(z - z_0)^k}{(w - z_0)^{k+1}} \right) dw \right| \\ &\leq \frac{1}{2\pi} \cdot \max\{|f(w)| \mid w \in C_R(z_0)\} \cdot \frac{1}{R} \cdot \frac{\rho^{n+1}}{1 - \rho} \cdot (2\pi R) \\ &\leq \max\{|f(w)| \mid w \in C_R(z_0)\} \cdot \frac{\rho^{n+1}}{1 - \rho} \end{aligned} \quad (17.29)$$

Für  $n \rightarrow \infty$  konvergiert also die Reihe wie behauptet gegen  $f(z)$ . □

*Bemerkungen.* Mittels (17.11) und (17.7) lässt sich als Verallgemeinerung von 17.4 zeigen, dass für einen einfach geschlossenen Weg  $\gamma$ , der mitsamt seinem Inneren ganz in  $U$  liegt, und einen Punkt  $z \in U$  im Inneren von  $\gamma$

$$\int_{\gamma} \frac{f(w)}{w - z} dw = \pm 2\pi i \cdot f(z) \quad (17.30)$$

je nach Orientierung von  $\gamma$ . Für  $z$  im Äusseren von  $\gamma$  ist das Integral 0. Und für einen ganz neugierigen geschlossenen Weg, der sich in  $U$  zu einem Punkt zusammenziehen lässt und  $z$  nicht trifft, gilt

$$\int_{\gamma} \frac{f(w)}{w - z} dw = 2\pi i \cdot f(z) \cdot W(\gamma, z) \quad (17.31)$$

wobei für jedes  $z \in \mathbb{C} \setminus \gamma([\alpha, \beta])$

$$W(\gamma, z) := \frac{1}{2\pi i} \int_{\gamma} \frac{1}{w - z} dw \in \mathbb{Z} \quad (17.32)$$

die ‘‘Windungszahl’’ von  $\gamma$  um  $z$  berechnet.

## § 18 Residuensatz

Im vorigen § hatten wir aus der Cauchy-Formel 17.4 gefolgert, dass holomorphe Funktionen unendlich oft differenzierbar sind, und sich die Ableitungen nach den Formeln (17.24) durch Mittelung über geeignete Kreise berechnen lassen. Nun wollen wir die Logik umkehren, um Integrale durch Ableiten zu berechnen. Dabei beruht nach wie vor alles auf

- (i) der Invarianz des Integrals einer holomorphen Funktion unter stetigen Deformationen des Integrationsweges, solange entweder die Endpunkte festgehalten werden oder der Weg geschlossen ist und ausserdem die Deformation innerhalb der Menge stattfindet, in der die Funktion holomorph ist, und
- (ii) der nun elementaren Feststellung, dass für  $n \in \mathbb{Z}$  und  $R > 0$

$$\int_{C_R^+(0)} w^n dw = \begin{cases} 0 & n \neq -1 \\ 2\pi i & n = -1 \end{cases} \quad (18.1)$$

Denn für  $n \neq -1$  hat die Funktion  $z \mapsto z^n$  auf  $\mathbb{C} \setminus \{0\}$  die Stammfunktion  $\frac{1}{n+1}z^{n+1}$ , und für  $n = -1$  gilt (17.6).

· Ist beispielsweise eine Funktion  $f$  durch eine auf  $\overline{B_R(0)}$  konvergente Potenzreihe  $\sum a_n z^n$  gegeben, so lässt sich die Formel (17.26) als direkte Konsequenz von (18.1) auffassen:

$$\begin{aligned} \int_{C_R^+(0)} \frac{f(w)}{w^{k+1}} dw &= \int_{C_R^+(0)} \sum_{n=0}^{\infty} a_n w^{n-k-1} dw = \int_{C_R^+(0)} \sum_{n=-k-1}^{\infty} a_{n+k+1} w^n dw \\ &= \underbrace{\sum_{n=-k-1}^{-2} a_{n+k+1} \int_{C_R^+(0)} w^n dw}_{=0} + \int_{C_R^+(0)} \frac{a_k}{w} dw + \underbrace{\int_{C_R^+(0)} \sum_{n=0}^{\infty} a_{n+k+1} w^n dw}_{=0 \text{ s. (17.11)}} \\ &= 2\pi i \cdot a_k \end{aligned} \quad (18.2)$$

da der Rest der Potenzreihe eine holomorphe Funktion darstellt.

· Und noch ein Physiker-Argument für (18.1): Das Integral ist wegen der im Beweis von 17.4 benutzten Version von (i) unabhängig von  $R$ , andererseits transformiert sich der Integrand unter der Substitution  $w \rightarrow R w$  mit einem Faktor  $R^{n+1}$ . Das Integral kann also nur für  $n = -1$  von null verschieden sein.

**Definition 18.1.** Für  $z_0 \in \mathbb{C}$  und zwei komplexe Folgen  $(a_n)_{n=0,1,\dots}$ ,  $(a_{-n})_{n=1,2,\dots}$  heisst das Paar von Reihen

$$\sum_{n=0}^{\infty} a_n (z - z_0)^n, \quad \sum_{n=1}^{\infty} a_{-n} (z - z_0)^{-n} \quad (18.3)$$

Laurent-Reihe mit Entwicklungspunkt  $z_0$ . Sie heisst konvergent, wenn beide Reihen in (18.3) getrennt konvergieren. In diesem Fall ist der Wert der Laurent-Reihe die Summe der Grenzwerte. Man schreibt für die Laurent-Reihe bzw. ihre Werte auch  $\sum_{n=-\infty}^{\infty} a_n (z - z_0)^n$ .

§18. RESIDUENSATZ

· Gemäss 5.19, (5.45) konvergiert die Reihe mit positiven Potenzen von  $z - z_0$  für

$$|z - z_0| < R := (\limsup |a_n|^{1/n})^{-1} \quad (18.4)$$

und analog die mit negativen Potenzen für

$$|z - z_0|^{-1} < S^{-1} := (\limsup |a_{-n}|^{1/n})^{-1} \quad (18.5)$$

und nach Anwendung von Proposition 15.3 folgt daraus, dass die Laurent-Reihe  $L(z) := \sum_{n=-\infty}^{\infty} a_n(z - z_0)^n$  auf dem Kreisring

$$A_{R,S}(z_0) := \{z \in \mathbb{C} \mid S < |z - z_0| < R\} \quad (18.6)$$

konvergiert, dort eine holomorphe Funktion darstellt, und gliedweise abgeleitet werden darf. (Interessant ist dies natürlich eigentlich nur dann, wenn  $S < R$ .) Ein wesentlicher Unterschied zu Potenzreihen ist, dass Laurent-Reihen mit  $a_{-1} \neq 0$  auf ihrem Konvergenz-Kreisring keine Stammfunktion besitzen.

Beispiel: Für  $|z| > 1$  :  $\sum_{n=0}^{\infty} z^{-n} = \frac{1}{1 - \frac{1}{z}} = \frac{z}{1 - z}$

· Für  $|z| > 0$  :  $\sum_{n=0}^{\infty} \frac{z^{-n}}{n!} = \exp\left(\frac{1}{z}\right)$

Umgekehrt gilt aber auch:

**Proposition 18.2.** Sei  $U \subset \mathbb{C}$  offen und  $f : U \rightarrow \mathbb{C}$  holomorph. Sei  $z_0 \in \mathbb{C}$  und  $\infty \geq R > S \geq 0$  mit  $A_{R,S}(z_0) \subset U$ . (Im Allgemeinen muss  $z_0$  nicht in  $U$  sein!) Dann besitzt  $f$  eine für alle  $z \in A_{R,S}(z_0)$  konvergente Entwicklung in eine Laurentreihe:

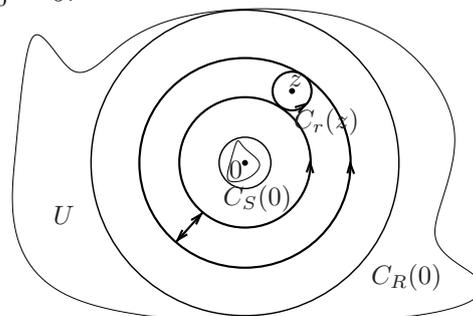
$$f(z) = \sum_{n=-\infty}^{\infty} a_n(z - z_0)^n \quad \forall z \in A_{R,S}(z_0) \quad (18.7)$$

*Beweis.* Wir nehmen der Einfachheit halber  $z_0 = 0$ .

Sei  $r > 0$  so, dass  $\overline{B_r(z)} \subset A_{R,S}(0)$ . Dann gilt nach 17.4

$$f(z) = \frac{1}{2\pi i} \int_{C_r^+(z)} \frac{f(w)}{w - z} dw \quad (18.8)$$

Durch Nachahmen von (17.13) anhand der Skizze wird dies gleich



$$f(z) = \frac{1}{2\pi i} \int_{C_{|z|+r}^+(0)} \frac{f(w)}{w - z} dw - \frac{1}{2\pi i} \int_{C_{|z|-r}^+(0)} \frac{f(w)}{w - z} dw \quad (18.9)$$

Im ersten Integranden gilt  $|z| < |w|$ , d.h.  $\frac{1}{w-z} = \sum_{n=0}^{\infty} \frac{z^n}{w^{n+1}}$  im zweiten  $|z| > |w|$  und  $\frac{1}{w-z} = -\sum_{n=1}^{\infty} \frac{w^{n-1}}{z^n} = -\sum_{n=-\infty}^{-1} \frac{z^n}{w^{n+1}}$ . Mit dem gleichen Konvergenzargument wie im Beweis von 17.5 folgt die Behauptung, wobei

$$a_n = \frac{1}{2\pi i} \cdot \left\{ \begin{array}{ll} \int_{C_{|z|+r}^+(0)} \frac{f(w)}{w^{n+1}} dw & n \geq 0 \\ \int_{C_{|z|-r}^+(0)} \frac{f(w)}{w^{n+1}} dw & n < 0 \end{array} \right\} = \frac{1}{2\pi i} \int_{C_\rho^+(0)} \frac{f(w)}{w^{n+1}} \quad \forall n \in \mathbb{Z} \quad (18.10)$$

für jedes  $S < \rho < R$ , da die Integrale von solchen  $\rho$  unabhängig sind. □

Mit Zuversicht erweitern wir nun erstmals unseren Begriff von Funktionen auf solche, die streng genommen gar nicht überall definiert sind.

**Definition 18.3.** Sei  $U \subset \mathbb{C}$  offen. Eine *holomorphe Funktion mit isolierten Singularitäten* auf  $U$  ist eine endliche oder abzählbare Teilmenge  $\{z_1, z_2, \dots\}$  von (verschiedenen) Punkten  $z_i \in U$ , welche keinen Häufungspunkt in  $U$  hat, zusammen mit einer holomorphen Funktion  $f : U \setminus \{z_1, z_2, \dots\} \rightarrow \mathbb{C}$ .

· Aus der Tatsache, dass die Ausnahmemenge  $\{z_1, z_2, \dots\}$  keinen Häufungspunkt in  $U$  hat, folgt, dass für jedes  $z_i$  ein  $R_i > 0$  existiert so, dass  $B_{R_i}(z_i) \subset U$  und  $f$  auf der “punktierten Scheibe”  $B_{R_i}(z_i) \setminus \{z_i\}$  definiert und holomorph ist. Aus 18.2 folgt dann, dass  $f$  auf  $B_{R_i}(z_i) \setminus \{z_i\} = A_{R_i,0}(z_i)$  durch eine Laurentreihe dargestellt wird, d.h.

$$f(z) = \sum_{n=-\infty}^{\infty} a_n^{(i)}(z - z_i)^n \quad \forall z \in B_{R_i}(z_i) \setminus \{z_i\} \quad (18.11)$$

Man nennt dabei

$$H_i(z) := \sum_{n=-\infty}^{-1} a_n^{(i)}(z - z_i)^n \quad (18.12)$$

den *Hauptteil* der Laurentreihe ( $\sum_{n=0}^{\infty}$  den Nebenteil), und

$$a_{-1}^{(i)} =: \text{Res}_{z_i}(f) \quad (18.13)$$

das *Residuum* von  $f$  in  $z_i$ . Gemäss (18.10) gilt nämlich

$$\text{Res}_{z_i}(f) = \frac{1}{2\pi i} \int_{C_{\rho}^+(z_i)} f(w) dw \quad (18.14)$$

für alle  $0 < \rho < R_i$  (was man aber für die Berechnung selten benutzt).

· Verschwindet der Hauptteil, d.h. gilt  $a_n^{(i)} = 0 \forall n < 0$ , so nennt man die zugehörige Singularität *hebbar*. Das ist genau dann der Fall, wenn sich  $f$  in  $z_i$  holomorph fortsetzen lässt. (Es liegt also gar keine Singularität vor.)

· Ist der Hauptteil endlich, d.h. gilt für ein  $k_i > 0$   $a_n^{(i)} = 0$  für  $n < -k_i$  und  $a_{-k_i}^{(i)} \neq 0$ , so nennt man die zugehörige Singularität einen *Pol der Ordnung  $k_i$* .

· Ist der Hauptteil unendlich, so heisst die Singularität *wesentlich* (oder auch “von unendlicher Ordnung”).

· Im Folgenden interessieren uns nur Pole endlicher Ordnung. Eine holomorphe Funktion mit isolierten Singularitäten auf  $U$ , welche allesamt Pole (sc. endlicher Ordnung) sind, heisst *meromorphe Funktion* auf  $U$ .

· Bei meromorphen Funktionen macht es Sinn, den Wertebereich von  $\mathbb{C}$  auf  $\mathbb{C} \cup \{\infty\}$  zu erweitern und den Polen formal den Wert  $\infty$  zuordnend  $f : U \rightarrow \mathbb{C} \cup \{\infty\}$  zu schreiben.

Diese Vorbereitungen versetzen uns nun in die Lage, eine für alle praktischen Zwecke hinreichende Version des Residuensatzes zu formulieren. Zuvor noch eine Formel zur Berechnung von Residuen meromorpher Funktionen, welche direkt aus der Laurententwicklung und der Taylorformel (17.26) herzuleiten ist.

§ 18. RESIDUENSATZ

**Lemma 18.4.** Ist  $z_0 \in U$  ein Pol der Ordnung  $k$  einer meromorphen Funktion  $f$  auf  $U$ , so gilt

$$\operatorname{Res}_{z_0}(f) = \lim_{z \rightarrow z_0} \frac{1}{(k-1)!} \left( \frac{d}{dz} \right)^{k-1} ((z-z_0)^k f(z)) \quad (18.15)$$

Persönlich finde ich das direkte Einsetzen von Potenzreihen ineinander ökonomischer, da es die für die Benutzung der Formel (18.15) sowieso benötigte Ordnung der Polstelle gleich mitliefert.

Beispiel: Für kleines  $z$  ist

$$\begin{aligned} \frac{1}{z(\exp(z) - 1)} &= \frac{1}{z(z + \frac{z^2}{2} + \dots)} = \frac{1}{z^2 + \frac{z^3}{2} + \dots} = \frac{1}{z^2} \cdot \frac{1}{1 + \frac{z}{2} + \dots} \\ &= \frac{1}{z^2} (1 - \frac{z}{2} + \mathcal{O}(z^2)) = \frac{1}{z^2} - \frac{1}{2z} + \mathcal{O}(1) \\ \Rightarrow \operatorname{Res}_0 \frac{1}{z(\exp(z) - 1)} &= -\frac{1}{2} \end{aligned} \quad (18.16)$$

**Theorem 18.5.** Sei  $f$  eine meromorphe Funktion auf einer offenen Menge  $U \subset \mathbb{C}$ . Dann gilt für jeden positiv orientierten einfach geschlossenen stückweise stetig differenzierbaren Weg  $\gamma$ , der mitsamt seinem Inneren ganz in  $U$  liegt und nicht die Pole von  $f$  trifft,

$$\int_{\gamma} f(w) dw = 2\pi i \cdot \sum_{\substack{\text{Pole } z_i \text{ im} \\ \text{Inneren von } \gamma}} \operatorname{Res}_{z_i}(f) \quad (18.17)$$

*Bemerkungen.* Der in der Aussage vorkommende Begriff des “Inneren” eines einfach geschlossenen Wegs kann mit Hilfe der Windungszahl (17.32) sauber definiert werden. In den Anwendungen ist es aber wichtiger, sich über die Orientierung Rechenschaft abzulegen. Es sei auch darauf hingewiesen, dass wir im Beweis die Aussagen (17.11), (17.30) benutzen, die wir nicht für alle fraglichen Wege streng bewiesen haben.

*Beweis.* · Zunächst wollen wir begründen, dass die Summe auf der rechten Seite von (18.17) endlich ist: Da  $\gamma$  zusammen mit seinem Inneren beschränkt und abgeschlossen sind (das folgt aus der genauen Definition des “Inneren”, siehe etwa (17.30)), so müssten, falls unendlich viele Polstellen von  $f$  im Inneren lägen, diese einen Häufungspunkt in  $U$  haben (s. 4.25), im Widerspruch zur Definition einer meromorphen Funktion.

· Es seien für  $r \in \mathbb{N}$  (oder allenfalls 0)  $z_1, \dots, z_r \in U$  die Polstellen von  $f$  im Inneren von  $\gamma$ , und für  $i = 1, \dots, r$

$$H_i(z) = \sum_{n=-k_i}^{-1} a_n^{(i)} (z - z_i)^n \quad (18.18)$$

$a_n^{(i)} \in \mathbb{C}$ , der Hauptteil von  $f$  am  $i$ -ten Pol.  $H_i(z)$  ist eine meromorphe Funktion auf  $\mathbb{C}$  mit einem Pol bei  $z_i$ , insbesondere holomorph auf  $\mathbb{C} \setminus \{z_i\}$ .

· In einer Umgebung eines jeden  $z_i$  verschwindet dann der Hauptteil von

$$\tilde{f}(z) := f(z) - \sum_{i=1}^r H_i(z) \quad (18.19)$$

so dass sich  $\tilde{f}$  zu einer holomorphen Funktion auf einer Umgebung von  $\gamma$  mitsamt seinem Inneren fortsetzen lässt. Aus (17.11) folgt  $\int_{\gamma} \tilde{f}(w)dw = 0$  und daher

$$\int_{\gamma} f(w)dw = \sum_{i=1}^r \int_{\gamma} H_i(w)dw \quad (18.20)$$

· Nun ist für jedes  $i$

$$\int_{\gamma} \sum_{n=-k_i}^{-2} a_n^{(i)}(z - z_i)^n = 0 \quad (18.21)$$

denn der Integrand besitzt auf ganz  $\mathbb{C}$  die Stammfunktion

$$\sum_{n=-k_i}^{-2} \frac{a_n^{(i)}}{n+1} (z - z_i)^{n+1} \quad (18.22)$$

Es folgt

$$\int_{\gamma} f(w)dw = \sum_{i=1}^r \int_{\gamma} \frac{a_{-1}^{(i)}}{w - z_i} dw \stackrel{(17.30)}{=} 2\pi i \cdot \sum_{i=1}^r \text{Res}_{z_i}(f) \quad (18.23)$$

□

*Bemerkungen.* Die Verallgemeinerung von 18.5 für einen nicht notwendig einfach geschlossenen Weg  $\gamma$ , der sich in  $U$  zu einem Punkt zusammenziehen lässt und die isolierten Singularitäten der (sonst) holomorphen Funktion  $f$  auf  $U$  vermeidet, lautet

$$\int_{\gamma} f(w)dw = 2\pi i \cdot \sum_i \text{Res}_{z_i}(f) \cdot W(\gamma, z_i) \quad (18.24)$$

wobei in der Summe die Windungszahl (17.32) weiter nur für endlich viele Singularitäten  $z_i$  von null verschieden ist.

### Auswertung reeller Integrale

Die folgenden Beispiele enthalten als gemeinsame Zutaten:

(1) die Identifikation eines gegebenen reellen Integrals als Kurvenintegral einer meromorphen Funktion. Dabei muss der reelle Integrand nicht unbedingt “erhalten” bleiben, sondern kann als Real-/Imaginärteil, oder auch eine sonstige Linearkombination von komplexen Integranden auftauchen.

(2) das Schliessen der Wege in der komplexen Ebene unter Berücksichtigung von (3). Bei uneigentlichen Integralen werden häufig Grenzwerte über Familien von Wegen gebildet.

(3) das Abschätzen von bei (2) eingeführten zusätzlichen Beiträgen, gegebenenfalls mit einer Anpassung der gewählten Schliessung.

(4) die Auswertung des geschlossenen Kurvenintegrals mit dem Residuensatz.

Die genauen Rezepte der gängigen Klassen unterscheiden sich allerdings in den Details, was man am besten durch eigenes Ausprobieren schätzen lernt.

§ 18. RESIDUENSATZ

**Beispiel 18.6.** Seien  $a, b \in \mathbb{R}$  mit  $a > |b|$ . Dann erkennen wir in

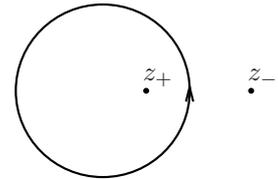
$$I(a, b) = \int_0^{2\pi} \frac{1}{a + b \cos \varphi} d\varphi \quad (18.25)$$

die Parametrisierung eines Integrals über den positiv orientierten Einheitskreis

$$\int_0^{2\pi} \frac{1}{a + \frac{b}{2}(e^{i\varphi} + e^{-i\varphi})} \frac{1}{i} e^{-i\varphi} \frac{d}{d\varphi} e^{i\varphi} d\varphi = \frac{1}{i} \int_{C_1^+(0)} \frac{1}{a + \frac{b}{2}(w + w^{-1})} \cdot \frac{1}{w} dw \quad (18.26)$$

eines auf  $\mathbb{C}$  meromorphen Integranden mit Polen bei

$$\frac{b}{2}z^2 + az + \frac{b}{2} = 0 \quad (18.27)$$



Für  $b = 0$  ist dies genau der Ursprung mit Residuum 1, d.h.

$$I(a, 0) = \frac{2\pi}{a} \quad (18.28)$$

Für  $b > 0$  lösen wir  $\frac{b}{2}z^2 + az + \frac{b}{2} = \frac{b}{2}(z - z_+)(z - z_-)$  mit  $z_{\pm} = \frac{-a \pm \sqrt{a^2 - b^2}}{b}$ . Wegen  $z_+ - z_- = \frac{2\sqrt{a^2 - b^2}}{b} \neq 0$  und  $z_+ z_- = 1$  liegt genau ein Pol (nämlich  $z_+$ ) in  $B_1(0)$ , und das Residuum dort ist

$$\lim_{z \rightarrow z_+} \frac{z - z_+}{\frac{b}{2}(z - z_+)(z - z_-)} = \frac{1}{\frac{b}{2}(z_+ - z_-)} = \frac{1}{\sqrt{a^2 - b^2}} \quad (18.29)$$

Es folgt

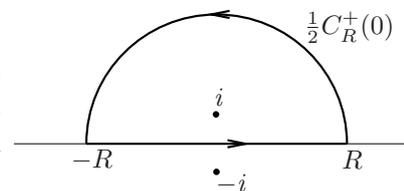
$$I(a, b) = \frac{2\pi}{\sqrt{a^2 - b^2}} \quad (18.30)$$

(was bei  $b \rightarrow 0$  noch in (18.28) übergeht).

**Beispiel 18.7.** Das uneigentliche Integral (vgl. (14.34))

$$\int_{-\infty}^{\infty} \frac{1}{1 + x^2} dx = \lim_{R \rightarrow \infty} \int_{-R}^R \frac{1}{1 + x^2} dx \quad (18.31)$$

dessen Konvergenz man durch Vergleich mit  $\int_1^{\infty} 1/x^2$  leicht verifiziert, und das sich auch elementar berechnen lässt, betrachten wir als den Grenzwert eines komplexen Kurvenintegrals von  $-R$  nach  $R$  der auf  $\mathbb{C}$  meromorphen Funktion  $z \mapsto \frac{1}{1+z^2}$  mit Polen bei  $\pm i$ .



Alsdann schliessen wir den Weg durch einen Halbkreis in der oberen Halbebene, auf dem (für  $R > 1$ )

$$\left| \frac{1}{1 + w^2} \right| \leq \frac{1}{R^2 - 1} \quad (18.32)$$

so dass durch Standardabschätzung

$$\left| \int_{\frac{1}{2}C_R^+(0)} \frac{1}{1 + w^2} dw \right| \leq \frac{1}{R^2 - 1} \cdot \pi R \rightarrow 0 \quad \text{für } R \rightarrow \infty \quad (18.33)$$

Es folgt

$$\int_{-\infty}^{\infty} \frac{1}{1+x^2} dx = 2\pi i \cdot \text{Res}_i \frac{1}{1+z^2} = 2\pi i \cdot \frac{1}{2i} = \pi \quad (18.34)$$

**Beispiel 18.8.** Das Integral

$$S := \int_{-1}^1 \sqrt{1-x^2} dx \quad (18.35)$$

ist Prototyp einer Klasse, die in der kanonischen Mechanik als Wirkungs- oder Periodenintegrale eine grosse Rolle spielen. Zur seiner Berechnung schieben wir (ganz unintuitiv) als erstes das Integral hinaus in die komplexe Ebene, d.h. wir schreiben unter Berufung auf 14.8

$$S = \lim_{\epsilon \rightarrow 0} \int_{-1}^1 \sqrt{1-(x-i\epsilon)^2} dx \quad (18.36)$$

wobei  $\sqrt{\cdot}$  die auf S. 129 diskutierte Fortsetzung der Wurzelfunktion in die komplexe Ebene ist:  $\sqrt{1-z^2} > 0$  für  $1-z^2 > 0$ , d.h.  $z \in (-1, 1)$ , und nicht definiert für  $z \in [1, \infty) \cup (-\infty, -1] \subset \mathbb{C}$ . Nach dem Hinausschieben ändert sich dann nichts, wenn wir den Integranden durch  $i\sqrt{z^2-1}$  ersetzen, wo jetzt  $\sqrt{z^2-1} > 0$  für  $z > 1$ ,  $\sqrt{z^2-1} < 0$  für  $z < -1$ , und nicht definiert für  $z \in [-1, 1]$ . Nach dem Rotieren des Schnitts ist also

$$S = \lim_{\epsilon \rightarrow 0} \int_{-1-i\epsilon}^{1-i\epsilon} i\sqrt{w^2-1} dw \quad (18.37)$$

oder bildlich

$$\int \xrightarrow{\quad} \begin{array}{c} \bullet \\ -1 \end{array} \xrightarrow{\quad} \begin{array}{c} \bullet \\ 1 \end{array} \xrightarrow{\quad} \quad = \int \begin{array}{c} \bullet \\ -1 \end{array} \xrightarrow{\quad} \begin{array}{c} \bullet \\ 1 \end{array} \xrightarrow{\quad}$$

$[-1, 1] - i\epsilon$   $[-1, 1] - i\epsilon$

Nun unterscheidet sich der Integrand auf der gegenüberliegenden Seite des Schnittes im Limes  $\epsilon \rightarrow 0$  gerade durch ein Vorzeichen, sodass

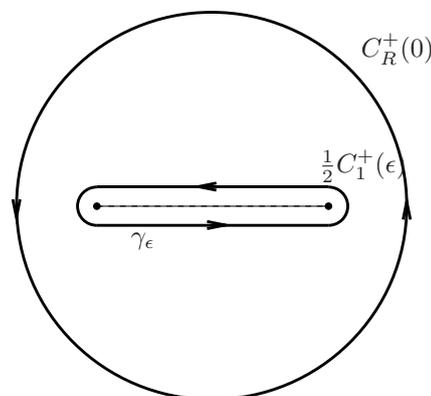
$$S = \frac{1}{2} \lim_{\epsilon \rightarrow 0} \left( \int_{-1-i\epsilon}^{1-i\epsilon} i\sqrt{w^2-1} dw - \int_{-1+i\epsilon}^{1+i\epsilon} i\sqrt{w^2-1} dw \right) \quad (18.38)$$

Auf den kleinen Halbkreisen  $\frac{1}{2}C_{\pm 1}^+(\epsilon)$  bleibt im Limes  $\epsilon \rightarrow 0$  der Integrand beschränkt, sodass

$$\lim_{\epsilon \rightarrow 0} \int_{\frac{1}{2}C_{\pm 1}^+(\epsilon)} i\sqrt{w^2-1} dw = 0 \quad (18.39)$$

Es folgt durch Schliessen des Weges

$$S = \frac{i}{2} \lim_{\epsilon \rightarrow 0} \int_{\gamma_\epsilon} \sqrt{w^2-1} dw \quad (18.40)$$



§ 18. RESIDUENSATZ

Nun liegt aber ein Integral einer holomorphen Funktion über einen geschlossenen Weg in  $\mathbb{C} \setminus [-1, 1]$  vor, sodass dieses unabhängig von  $\epsilon$  ist, und auch gleich dem Integral über den Kreis  $C_R^+(0)$  für  $R > 1$ . Es konvergiert dort die Entwicklung

$$\sqrt{w^2 - 1} = w - \frac{1}{2w} - \frac{1}{8w^3} + \dots \tag{18.41}$$

die wir gliedweise integrieren können, um als Endergebnis

$$S = \frac{i}{2} \int_{C_R^+(0)} \left( w - \frac{1}{2w} - \frac{1}{8w^3} + \dots \right) dw = \frac{i}{2} \int_{C_R^+(0)} \frac{-1}{2w} dw = -\frac{i}{2} \cdot \frac{2\pi i}{2} = \frac{\pi}{2} \tag{18.42}$$

zu erhalten. (Auch dieses Integral geht natürlich mit elementaren Mitteln...)

**Beispiel 18.9.** Das Gausssche Integral

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi} \tag{18.43}$$

spielt eine grundlegende Rolle unter anderem in der Fouriertransformation, der Wellenmechanik, der Wahrscheinlichkeitstheorie und dem Pfadintegralzugang zur Quantenfeldtheorie. Es gibt eine Vielzahl von Tricks zur Berechnung von (18.43), von denen der Residuensatz einer der weniger bekannten (und auch wenigst offensichtlichen) ist. Andererseits ist die Methode ziemlich rein ein-dimensional<sup>7</sup>, sodass wir sie hier vorführen können.

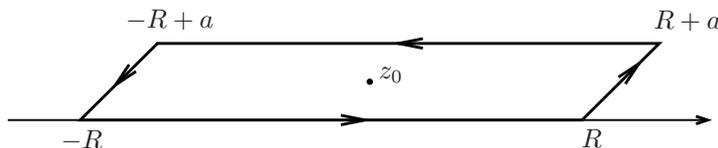
Wir setzen unser Vorwissen über das Ergebnis ein in die Zahl

$$a := \sqrt{\pi} e^{\pi i/4} = \sqrt{\frac{\pi}{2}}(1 + i) \quad \text{mit der Eigenschaft} \quad a^2 = \pi i \tag{18.45}$$

Dann erfüllt die auf  $\mathbb{C}$  meromorphe Funktion  $f(z) = \frac{e^{-z^2}}{1 + e^{-2az}}$  die Gleichung

$$\begin{aligned} f(z) - f(z + a) &= \frac{e^{-z^2}}{1 + e^{-2az}} - \frac{e^{-z^2 - 2az - a^2}}{1 + e^{-2az - 2a^2}} = \frac{e^{-z^2}}{1 + e^{-2az}} - \frac{e^{-z^2 - 2az - \pi i}}{1 + e^{-2az - 2\pi i}} \\ &= \frac{e^{-z^2}(1 + e^{-2az})}{1 + e^{-2az}} = e^{-z^2} \end{aligned} \tag{18.46}$$

Wir betrachten nun für  $R > 0$  das Integral von  $f$  über den Rand  $\square(R)$  des Parallelogramms mit den Ecken  $-R, R, R + a, -R + a$ .



<sup>7</sup>Die schnellste Methode wertet das Quadrat des Integrals durch zwei-dimensionale Integration in Polarkoordinaten aus:

$$\left( \int_{-\infty}^{\infty} e^{-x^2} dx \right)^2 = \int_{\mathbb{R}^2} e^{-x^2 - y^2} dx dy = \int_0^{\infty} \int_0^{2\pi} e^{-r^2} r d\varphi dr = 2\pi \int_0^{\infty} e^{-r^2} r dr = \pi \tag{18.44}$$

Einerseits enthalten dann wegen (18.46) die langen Seiten das gewünschte Integral:

$$\int_{-R}^R f(w)dw + \int_{R+a}^{-R+a} f(w)dw = \int_{-R}^R f(x)dx - \int_{-R}^R f(x+a)dx = \int_{-R}^R e^{-x^2} dx \quad (18.47)$$

Auf der rechten kurzen Seite ist mit  $w = R + ta$ ,  $t \in [0, 1]$

$$\begin{aligned} |e^{-2aw}| &= e^{\operatorname{Re}(-2a(R+ta))} = e^{-R\sqrt{2\pi}} \\ \text{und } |e^{-w^2}| &= e^{-\operatorname{Re}(R+ta)^2} = e^{-R^2 - tR\sqrt{2\pi}} \leq e^{-R^2} \end{aligned} \quad (18.48)$$

und daher gilt für  $R$  gross genug

$$\left| \frac{e^{-w^2}}{1 + e^{aw}} \right| \leq \frac{e^{-R^2}}{1 - e^{-R\sqrt{2\pi}}} \longrightarrow 0 \text{ für } R \rightarrow \infty \quad (18.49)$$

Nach einer analogen (bitte nachzuprüfenden) Abschätzung auf der linken kurzen Seite von  $\square(R)$  folgt mittels Standardabschätzung des Integrals

$$\lim_{R \rightarrow \infty} \left( \int_R^{R+a} f(w)dw + \int_{-R+a}^{-R} f(w)dw \right) = 0 \quad (18.50)$$

Andererseits liegen die Pole von  $f(z)$  bei  $-2az_n = (2n+1)\pi i = (2n+1)a^2$  für  $n \in \mathbb{Z}$ , d.h.  $z_n = \frac{2n+1}{2}a$ . Im Inneren des Parallelogramms befindet sich genau  $z_0 = \frac{a}{2}$ , und das Residuum ist

$$\operatorname{Res}_{z_0} f = \frac{e^{-z_0^2}}{\frac{d}{dz} \Big|_{z=z_0} (1 + e^{-2az})} = \frac{e^{-\pi i/4}}{-2ae^{-a^2}} = -\frac{i}{2\sqrt{\pi}} \quad (18.51)$$

Es folgt aus obigen Beziehungen und dem Residuensatz

$$\begin{aligned} \int_{-\infty}^{\infty} e^{-x^2} dx &= \lim_{R \rightarrow \infty} \int_{\square(R)} f(w)dw = 2\pi i \cdot \operatorname{Res}_{z_0} f = 2\pi i \cdot \frac{-i}{2\sqrt{\pi}} \\ &= \sqrt{\pi} \end{aligned} \quad (18.52)$$

wie behauptet. □

# KAPITEL 7

## SYMMETRIEN

Vektorräume spielen zur “Linearisierung der Physik” eine wichtige Rolle sowohl in der klassischen als auch in der modernen Theorie.

(i) In der klassischen Physik (und auch noch in der Relativitätstheorie) dient die Lineare Algebra im Sinne der Differential- und Integralrechnung vor allem zur “Approximation” quantitativer Zusammenhänge. Siehe hierzu insbesondere die Interpretation der mehr-dimensionalen Ableitung als lineare Abbildung auf den Tangentialräumen der Richtungen im §10, und die iterative Lösung gewöhnlicher Differentialgleichungen in §15. Die relevanten Vektorräume sind hauptsächlich reell und endlich-dimensional.

(ii) Im Unterschied dazu ist die lineare Struktur des quantenmechanischen Zustandsraums *fundamental* und *exakt*. Um dennoch Platz für interessante nichtlineare Phänomene zu bieten, müssen allerdings die Vektorräume *komplex* und meistens *unendlich-dimensional* sein.

(iii) Beiden gemeinsam ist die zentrale Rolle von Symmetrien und einer positiven (metrischen) Struktur.

### § 19 Multi-lineare Algebra

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

**Definition 19.1.** Ein Vektorraum über  $\mathbb{K}$  ist eine Menge  $V$  mit einem ausgezeichneten Element  $0 \in V$  und zwei binären Verknüpfungen,

$$\text{die Addition: } V \times V \ni (v_1, v_2) \mapsto v_1 + v_2 \in V \quad (19.1)$$

$$\text{und die Skalarmultiplikation: } \mathbb{K} \times V \ni (\lambda, v) \mapsto \lambda \cdot v := \lambda v \in V$$

welche die folgenden Rechenregeln erfüllen:

$$\text{Kommutativität: } \forall v_1, v_2 \in V : v_1 + v_2 = v_2 + v_1,$$

$$\text{Assoziativität: } \forall v_1, v_2, v_3 \in V : (v_1 + v_2) + v_3 = v_1 + (v_2 + v_3)$$

$$\forall v \in V, \lambda_1, \lambda_2 \in \mathbb{K} : \lambda_2 \cdot (\lambda_1 \cdot v) = (\lambda_2 \cdot \lambda_1) \cdot v \quad (19.2)$$

$$\text{Neutrale Elemente: } \forall v \in V : v + 0 = v, 1 \cdot v = v$$

$$\text{Distributivgesetze: } \forall v_1, v_2 \in V, \lambda \in \mathbb{K} : \lambda \cdot (v_1 + v_2) = \lambda \cdot v_1 + \lambda \cdot v_2$$

$$\forall v \in V, \lambda_1, \lambda_2 \in \mathbb{K} : (\lambda_1 + \lambda_2) \cdot v = \lambda_1 \cdot v + \lambda_2 \cdot v$$

derart, dass die Gleichung

$$v + x = 0 \quad (19.3)$$

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

Übungsaufgabe:  $0 \cdot v = 0 \in V \forall v \in V$ .

**Bemerkungen/Beispiele 19.2.** · 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

$$\mathbb{K}^n = \underbrace{\mathbb{K} \times \mathbb{K} \times \dots \times \mathbb{K}}_{n \text{ mal}} = \{(x^1, \dots, x^n)^T \mid x^i \in \mathbb{K}\} \quad (19.4)$$

mit komponentenweiser Addition und Skalarmultiplikation ein Vektorraum. Dabei ist  $\mathbb{K}^0 := \{0\}$ , der “triviale” Vektorraum. Der  $\mathbb{R}^n$  ist das kanonische Model des klassischen Konfigurationsraums.

· Für eine beliebige Menge  $M$  ist die Menge der Abbildungen

$$\mathcal{F}(M, \mathbb{K}) = \{f : M \rightarrow \mathbb{K}\} \quad (19.5)$$

mit punktweise Addition und Skalarmultiplikation ein  $\mathbb{K}$ -Vektorraum. (Wobei per Definition  $\mathcal{F}(\emptyset, \mathbb{K}) \cong \{0\}$ .) Der quantenmechanische Zustandsraum ist (moralisch gesprochen)  $\mathcal{F}(\mathcal{R}, \mathbb{C})$ , wobei  $\mathcal{R}$  die Menge möglicher Realtexte ist.

**Definition 19.3.** 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 (19.5) 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.

· Eine andere Variante von (19.5), welche gleich noch eine wichtige Rolle spielen wird, ist der Unterraum der *Abbildungen mit endlichem Träger*

$$\begin{aligned} \mathcal{F}^{\text{fin}}(M, \mathbb{K}) &:= \{f : M \rightarrow \mathbb{K} \mid f(m) = 0 \text{ ausser für endlich viele } m\} \\ &= \{f : M \rightarrow \mathbb{K} \mid |\text{supp}(f)| = \#\{m \in M \mid f(m) \neq 0\} < \infty\} \end{aligned} \quad (19.6)$$

### Struktur von Vektorräumen

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

**Definition 19.4.** · Für eine beliebige Teilmenge  $A \subset V$  eines Vektorraums ist die *lineare Hülle von A* die Menge

$$\text{span}(A) = \left\{ \sum_{i=1}^r \lambda_i v_i \mid r \in \mathbb{N}_0, \lambda_i \in \mathbb{K}, v_i \in A \right\} \quad (19.7)$$

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

## § 19. MULTI-LINEARE ALGEBRA

- 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

$$\sum_{i=1}^r \lambda_i v_i = 0 \quad (19.8)$$

folgt, dass

$$\lambda_i = 0 \forall i \quad (19.9)$$

(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

$$v = \sum_{i=1}^r v^i b_i \quad \text{mit } v^i \in \mathbb{K} \text{ und } b_i \in B \quad (19.10)$$

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

Ein wichtiges Ergebnis zur Struktur von Vektorräumen ist:

**Theorem 19.5.** (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  $\dim(V)$ .*

*Beweisidee.* Eine Basis ist ein “minimales Erzeugendensystem” und auch ein “maximales linear unabhängiges System”. Dies erlaubt die “Konstruktion” einer Basis aus “elementaren” mengentheoretischen Überlegungen. Für Details siehe Lehrbücher zur Linearen Algebra.  $\square$

- 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 (19.10) mit Nullen auf.

- Man beachte, dass in (19.7) 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. 4.2, unter der man dann beispielsweise wie in §5 Reihen von Vektoren in  $V$  auf Konvergenz untersuchen kann. Man erinnere sich aber daran, dass genau im unendlich-dimensionalen Fall nicht alle Normen äquivalent sind, und daher insbesondere auch davon abhängige Verallgemeinerungen des Basisbegriffs nicht eindeutig. Hierzu später mehr.

**Bemerkungen/Beispiele 19.6.** · Anfänglich scheinen die meisten Vektorräume “mit einer Basis geboren” zu werden. Beispielsweise ist

$$B = \left( e_i = (0, \dots, \underset{\substack{\uparrow \\ i\text{-te Stelle}}}{1}, \dots, 0)^T \right)_{i=1, \dots, n} \quad (19.11)$$

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

$$\text{Free}(M) = \{ \lambda_1 m_1 + \lambda_2 m_2 + \dots + \lambda_r m_r \mid \lambda_i \in \mathbb{K}, m_i \in M, m_i \neq m_j \text{ falls } i \neq j \} \quad (19.12)$$

mit Vektorraumoperationen, welche, im scheinbaren Unterschied zu allem Vorhergegangenen, erst “nachträglich” erklärt werden, ein Vektorraum mit Basis  $M$ .

Übungsaufgabe: Die Abbildung

$$\begin{aligned} \text{Free}(M) &\rightarrow \mathcal{F}^{\text{fin}}(M, \mathbb{K}) \\ \sum_{i=1}^r \lambda_i m_i &\mapsto \left( f : m \mapsto \begin{cases} \lambda_i & \text{falls } m = m_i \text{ für ein } i \\ 0 & \text{sonst} \end{cases} \right) \end{aligned} \quad (19.13)$$

ist ein Vektorraumisomorphismus.

In der Linearen Algebra lernt man jedoch zu schätzen, über Vektorräume “basisfrei” zu reden, bzw. zwischen verschiedenen Basen zu wechseln. Auch für das Verständnis der Quantenmechanik ist es unerlässlich zu akzeptieren, dass Basen von Vektorräumen zusätzliche physikalische Information darstellen<sup>8</sup> und keine *a priori* bevorzugt ist. Tatsächlich ist es für unendlich-dimensionale Vektorräume i.A. gar nicht *möglich*, eine Basis explizit anzugeben. (Die Konstruktion in 19.5 hängt vom Auswahlaxiom bzw. Zornschen Lemma ab.)

Ein lehrreiches Beispiel ergibt sich aus der

**Definition 19.7.** Sei  $V$  ein  $\mathbb{K}$ -Vektorraum und  $W \subset V$  ein Untervektorraum. Dann wird auf  $V$  durch

$$v_1 \sim v_2 :\Leftrightarrow v_1 - v_2 \in W \quad (19.14)$$

eine Äquivalenzrelation erklärt, und auf dem mengentheoretischen Quotienten

$$V/\sim = \{ [v] := \{ \tilde{v} \sim v \} \mid v \in V \} \quad (19.15)$$

erfüllen die Verknüpfungen

$$[v_1] + [v_2] := [v_1 + v_2], \quad \lambda[v] := [\lambda v] \quad (19.16)$$

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  $\dim W \geq 1$ , so ist  $\{[b_1], \dots, [b_n]\} \subset V/W$  nicht mehr linear unabhängig und bildet daher keine Basis mehr. Es gibt im Allgemeinen keine weitere Information zur Auswahl einer Basis von  $V/W$ .

<sup>8</sup>z.B. die Wahl zwischen der Orts- und Impulsdarstellung

**Basisdarstellung linearer Abbildungen**

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 mit  $i = 1, \dots, n$  und  $j = 1, \dots, m$  eindeutige  $A_i^j \in \mathbb{K}$  so, dass

$$\hat{A}(b_i) = \sum_{j=1}^m c_j A_i^j \quad \forall i = 1, \dots, n \tag{19.17}$$

und diese “Darstellung von  $\hat{A}$  bezüglich der ausgezeichneten Basen” legt die lineare Abbildung bereits vollständig fest. Umgekehrt existiert zu jeder solchen Matrix eine (eindeutige) lineare Abbildung  $V \rightarrow W$ .

· 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  $v$ ”:

$$\begin{aligned} \hat{A}(v^i b_i) &= v^i \hat{A}(b_i) = c_j A_i^j v^i = c_j w^j \\ \Rightarrow w^j &= \sum_{i=1}^n A_i^j v^i \end{aligned} \tag{19.18}$$

· Fasst man ähnlich wie auf S. 86 die geordnete Basis  $B$  auf als Isomorphismus<sup>9</sup>  $\mathbb{K}_v^n \xrightarrow{\cong} V$ ,  $B((v^1, \dots, v^n)^T) = \sum b_i v^i$ , und entsprechend  $C : \mathbb{K}_w^m \xrightarrow{\cong} W$ , so gilt  $A = C^{-1} \hat{A} B$ , oder diagrammatisch

$$\begin{array}{ccc} V & \xrightarrow{\hat{A}} & W \\ B \uparrow & & \uparrow C \\ \mathbb{K}_v^n & \xrightarrow{A} & \mathbb{K}_w^m \end{array} \tag{19.19}$$

· Basen von  $V$  und  $W$  stiften damit einen Isomorphismus

$$\text{Hom}_{\mathbb{K}}(V, W) \cong \text{Mat}_{m \times n}(\mathbb{K}) \tag{19.20}$$

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

· 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

$$\tilde{b}_k = \sum_{i=1}^n b_i R_k^i \tag{19.21}$$

---

<sup>9</sup>Der allerdings bezüglich der kanonischen Basis von  $\mathbb{K}_v^n$  und  $B$  als Basis von  $V$  durch die Einheitsmatrix dargestellt wird.

Die Komponenten bezüglich der neuen Basis hängen über

$$\tilde{v}^k = \sum_{i=1}^n (R^{-1})_i^k v^i \tag{19.22}$$

mit denen der alten zusammen: Aus

$$\sum_k R_k^i (R^{-1})_j^k = \delta_j^i = \begin{cases} 1 & i = j \\ 0 & \text{sonst} \end{cases} \tag{19.23}$$

folgt

$$\sum_{k=1}^n \tilde{b}_k \tilde{v}^k = \sum_{i=1}^n b_i v^i = v \tag{19.24}$$

Diagrammatisch ist

$$\begin{array}{ccc} V & \xrightarrow{\text{id}} & V \\ \tilde{B} \uparrow & & \uparrow B \\ \mathbb{K}_v^n & \xrightarrow{R} & \mathbb{K}_v^n \end{array} \tag{19.25}$$

d.h. die ‘‘Basiswechselform’’  $R = B^{-1}\tilde{B}$  ist die Darstellung der Identität auf  $V$  bezüglich den Basen  $B$  und  $\tilde{B}$ .

· 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

$$\tilde{A} = S^{-1} \cdot A \cdot R \tag{19.26}$$

Diagrammatisch

$$\begin{array}{ccc} \mathbb{K}_v^n & \xrightarrow{\tilde{A}} & \mathbb{K}_w^m \\ \left( \begin{array}{c} \downarrow \tilde{B} \\ V \xrightarrow{\hat{A}} W \\ \uparrow B \end{array} \right) & & \left( \begin{array}{c} \downarrow \tilde{C} \\ \uparrow C \end{array} \right) \\ \mathbb{K}_v^n & \xrightarrow{A} & \mathbb{K}_w^m \end{array} \tag{19.27}$$

### Operationen mit Vektorräumen

Ein Vektorraum  $V$  mit gegebener Basis  $B$  ist also nichts anderes als der Raum der  $\mathbb{K}$ -wertigen Funktionen auf  $B$  mit endlichem Träger:

$$V \cong \mathcal{F}^{\text{fin}}(B, \mathbb{K}) \tag{19.28}$$

Dieser ‘‘operationelle’’ Standpunkt zahlt sich auf lange Zeit auch gerade in der Physik aus,<sup>10</sup> solange man sicherstellt, dass die Betrachtungen entweder gar nicht, oder aber in ‘‘natürlicher Weise’’, von der Wahl der Basis abhängen. Wir illustrieren dies bei der Zusammensetzung neuer Vektorräume aus gegebenen.

<sup>10</sup>Im 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.

§ 19. MULTI-LINEARE ALGEBRA

**Definition 19.8.** Ist  $V$  ein  $\mathbb{K}$ -Vektorraum, so heisst der  $\mathbb{K}$ -Vektorraum der linearen Abbildung von  $V$  nach  $\mathbb{K}$  (als Vektorraum) der *Dualraum* von  $V$ .

$$V^\vee := \text{Hom}_{\mathbb{K}}(V, \mathbb{K}) \quad (19.29)$$

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

$$\beta^j : V \rightarrow \mathbb{K} \quad \beta^j(b_i) := \delta_i^j \quad (19.30)$$

eine Basis von  $V^\vee$ , welche die zu  $B$  duale Basis genannt wird. Elemente von  $V^\vee$  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

$$e_i = (0, \dots, 0, 1, 0, \dots, 0, \dots) \quad (19.31)$$

↑  
i-te Stelle

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

$$\epsilon^j(e_i) = \delta_i^j \quad (19.32)$$

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  $\epsilon^j$ .

Übungsaufgabe:  $\mathcal{F}^{\text{fin}}(M, \mathbb{K})^\vee \cong \mathcal{F}(M, \mathbb{K})$

· 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  $\mathbb{K}^n \stackrel{B^\vee}{\cong} V^\vee$ , so dass die scheinbar harmlose Identifikation  $\mathbb{K}^n = \mathbb{K}^n$  einen Isomorphismus  $V \cong V^\vee$  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. (19.22))

$$\tilde{\beta}^l = \sum_{j=1}^n (R^{-1})_j^l \beta^j \quad (19.33)$$

Zusammen mit (19.26) folgt daraus: Der von  $B, B^\vee$  induzierte Isomorphismus  $V \cong V^\vee$ , der bezüglich der Basen  $B, B^\vee$  durch die Einheitsmatrix dargestellt wird, hat bezüglich der Basen  $\tilde{B}, \tilde{B}^\vee$  die darstellende Matrix  $R^T R$ . Im Allgemeinen unterscheidet er sich damit von dem Isomorphismus, welcher bezüglich der Basen  $\tilde{B}, \tilde{B}^\vee$  natürlich erscheint (weil er bezüglich dieser Basen durch die Einheitsmatrix dargestellt wird).

· Man sagt zu (19.33) auch salopp: “Kovektoren transformieren sich unter Basiswechsel kontragredient zu den Vektoren”.

**Definition 19.9.** 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

$$(v_1, w_1) + (v_2, w_2) := (v_1 + v_2, w_1 + w_2), \quad \lambda \cdot (v, w) := (\lambda \cdot v, \lambda \cdot w) \quad (19.34)$$

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

$$V \oplus W = (V \times W, +, \cdot, (0, 0)) \quad (19.35)$$

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

$$\mathcal{F}^{\text{fin}}(B \cup C, \mathbb{K}) = \mathcal{F}^{\text{fin}}(B, \mathbb{K}) \oplus \mathcal{F}^{\text{fin}}(C, \mathbb{K}) \quad (19.36)$$

und im endlich-dimensionalen Fall insbesondere

$$\dim(V \oplus W) = \dim(V) + \dim(W) \quad (19.37)$$

· 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 19.9 dann “äussere” direkte Summe. Allerdings ist der Unterschied letztlich klein, da in der Situation von 19.9  $V$  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 19.10.** Seien  $V_1, \dots, V_k$  und  $X$  Vektorräume über  $\mathbb{K}$ . Eine Abbildung

$$T : V_1 \times \dots \times V_k \rightarrow X \quad (19.38)$$

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

$$V_i \ni v \mapsto T(v_1, \dots, v_{i-1}, v, v_{i+1}, \dots, v_k) \in X \quad (19.39)$$

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

§ 19. MULTI-LINEARE ALGEBRA

**Definition 19.11.** Seien  $V$  und  $W$  Vektorräume über  $\mathbb{K}$ . Für gegebene Basen  $B$  von  $V$  und  $C$  von  $W$  gilt

$$V = \mathcal{F}^{\text{fin}}(B, \mathbb{K}) \quad W = \mathcal{F}^{\text{fin}}(C, \mathbb{K}) \quad (19.40)$$

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

$$V \otimes W = \mathcal{F}^{\text{fin}}(B \times C, \mathbb{K}) \quad (19.41)$$

*Bemerkungen.* · Vor allem Anderen vergleiche man (19.41) mit (19.36).

· Dann aber beachte man, dass 19.11 eigentlich das Tensorprodukt von Vektorräumen mit Basis definiert, und zwar wieder als Vektorraum mit Basis. Um die Unabhängigkeit von der Wahl der Basis sicherzustellen, d.h. die Definition von  $V \otimes W$  als "Vektorraum ohne ausgezeichnete Basis", ergänzt man die Konstruktion

1. um eine *bilineare Abbildung*

$$V \times W \rightarrow V \otimes W \quad (v, w) \mapsto v \otimes w \quad (19.42)$$

welche in den gegebenen Identifikationen (19.40), (19.41) definiert ist durch

$$(f \otimes g)(b, c) := f(b) \cdot g(c) \quad (19.43)$$

Übungsaufgabe: Man prüfe, dass diese Abbildung bilinear ist.

Mit Hilfe dieser bilinearen Abbildung (dem Tensorprodukt von Vektoren) lässt sich die Basis von  $V \otimes W$  schreiben als

$$\{b \otimes c \mid b \in B, c \in C\} \quad (19.44)$$

und es gilt (alles endliche Summen!)

$$\left( \sum_i \lambda_i \cdot b_i \right) \otimes \left( \sum_k \mu_k \cdot b_k \right) = \sum_{i,k} \lambda_i \mu_k \cdot b_i \otimes c_k \quad (19.45)$$

2. um eine *Vorschrift*, wie ein Basiswechsel in  $V$  und  $W$  einen Basiswechsel im Tensorprodukt induziert: Sind  $\tilde{B}$  und  $\tilde{C}$  andere Basen von  $V$  bzw.  $W$ , so identifizieren wir das zugehörige, a priori verschiedene, Tensorprodukt  $V \tilde{\otimes} W$  mittels  $\tilde{b} \tilde{\otimes} \tilde{c} \mapsto \tilde{b} \otimes \tilde{c}$  mit  $V \otimes W$ , so dass die bezüglich den beiden Basen definierten Abbildungen  $\otimes$  und  $\tilde{\otimes}$  übereinstimmen.

Beh.:  $\{\tilde{b} \otimes \tilde{c} \mid (\tilde{b}, \tilde{c}) \in \tilde{B} \times \tilde{C}\}$  ist eine Basis von  $V \otimes W$ .

Begr.: (Der Einfachheit halber, im endlich-dimensionalen Fall) Sind

$$\tilde{b}_j = \sum_{i=1}^n b_i R_j^i, \quad \tilde{c}_l = \sum_{k=1}^m c_k S_l^k \quad (19.46)$$

mit invertierbaren Matrizen  $R$  und  $S$ , so gilt gemäss (19.45)

$$\tilde{b}_j \otimes \tilde{c}_l = \sum_{i,k} b_i \otimes c_k R_j^i S_l^k \quad (19.47)$$

und die “Matrix”  $(R_j^i S_l^k)_{(i,k),(j,l) \in \{1,\dots,n\} \times \{1,\dots,m\}}$  ist invertierbar:

$$\sum_{(j,l)} (R_j^i S_l^k) ((R^{-1})_p^j (S^{-1})_q^l) = \sum_j R_j^i (R^{-1})_p^j \sum_l S_l^k (S^{-1})_q^l = \delta_p^i \delta_q^k = \delta_{(p,q)}^{(i,k)} \quad (19.48)$$

· Das Tensorprodukt zweier Vektorräume als der Vektorraum (19.41) zusammen mit der bilinearen Abbildung (19.42) ist damit so kanonisch definiert, wie es mit Hilfe von Basen geht. Bei weiterbestehenden oder wiederkehrenden Zweifeln merke man sich (19.44), im endlich-dimensionalen Fall noch (19.47) und vor allem

$$\dim(V \otimes W) = \dim(V) \cdot \dim(W) \quad (19.49)$$

im Unterschied zu (19.37).

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

$$\mathcal{F}^{\text{fin}}(V \times W, \mathbb{K})/L \quad (19.50)$$

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

$$A_1 \times A_2 : V_1 \times V_2 \rightarrow W_1 \times W_2 \rightarrow W_1 \otimes W_2 \quad (19.51)$$

eine eindeutige lineare Abbildung

$$A_1 \otimes A_2 : V_1 \otimes V_2 \rightarrow W_1 \otimes W_2 \quad (19.52)$$

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

$$(A_1 \otimes A_2)(b_1 \otimes b_2) := A_1(b_1) \otimes A_2(b_2) \quad (19.53)$$

von einer Produkt-Basis<sup>11</sup>  $\{b_1 \otimes b_2 \mid b_1 \in B_1, b_2 \in B_2\}$  auf ganz  $V_1 \otimes V_2$ .) (19.52) ist analog zur wohl schon bekannten direkten Summe von linearen Abbildungen,

$$\begin{aligned} A_1 \oplus A_2 : V_1 \oplus V_2 &\rightarrow W_1 \oplus W_2 \\ (A_1 \oplus A_2)(v_1 \oplus v_2) &:= A_1(v_1) \oplus A_2(v_2) \end{aligned} \quad (19.54)$$

· 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

$$(\lambda A_1 + B_1) \otimes A_2 = \lambda A_1 \otimes A_2 + B_1 \otimes A_2 \in \text{Hom}_{\mathbb{K}}(V_1 \otimes V_2, W_1 \otimes W_2) \quad (19.55)$$

<sup>11</sup>Dass nicht jede Basis von  $V_1 \otimes V_2$  von dieser Form ist, lernen wir hoffentlich in den Übungen.

## § 19. MULTI-LINEARE ALGEBRA

· 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

$$\begin{aligned} (V_1 \oplus V_2) \otimes V_3 &\cong (V_1 \otimes V_3) \oplus (V_2 \otimes V_3) \\ (V_1 \otimes V_2) \otimes V_3 &\cong V_1 \otimes (V_2 \otimes V_3) \end{aligned} \quad (19.56)$$

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

$$c_{V,W} : V \otimes W \ni v \otimes w \mapsto w \otimes v \in W \otimes V \quad (19.57)$$

definiert aber einen linearen Isomorphismus, welcher kanonisch ist zumindestens in dem Sinne, dass er nicht von weiteren Wahlen abhängt. Falls allerdings von vorne herein  $V = W$ , 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

$$\mathbb{K} \otimes V \cong V, \quad \mathbb{K}^0 \otimes V \cong \mathbb{K}^0, \quad \mathbb{K}^0 \oplus V \cong V \quad (19.58)$$

Wir untersuchen zuletzt noch das Zusammenspiel zwischen Tensorprodukt und Dualraum, das aber wegen der Bemerkungen unter 19.8 nur im endlich-dimensionalen Fall gut funktioniert.

**Proposition 19.12.** *Seien  $V$  und  $W$  endlich-dimensionale  $\mathbb{K}$ -Vektorräume. Dann ist die (kanonische) Abbildung*

$$\text{ev} : V^\vee \otimes W \rightarrow \text{Hom}_{\mathbb{K}}(V, W), \quad (19.59)$$

definiert als lineare Fortsetzung von

$$\lambda \otimes w \mapsto (v \mapsto \lambda(v) \cdot w) \quad (19.60)$$

ein Isomorphismus von Vektorräumen.

*Beweis.* Man prüft leicht, dass  $\text{ev}$  wohldefiniert ist. (Dazu genügt es zu zeigen, dass (19.60) linear in  $\lambda \in V^\vee$  und  $w \in W$  ist.) Um zu zeigen, dass (19.59) 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. (19.20), (19.49)). 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

$$\text{ev} \left( \sum_{j=1}^n \sum_{k=1}^m \beta^j \otimes c_k A_j^k \right) (b_i) = \sum_{k=1}^m c_k A_i^k = A(b_i) \quad \forall i = 1, \dots, n \quad (19.61)$$

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

*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

$$(V^\vee)^\vee = V \tag{19.62}$$

sowie

$$(V \otimes W)^\vee = V^\vee \otimes W^\vee \tag{19.63}$$

wobei die linke Seite definitionsgemäss auch gleich dem Raum der bilinearen Abbildungen  $V \times W \rightarrow \mathbb{K}$  ist (vgl. 19.10).

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

$$L(V_1, \dots, V_k; \mathbb{K}) = \text{Hom}_{\mathbb{K}}(V_1 \otimes \dots \otimes V_k, \mathbb{K}) = V_1^\vee \otimes \dots \otimes V_k^\vee \tag{19.64}$$

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

**Definition 19.13.** 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

$$V^{\mathbb{C}} = V \otimes \mathbb{C} \tag{19.65}$$

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

$$\lambda(v \otimes \mu) := v \otimes (\lambda\mu) \quad v \in V, \lambda, \mu \in \mathbb{C} \tag{19.66}$$

definiert ist (m.A.W. ist dies  $\text{id}_V \otimes \lambda$ ), zusammen mit der auf (19.65) 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 19.14.** Seien  $V$  und  $W$  Vektorräume über  $\mathbb{K}$ , und  $A \in \text{Hom}_{\mathbb{K}}(V, W)$ . Dann heisst die lineare Abbildung

$$\begin{aligned} A^\vee : W^\vee &\rightarrow V^\vee \\ A^\vee(\mu)(v) &:= \mu(A(v)) \end{aligned} \tag{19.67}$$

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

## § 20 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 20.1.** (i) Eine Gruppe ist eine Menge  $G$  mit einem ausgezeichneten Element  $e \in G$  und einer binären Verknüpfung

$$\cdot : G \times G \rightarrow G, \quad (g_1, g_2) \mapsto g_1 \cdot g_2 \quad (20.1)$$

mit den Rechenregeln

$$\begin{aligned} \forall g_1, g_2, g_3 \in G : (g_1 \cdot g_2) \cdot g_3 &= g_1 \cdot (g_2 \cdot g_3) \\ \forall g \in G : e \cdot g &= g \end{aligned} \quad (20.2)$$

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

$$h \cdot g = e \quad (20.3)$$

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

$$\phi(g_1 \cdot g_2) = \phi(g_1) \cdot \phi(g_2) \quad \forall g_1, g_2 \in G \quad (20.4)$$

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 20.2.** · 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  $S_n$ ,

$$S_n \ni \sigma : N \ni a \mapsto \sigma(a) \quad (20.5)$$

als zwei-zeilige Matrix schreiben:

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix} \quad (20.6)$$

In  $S_5$  gilt etwa

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 3 & 4 & 5 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 5 & 1 & 3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 5 & 2 & 3 & 1 \end{pmatrix} \quad (20.7)$$

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

$$L_g : G \ni h \mapsto g \cdot h \in G, \quad (20.8)$$

eine Permutation der  $G$  zugrundeliegenden Menge, mit Inversum  $L_{g^{-1}}$ .  $g \mapsto L_g$  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.)

· Im Unterschied hierzu ist  $GL(n, \mathbb{R})$  ein Beispiel einer *kontinuierlichen Gruppe*: Da eine linearer Endomorphismus eines endlich-dimensionalen Vektorraums genau dann invertierbar ist, wenn die Determinante ihrer darstellenden Matrix nicht verschwindet, ist

$$GL(n, \mathbb{R}) = \{A \in \text{Mat}(n, \mathbb{R}) \mid \det A \neq 0\} \quad (20.9)$$

und, da  $\det$  bezüglich einer beliebigen Norm auf  $GL(n, \mathbb{R}) \cong \mathbb{R}^{n^2}$  stetig ist, eine offene Teilmenge von  $GL(n, \mathbb{R})$ . Bezüglich dieser metrischen Struktur (s. 4.6) sind dann die Operationen in der Gruppe (Multiplikation und Inversenbildung) *stetige* Abbildungen.  $GL(n, \mathbb{R})$  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, und auch hier hilft dies nur bedingt zum Verständnis aller kontinuierlichen Gruppen.)

· Für die Physik sind kontinuierliche Gruppen letztendlich deutlich wichtiger. Endliche Gruppen sind zur Bildung einer Intuition und zur Einführung der Darstellungstheorie aber deutlich einfacher.

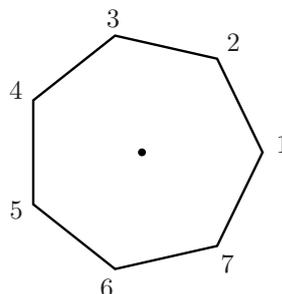
**Beispiel 20.4.** Wie bereits skizziert sind die meisten speziellen Gruppen interessant als Symmetriegruppen algebraischer Strukturen (z.B. eines inneren Produkts auf einem reellen Vektorraum) oder geometrischer Objekte.

· 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*  $\mathbb{Z}/n\mathbb{Z}$ , da sie als Untergruppe  $C_n \subset S_n$  die Zahlen  $\{1, \dots, n\}$  “im Kreis dreht”. Wir notieren sie dann multiplikativ

$$C_n = \{r_k \mid k \in \mathbb{Z}\} / (r_{n+k} \sim r_k) \text{ mit } r_k \cdot r_l = r_{k+l}. \quad (20.10)$$

· Als nicht-triviales Paradebeispiel 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 Matrizen<sup>12</sup>

$$r_k = \begin{pmatrix} \cos \frac{2\pi k}{n} & -\sin \frac{2\pi k}{n} \\ \sin \frac{2\pi k}{n} & \cos \frac{2\pi k}{n} \end{pmatrix}, \quad s_k = \begin{pmatrix} \cos \frac{2\pi k}{n} & \sin \frac{2\pi k}{n} \\ \sin \frac{2\pi k}{n} & -\cos \frac{2\pi k}{n} \end{pmatrix} \quad (20.11)$$

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

$$s_k = r_k \cdot s_0, \quad s_0 r_k s_0 = r_{-k} = (r_k)^{-1}, \quad s_k \cdot s_k = r_0 = e \quad (20.12)$$

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 (19.15) der Quotient einer Gruppe nach einer Untergruppe im Allgemeinen keine natürliche Gruppenstruktur: Zwar ist für eine Untergruppe  $H \subset G$

$$g \sim g' :\Leftrightarrow g^{-1} \cdot g' \in H \Leftrightarrow \exists h \in H : g' = g \cdot h \quad (20.13)$$

<sup>12</sup>Für 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.

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

$$G/H := \{gH := \{g \cdot h \mid h \in H\} \mid g \in G\} \quad (20.14)$$

eine Multiplikation zu definieren,

$$(g_1H) \cdot (g_2H) \stackrel{?}{=} (g_1 \cdot g_2)H \quad (20.15)$$

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  $g'_1H = g_1H$ , 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  $g'_1g_2 = g_1g_2h'$ , oder

$$(g_1g_2)^{-1}(g'_1g_2) = g_2^{-1}g_1^{-1}g_1hg_2 = g_2^{-1} \cdot h \cdot g_2 = h' \quad (20.16)$$

**Definition 20.5.** Eine Untergruppe mit dieser Eigenschaft, das heisst

$$g^{-1} \cdot H \cdot g \subset H \quad \forall g \in G \quad (20.17)$$

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

**Proposition 20.6.** Ist  $H \subset G$  eine normale Untergruppe, so ist auf dem (mengen-theoretischen) Quotienten  $G/H = \{g \cdot H \mid g \in G\}$  die Verknüpfung

$$[g_1] \cdot [g_2] = [g_1 \cdot g_2] \quad (20.18)$$

wohldefiniert und erfüllt die Axiome aus Def. 20.1 (i).

· Beispielsweise ist in  $D_n$  die zyklische Untergruppe  $C_n$  der Drehungen eine normale Untergruppe (vgl. insbes. die zweite Relation in (20.12)). Andererseits ist die von  $s_0$  erzeugte Untergruppe  $\{e, s_0\} \cong \mathbb{Z}/2\mathbb{Z}$  kein Normalteiler:

$$r_1^{-1}s_0r_1 = r_1^{-1}r_1^{-1}s_0 = r_2^{-1}s_0 \neq s_0 \quad (20.19)$$

(für  $n > 2$ ).

Zur Vorbereitung der nächsten Definition bemerken wir, dass man die erste Klasse von Beispielen in 20.2 in die zweite einbetten kann, indem man die Menge  $M$  im Sinne von (19.6) (bzw. (19.12)) zu einer Basis eines Vektorraums erklärt. Konkret ist dies für  $S_n = \text{Bij}(\{1, 2, \dots, n\})$  der injektive Gruppenhomomorphismus

$$\begin{aligned} \phi : S_n &\rightarrow GL(n, \mathbb{R}) \\ \phi(\sigma)(e_i) &:= e_{\sigma(i)} \end{aligned} \quad (20.20)$$

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 (20.11).

Jede endliche Gruppe wird also durch Verkettung von (20.8) und (20.20) 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.

## § 20. GRUPPEN UND IHRE DARSTELLUNGEN

**Definition 20.7.** Eine (*lineare*) *Darstellung* einer Gruppe  $G$  ist ein Vektorraum  $V$  zusammen mit einem Gruppenhomomorphismus

$$\rho : G \rightarrow GL(V) \quad (20.21)$$

· In dieser Definition wird nicht angenommen, dass  $G$  endlich ist oder  $\rho$  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  $(V_1, \rho_1)$  und  $(V_2, \rho_2)$  ist eine  $\mathbb{K}$ -lineare Abbildung  $\phi : V_1 \rightarrow V_2$  von Vektorräumen derart, dass für alle  $g \in G$

$$\phi \circ \rho_1(g) = \rho_2(g) \circ \phi \in \text{Hom}_{\mathbb{K}}(V_1, V_2) \quad (20.22)$$

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

·  $(V_1, \rho_1)$  und  $(V_2, \rho_2)$  heißen isomorph (oder äquivalent) falls eine invertierbare  $G$ -lineare Abbildung  $V_1 \xrightarrow{\cong} V_2$  existiert.

· Eine *Unterdarstellung* einer Darstellung  $(V, \rho)$  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

$$V^G := \{v \in V \mid \rho(g)(v) = v \forall g \in G\} \quad (20.23)$$

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

$$g \mapsto (\rho_1 \oplus \rho_2)(g) := \rho_1(g) \oplus \rho_2(g) \quad (20.24)$$

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

$$g \mapsto (\rho_1 \otimes \rho_2)(g) := \rho_1(g) \otimes \rho_2(g) \quad (20.25)$$

· 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  $(V, \rho)$  von  $G$  sind der triviale Vektorraum  $\{0\}$  und das gesamte  $V$  invariante Unterräume.

**Definition 20.8.** Eine Darstellung  $(V, \rho)$  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:  $\phi$  aus (20.20) 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 20.9.** *Seien  $(V_1, \rho_1)$  und  $(V_2, \rho_2)$  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  $(V_1, \rho_1)$  und  $\text{Im}(\phi) := \{\phi(v) \mid v \in V_1\} \subset V_2$  ist eine Unterdarstellung von  $(V_2, \rho_2)$ . Ausserdem induziert  $\rho_2$  eine Darstellung auf  $\text{Coker}(\phi) = V_2 / \text{Im}(\phi)$ .

(ii) Falls  $(V_1, \rho_1)$  und  $(V_2, \rho_2)$  irreduzibel sind, so ist  $\phi$  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  $(V_1, \rho_1)$  und  $(V_2, \rho_2)$  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  $\phi$  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  $\phi$  surjektiv, mithin invertierbar.

(iii) Wegen (ii) und  $\phi \neq 0$  ist  $\phi$  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 7.7 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

$$\text{Hom}_G(V_1, V_2) = (V_1^* \otimes V_2)^G = \begin{cases} \mathbb{C} & \text{falls } (V_1, \rho_1) \text{ isomorph zu } (V_2, \rho_2) \\ \{0\} & \text{sonst} \end{cases} \quad (20.26)$$

Hier haben wir in der ersten Gleichung zunächst 19.12 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. (19.14))  $V_1^*$  zu einer Darstellung aufgerüstet. Man prüft dann, dass unter dieser Identifikation  $\text{Hom}_G(V_1, V_2)$  (vgl. (20.22)) mit dem Raum der invarianten Vektoren (vgl. (20.23)) im Tensorprodukt übereinstimmt.

**Theorem 20.10.** *Sei  $G$  abelsch und  $(V, \rho)$  eine endlich-dimensionale irreduzible Darstellung von  $G$  über  $\mathbb{C}$ . Dann ist  $V$  ein-dimensional, d.h.  $\rho \in \text{Hom}(G, \mathbb{C}^\times)$ .*

## § 20. GRUPPEN UND IHRE DARSTELLUNGEN

*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  $g \in G$

$$\rho(g_0)\rho(g) = \rho(g_0g) = (G \text{ abelsch}) = \rho(gg_0) = \rho(g)\rho(g_0) \quad (20.27)$$

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  $(V, \rho)$  irreduzibel ist, muss  $V$  ein-dimensional sein.  $\square$

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

**Beispiel 20.11.** Da  $C_n$  von  $r_1$  erzeugt wird (vgl. (20.10)), ist eine endlich-dimensionale irreduzible Darstellung über  $\mathbb{C}$ , welche nach 20.10 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  $(\mathbb{C}, \rho_l)$  mit

$$\rho_l(r_k) = (\rho_l(r_1))^k = \exp\left(\frac{2\pi i l \cdot k}{n}\right) \quad (20.28)$$

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

**Beispiel 20.12.** Wir verweisen auf die Präsentation der Gruppe in (20.12).

1. Ist  $(V, \rho)$  eine irreduzible Darstellung von  $D_n$  über  $\mathbb{C}$ , dann erhalten wir durch Einschränkung von  $\rho$  (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 20.11 existiert also ein  $l \in \{0, \dots, n-1\}$  und ein ein-dimensionaler Unterraum  $V_l \subset V$  mit

$$\rho(r_k)|_{V_l} = \exp\left(\frac{2\pi i l \cdot k}{n}\right) \text{id}_{V_l}. \quad (20.29)$$

$\rho(s_0)(V_l)$  ist wieder ein ein-dimensionaler Unterraum, der wegen (20.12), speziell wegen  $r_k s_0 = s_0 r_{-k}$ , invariant ist mit

$$\rho(r_k)|_{\rho(s_0)(V_l)} = \rho(s_0)\rho(r_{-k})|_{V_l} = \exp\left(\frac{-2\pi i l \cdot k}{n}\right) \text{id}_{\rho(s_0)(V_l)} \quad (20.30)$$

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

$$\rho(r_k) = \begin{pmatrix} \exp\left(\frac{2\pi i k \cdot l}{n}\right) & 0 \\ 0 & \exp\left(\frac{-2\pi i k \cdot l}{n}\right) \end{pmatrix}, \quad \rho(s_0) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad (20.31)$$

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.<sup>13</sup>

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. 19.10 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. (19.64))

$$L(V, \dots, V; \mathbb{K}) = \underbrace{(V \otimes \dots \otimes V)}_{k \text{ Mal}}^\vee = \underbrace{V^\vee \otimes \dots \otimes V^\vee}_{k \text{ Mal}} =: (V^\vee)^{\otimes k} \quad (20.32)$$

was sich mit (19.58) 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

$$S_k \ni \sigma \mapsto c_\sigma(\lambda_1 \otimes \lambda_2 \otimes \dots \otimes \lambda_k) := \lambda_{\sigma(1)} \otimes \lambda_{\sigma(2)} \dots \otimes \lambda_{\sigma(k)} \quad (20.33)$$

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

$$(c_\sigma(T))(v_1, \dots, v_k) := T(v_{\sigma(1)}, \dots, v_{\sigma(k)}) \quad (20.34)$$

Beispielsweise wird für  $k = 2$  das nicht-triviale Element von  $S_2 = \mathbb{Z}/2\mathbb{Z}$  durch die Abbildung  $c_{V^\vee, V^\vee}$  aus (19.57) dargestellt.

· Die Darstellung (20.33) ist im Allgemeinen nicht irreduzibel. Sie enthält als Unterdarstellung den Raum der invarianten oder symmetrischen Tensoren, welchen wir mit Hilfe von (20.32) schreiben können als

$$\text{Sym}^k V^\vee := \left\{ T \in L(V, \dots, V; \mathbb{K}) \mid T(v_1, \dots, v_k) = T(v_{\sigma(1)}, \dots, v_{\sigma(k)}) \forall \sigma \in S_k \right\} \quad (20.35)$$

<sup>13</sup>Bezüglich der obigen Basen stellt  $\phi := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$  den Isomorphismus zwischen  $V_l$  und  $V_{n-l}$  dar.

## § 20. GRUPPEN UND IHRE DARSTELLUNGEN

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

$$\wedge^k V^\vee := \{T \in L(V, \dots, V; \mathbb{K}) \mid T = \text{sgn}(\sigma)c_\sigma(T) \forall \sigma \in S_k\} \quad (20.36)$$

wobei  $\text{sgn} : S_k \rightarrow \{\pm 1\}$  den Signum- oder Paritätshomomorphismus in die multiplikative Gruppe  $\{\pm 1\}$  bezeichnet.  $((V^\vee)^{\otimes k}, c)$  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 (20.20) definierten Darstellung von  $S_n$ .

- Elemente  $\omega \in \wedge^k V^\vee$  heissen meistens auch einfach  $k$ -Formen auf  $V$ .
- Wir definieren dann für  $\lambda_1, \dots, \lambda_k \in V^\vee$  das *Dachprodukt*  $\lambda_1 \wedge \dots \wedge \lambda_k \in \wedge^k V^\vee$  durch

$$\lambda_1 \wedge \dots \wedge \lambda_k := \sum_{\sigma \in S_k} \text{sgn}(\sigma)c_\sigma(\lambda_1 \otimes \dots \otimes \lambda_k) \quad (20.37)$$

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

$$\begin{aligned} (\lambda_1 \wedge \dots \wedge \lambda_k)(v_1, \dots, v_k) &:= \sum_{\sigma \in S_k} \text{sgn}(\sigma)\lambda_1(v_{\sigma(1)}) \cdot \lambda_2(v_{\sigma(2)}) \cdot \dots \cdot \lambda_k(v_{\sigma(k)}) \\ &= \det(\lambda_i(v_j))_{\substack{i=1, \dots, k \\ j=1, \dots, k}} \end{aligned} \quad (20.38)$$

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  $V^\vee$ , so ist

$$\{\epsilon^{i_1} \wedge \dots \wedge \epsilon^{i_k} \mid 1 \leq i_1 < \dots < i_k \leq n\} \quad (20.39)$$

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

$$\dim \wedge^k V^\vee = \binom{n}{k} \quad (20.40)$$

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

- Wegen (20.39) können wir dann für beliebige  $k, l \geq 1$  das Dachprodukt

$$\begin{aligned} \cdot \wedge \cdot : \wedge^k V^* \times \wedge^l V^\vee &\rightarrow \wedge^{k+l} V^\vee \\ (\omega, \alpha) &\mapsto \omega \wedge \alpha \end{aligned} \quad (20.41)$$

definieren durch die Formel

$$(\lambda_1 \wedge \dots \wedge \lambda_k) \wedge (\mu_1 \wedge \dots \wedge \mu_l) := \lambda_1 \wedge \dots \wedge \lambda_k \wedge \mu_1 \wedge \dots \wedge \mu_l \quad (20.42)$$

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

$$\omega \wedge \alpha(v_1, \dots, v_{k+l}) = \frac{1}{k!l!} \sum_{\sigma \in S_{k+l}} \text{sgn}(\sigma)\omega(v_{\sigma(1)}, \dots, v_{\sigma(k)}) \cdot \alpha(v_{\sigma(k+1)}, \dots, v_{\sigma(k+l)}) \quad (20.43)$$

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

$$\begin{aligned} (\omega_1 \wedge \omega_2) \wedge \omega_3 &= \omega_1 \wedge (\omega_2 \wedge \omega_3) \\ \omega_1 \wedge \omega_2 &= (-1)^{k_1 k_2} \omega_2 \wedge \omega_1 \end{aligned} \quad (20.44)$$

Dies bedeutet: Der Raum

$$\wedge V^\vee = \bigoplus_{k=0}^n \wedge^k V^\vee, \quad (20.45)$$

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

$$\dim \wedge V^\vee = \sum_{k=0}^n \binom{n}{k} = 2^n \quad (20.46)$$

Geometrische Interpretation: Für einen  $n$ -dimensionalen reellen Vektorraum  $V$  ist wegen (20.40)  $\wedge^n V^\vee$  ein-dimensional. Man nennt ein nicht-verschwindendes Element  $\omega \in \wedge^n V^\vee$ ,  $\omega \neq 0$  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  $\alpha \in \wedge^n W^\vee$ , und  $A : W \rightarrow V$  eine lineare Abbildung, so wird durch

$$A^\vee(\omega)(c_1, \dots, c_n) := \omega(Ac_1, \dots, Ac_n) \quad (20.47)$$

eine alternierende  $n$ -Form  $A^\vee(\omega)$  definiert. (Übungsaufgabe). Wegen  $\dim \wedge^n W^\vee = 1$  gilt  $A^\vee(\omega) = d \cdot \alpha$  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

$$\omega = \epsilon^1 \wedge \epsilon^2 + 2\epsilon^2 \wedge \epsilon^3 \quad (20.48)$$

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

$$\begin{aligned} \omega(e_1 + e_2 + e_3, 2e_3) &= (\epsilon^1 \wedge \epsilon^2)(e_1 + e_2 + e_3, 2e_3) + 2(\epsilon^2 \wedge \epsilon^3)(e_1 + e_2 + e_3, 2e_3) \\ &= (\epsilon^1 \otimes \epsilon^2 - \epsilon^2 \otimes \epsilon^1)(e_1 + e_2 + e_3, 2e_3) \\ &\quad + 2(\epsilon^2 \otimes \epsilon^3 - \epsilon^3 \otimes \epsilon^2)(e_1 + e_2 + e_3, 2e_3) \\ &= 0 + 2\epsilon^2(e_2) \cdot \epsilon^3(2e_3) = 4 \end{aligned} \quad (20.49)$$

## § 21 Euklidische und unitäre Räume

In den Beispielen 20.11 und 20.12 leuchtete 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  $\rho$  verträglichen inneren Produkt ausgerüstet ist. Solche inneren Produkte und die daraus abgeleiteten Normen (s. §4 und §10) sind ganz allgemein das Werkzeug, um Objekte aus Vektorräumen mit physikalischen Messgrößen in Verbindung zu setzen. Wir wiederholen zunächst die wichtigsten Aussagen aus der linearen Algebra (siehe S. 85 ff.), unter Einbeziehung ihrer komplexen Analoga.

**Definition 21.1.** 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

$$\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{C} \quad (21.1)$$

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:

$$\langle \lambda v_1 + v_2, \mu w_1 + w_2 \rangle = \bar{\lambda} \mu \langle v_1, w_1 \rangle + \bar{\lambda} \langle v_1, w_2 \rangle + \mu \langle v_2, w_1 \rangle + \langle v_2, w_2 \rangle \quad (21.2)$$

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

$$\langle w, v \rangle = \overline{\langle v, w \rangle} \quad (21.3)$$

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

$$\langle v, v \rangle > 0 \quad \text{für } v \neq 0 \quad (21.4)$$

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

$$V \mapsto V^\vee \quad v \mapsto v^\flat : w \mapsto \langle v, w \rangle \quad (21.5)$$

injektiv ist. Für  $\mathbb{K} = \mathbb{R}$  ist diese Abbildung linear, für  $\mathbb{K} = \mathbb{C}$  konjugiert linear: Da Skalare beim Herausziehen aus dem ersten Argument von  $\langle \cdot, \cdot \rangle$  komplex konjugiert werden, gilt

$$(v_1 + \lambda v_2)^\flat = v_1^\flat + \bar{\lambda} v_2^\flat \quad (21.6)$$

Dennoch folgt wie auch für  $\mathbb{K} = \mathbb{R}$  im Falle  $\dim(V) < \infty$  bereits aus der Injektivität, dass (21.5) ein Isomorphismus ist, der allerdings für  $\mathbb{K} = \mathbb{C}$  nur konjugiert linear ist.

*Beispiel.* Auf  $V = \mathbb{R}^n$  ist

$$\langle x, y \rangle := \sum_{i=1}^n x^i y^i \quad (21.7)$$

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

$$\langle z, w \rangle := \sum_{i=1}^n \bar{z}^i w^i \quad (21.8)$$

ein hermitesches inneres Produkt.

**Proposition 21.2.** *Jedes euklidische bzw. hermitesche innere Produkt auf einem endlich-dimensionalen Vektorraum ist äquivalent zu den obigen Standardformen, d.h. es existiert jeweils eine Orthonormalbasis (ONB)  $B : \mathbb{K}^n \cong V$ , bezüglich der  $\langle \cdot, \cdot \rangle$  die Form (21.7) bzw. (21.8) hat.*

*Beweis.* Gram-Schmidt Verfahren, s. auch Beweis von 21.4. □

### Selbstadjungiertheit

**Definition 21.3.** 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* falls

$$\langle v, Aw \rangle = \langle Av, w \rangle \quad \forall v, w \in V \quad (21.9)$$

*Bemerkungen.* Für 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 (21.5), dass ein eindeutiger Vektor  $A^{\text{ad}}v \in V$  existiert, mit dem

$$\langle v, Aw \rangle = \langle A^{\text{ad}}v, w \rangle \quad \forall w \in V \quad (21.10)$$

Die Zuordnung  $v \mapsto A^{\text{ad}}v$  ist  $\mathbb{K}$ -linear (auch für  $\mathbb{K} = \mathbb{C}$ !) und eindeutig durch die Eigenschaft (21.10) bestimmt. Sie heisst die zu  $A$  adjungierte Abbildung. Die Bedingung (21.9) 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  $A^{\text{ad}} = A^\dagger$ . (Transposition + komplexe Konjugation):

$$\begin{aligned} \langle v, Aw \rangle &= \sum_i \bar{v}^i A_j^i w^j \quad (\text{wo bei } A_j^i \text{ der Zeilenindex } i \text{ und der Spaltenindex } j \text{ ist}) \\ &= \sum_i \overline{A_j^i v^i} w^j \\ &= \sum_j \overline{A_i^\dagger{}^j v^i} w^j \quad (\text{wo bei } A_i^\dagger{}^j \text{ der Zeilenindex } j \text{ und der Spaltenindex } i \text{ ist}) \\ &= \langle A^\dagger v, w \rangle \end{aligned} \quad (21.11)$$

## § 21. EUKLIDISCHE UND UNITÄRE RÄUME

**Theorem 21.4.** *Es sei nun  $V$  ein endlich-dimensionaler euklidischer oder unitärer Raum, 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  $A$ , und die Eigenwerte sind alle reelle. Mit Zahlen  $\lambda_i \in \mathbb{R} \subset \mathbb{K}$  gilt also*

$$A(b_i) = \lambda_i \cdot b_i \quad \forall i \quad \text{und} \quad \langle b_i, b_j \rangle = \begin{cases} 1 & \text{für } i = j \\ 0 & \text{sonst} \end{cases} \quad (21.12)$$

*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 20.9 aus dem Fundamentalsatz der Algebra 7.7 zunächst die Existenz eines komplexen Eigenwertes  $\lambda \in \mathbb{C}$ . Mit einem Eigenvektor  $v \neq 0$  zum Eigenwert  $\lambda$  folgt aber aus

$$\bar{\lambda} \langle v, v \rangle = \langle Av, v \rangle = \langle v, Av \rangle = \lambda \langle v, v \rangle \quad (21.13)$$

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. 19.13) 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

$$\hat{v}^{\perp} := \{w \in V \mid \langle \hat{v}, w \rangle = 0\} \quad (21.14)$$

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

$$w \in \hat{v}^{\perp} \Rightarrow \langle \hat{v}, w \rangle = 0 \Rightarrow \langle \hat{v}, Aw \rangle = \langle A\hat{v}, w \rangle = \lambda \langle \hat{v}, w \rangle = 0 \Rightarrow Aw \in \hat{v}^{\perp} \quad (21.15)$$

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.  $\square$

### Unitarität

**Definition 21.5.** Sei  $(V, \langle \cdot, \cdot \rangle)$  ein euklidischer oder unitärer Raum. Eine lineare Abbildung  $U \in \text{Hom}_{\mathbb{K}}(V, V)$  heisst orthogonal/unitär, falls  $U^{\text{ad}} = U^{-1}$ , d.h.  $\langle Uv, Uw \rangle = \langle v, w \rangle \forall v, w \in V$ .

**Theorem 21.6.** *Ist  $\mathbb{K} = \mathbb{C}$  und  $U \in \text{Hom}_{\mathbb{C}}(V, V)$  unitär, so existiert eine ONB  $(b_1, \dots, b_n)$  von  $V$  aus Eigenvektoren von  $U$ , und die Eigenwerte sind alle komplexe Zahlen mit Absolutbetrag 1.*

*Beweis.* Gleiche Ideen wie bei 21.4, wobei die Aussage zu den Eigenwerten folgt aus der Rechnung

$$0 \neq \langle v, v \rangle = \langle Uv, Uv \rangle = \langle \lambda v, \lambda v \rangle = \bar{\lambda} \lambda \langle v, v \rangle = |\lambda|^2 \langle v, v \rangle \quad (21.16)$$

$\square$

**Definition 21.7.** Eine komplexe Darstellung  $(V, \rho)$  von  $G$  heisst *unitär*, falls auf  $V$  ein hermitesches inneres Produkt  $\langle \cdot, \cdot \rangle$  existiert, welches  $G$ -invariant ist, d.h.

$$\langle \rho(g)v, \rho(g)w \rangle = \langle v, w \rangle \quad \forall v, w \in V, \forall g \in G \quad (21.17)$$

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

**Proposition 21.8.** Sei  $(V, \rho)$  eine endlich-dimensionale unitäre Darstellung von  $G$ , und  $U \subset V$  ein invarianter Unterraum. Dann existiert ein komplementärer invarianter Unterraum  $W \subset V$  d.h. als Darstellungen von  $G$

$$V = U \oplus W \quad (21.18)$$

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

$$\langle u, \rho(g)w \rangle = \langle \rho(g)^{-1}u, w \rangle = 0 \quad (21.19)$$

(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.

**Proposition 21.9.** Sei jetzt  $G$  eine endliche Gruppe und  $(V, \rho)$  eine komplexe Darstellung von  $G$ . Dann ist  $(V, \rho)$  unitär.

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

$$\langle v, w \rangle := \sum_{h \in G} (\rho(h)v, \rho(h)w) \quad (21.20)$$

ein  $G$ -invariantes hermitesches inneres Produkt:

$$\langle \rho(g)v, \rho(g)w \rangle = \sum_{h \in G} (\rho(hg)v, \rho(hg)w) = \sum_{hg=h' \in G} (\rho(h')v, \rho(h')w) = \langle v, w \rangle \quad (21.21)$$

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 21.10.** 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

$$V \cong \bigoplus_{i=1}^k V_i \otimes \mathbb{C}^{a_i} = \bigoplus_{i=1}^k V_i^{\oplus a_i} \quad (21.22)$$

und diese Zerlegung ist bis auf Reihenfolge und Isomorphismus eindeutig.

## § 21. EUKLIDISCHE UND UNITÄRE RÄUME

*Beweis.* Die Existenz folgt durch wiederholtes Anwenden aus Prop. 21.8. Für die Eindeutigkeit sei

$$V = \bigoplus_{j=1}^l W_j \otimes \mathbb{C}^{b_j} \quad (21.23)$$

eine weitere solche Zerlegung. Da  $\text{id}_V|_{V_i} \rightarrow V$  ein  $G$ -Morphismus ist, muss nach dem Lemma von Schur 20.9 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.

$$(\mathbb{R}, +) \ni a \mapsto \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \in GL(2, \mathbb{R}) \quad (21.24)$$

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. 163)

Zur Erinnerung: Wir hatten im §5 die Exponentialfunktion  $\exp$  durch die Potenzreihe (5.50) 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 4.27. Hierbei ist die Operatornorm eines solchen Endomorphismus definiert als (4.48)

$$\|A\| := \sup\{\|Av\| \mid \|v\|_V = 1\} \quad (21.25)$$

· Seinerseits war das motivierende Beispiel für einen normierten Vektorraum der  $\mathbb{R}^n$  mit seiner vom euklidischen Standardprodukt (21.7) 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 (21.8) eine Norm auf  $\mathbb{C}^n$ , und folglich auf  $\text{Hom}_{\mathbb{C}}(\mathbb{C}^n, \mathbb{C}^n)$ : Die Argumente auf S. 29 bzw. 40, insbesondere die für die Gültigkeit der Cauchy-Schwarz-Ungleichung (C.S.U.),

$$|\langle v, w \rangle| \leq \|v\| \cdot \|w\| \quad \forall v, w \in V \quad (21.26)$$

lassen sich leicht für einen allgemeinen unitären Raum 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. 19.13) (21.8) identisch zu (21.7) 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 (21.7), bzw. für  $\mathbb{C}^n$  mit (21.8).

· 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.<sup>14</sup> Gemäss (15.18) ist das Differential bei  $0 \in \text{Mat}(n, \mathbb{K})$  die Identität, d.h. ausgedrückt durch die Richtungsableitungen (vgl. 10.9) gilt

$$D \exp(0)(A) = \left. \frac{d}{dt} \right|_{t=0} \exp(tA) = A \quad \forall A \in \text{Mat}(n, \mathbb{K}) \quad (21.27)$$

· Die hierfür und auch sonst wichtigste charakteristische Eigenschaft der Exponentialfunktion ist die “Funktionalgleichung” genannte Identität (s. 5.22)

$$\exp(A) \cdot \exp(B) = \exp(A + B) \quad \forall A, B \in \text{Mat}(n, \mathbb{K}) \mid AB = BA \quad (21.28)$$

· Aus dieser Gleichung folgt nämlich (vgl. (5.60)), dass für alle  $A \in \text{Mat}(n, \mathbb{K})$

- (i)  $\exp(A)$  invertierbar ist, und
- (ii) für alle  $t, s \in \mathbb{R}$

$$\exp(tA) \exp(sA) = \exp((t + s)A) \quad (21.29)$$

Man sagt in der Sprache der Gruppentheorie: Die Abbildung

$$\mathbb{R} \ni t \mapsto \exp(tA) \in GL(n, \mathbb{K}) \quad (21.30)$$

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 (20.12)).

· Denn wie wir schon im Zusammenhang des Umkehrsatzes im 13 (s. um (13.16) 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  $GL(n, \mathbb{K})$ .<sup>15</sup>

· 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.

<sup>14</sup>Im 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.

<sup>15</sup>Tatsächlich erzeugt dann eine offene Umgebung von  $\text{id}$  (evtl. zusammen mit einer endlichen Menge)  $GL(n, \mathbb{K})$  algebraisch.

§ 21. EUKLIDISCHE UND UNITÄRE RÄUME

**Definition 21.11.** Für  $\mathbb{K} = \mathbb{R}$  heisst die Gruppe

$$\begin{aligned} O(n) &:= \{R \in GL(n, \mathbb{R}) \mid \langle Rv, Rw \rangle = \langle v, w \rangle \forall v, w \in \mathbb{R}^n\} \\ &= \{R \in GL(n, \mathbb{R}) \mid R^T = R^{-1}\} \end{aligned} \quad (21.31)$$

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

$$\mathfrak{so}(n) := \{A \in \text{Mat}(n, \mathbb{R}) \mid \exp(tA) \in O(n) \forall t \in \mathbb{R}\} \quad (21.32)$$

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

$$\begin{aligned} U(n) &:= \{U \in GL(n, \mathbb{C}) \mid \langle Uv, Uw \rangle = \langle v, w \rangle \forall v, w \in \mathbb{C}^n\} \\ &= \{U \in GL(n, \mathbb{C}) \mid U^\dagger = U^{-1}\} \end{aligned} \quad (21.33)$$

die *unitäre Gruppe*. (Zur Notation siehe Fussnote 21.) Wir schreiben für die Menge ihrer infinitesimal Erzeugenden

$$\mathfrak{u}(n) := \{A \in \text{Mat}(n, \mathbb{C}) \mid \exp(tA) \in U(n) \forall t \in \mathbb{R}\} \quad (21.34)$$

**Proposition 21.12.** (i) Für  $\mathbb{K} = \mathbb{R}$  gilt:

$$\mathfrak{so}(n) = \{A \in \text{Mat}(n, \mathbb{R}) \mid A^T = -A\} \quad (21.35)$$

Für  $\mathbb{K} = \mathbb{C}$  gilt:

$$\mathfrak{u}(n) = \{A \in \text{Mat}(n, \mathbb{C}) \mid A^\dagger = -A\} \quad (21.36)$$

(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

$$[A_1, A_2] := A_1 \cdot A_2 - A_2 \cdot A_1 \in \mathfrak{so}(n) \text{ (bzw. } \mathfrak{u}(n)) \quad (21.37)$$

*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  $\forall A \in \text{Mat}(n, \mathbb{C})$ :

$$\begin{aligned} (\exp(A))^\dagger &= \left( \lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{A^k}{k!} \right)^\dagger = \lim_{n \rightarrow \infty} \left( \sum_{k=0}^n \frac{A^k}{k!} \right)^\dagger = \lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{(A^\dagger)^k}{k!} \\ &= \exp(A^\dagger) \end{aligned} \quad (21.38)$$

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

· “ $\Rightarrow$ ”: Gilt

$$\exp(tA^\dagger) = \exp(tA)^\dagger = \exp(tA)^{-1} = \exp(-tA) \quad \forall t \in \mathbb{R}, \quad (21.39)$$

so folgt durch Ableiten

$$\left. \frac{d}{dt} \right|_{t=0} \exp(tA^\dagger) = \left. \frac{d}{dt} \right|_{t=0} \exp(-tA) \Rightarrow A^\dagger = -A \quad (21.40)$$

· “ $\Leftarrow$ ”: Gilt umgekehrt  $A^\dagger = -A$ , so folgt mit Hilfe von (21.38)

$$(\exp(tA))^\dagger = \exp(tA^\dagger) = \exp(-tA) = (\exp(tA))^{-1} \quad \forall t \in \mathbb{R} \quad (21.41)$$

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

*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 21.4 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 (21.37) ist anti-symmetrisch

$$[A_1, A_2] = -[A_2, A_1] \tag{21.42}$$

und erfüllt die Jacobi-Identität

$$[A_1, [A_2, A_3]] + [A_2, [A_3, A_1]] + [A_3, [A_1, A_2]] = 0 \tag{21.43}$$

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 (21.7).

**Theorem 21.13.** *Für eine symmetrische Matrix  $A \in \text{Mat}(n, \mathbb{R})$ ,  $A^T = A$  ist*

$$\lambda := \min\{\langle x, Ax \rangle \mid \langle x, x \rangle = 1\} \tag{21.44}$$

*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  $\lambda$ .*

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

$$S^{n-1} := \left\{ x \in \mathbb{R}^n \mid \langle x, x \rangle = \sum_{i=1}^n (x^i)^2 = 1 \right\} \tag{21.45}$$

ist beschränkt und abgeschlossen, und daher nach dem Satz von Bolzano-Weierstrass 4.25 folgenkompakt im Sinne von 7.6. 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 §13 zu wiederholen. Die folgenden Betrachtungen hängen zunächst nicht von der euklidischen Metrik auf  $\mathbb{R}^n$  ab.

## § 21. EUKLIDISCHE UND UNITÄRE RÄUME

**Fakten 21.14.** 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

$$\varphi : U \cap M \rightarrow \mathbb{R}^k \quad (21.46)$$

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

(i)  $\varphi$  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 oft<sup>16</sup>) 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  $U = \mathbb{R}^n$ ,  $\varphi(x, f(x)) = x$ .)

· Der Inhalt des Satzes 13.4 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 (??) 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:

$$T_p M := D\psi(q)(\mathbb{R}^k) = \{D\psi(q)(w) \mid w \in \mathbb{R}^k\} \quad (21.47)$$

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

$$v = D\psi(q)(w) \Rightarrow DF(p)(v) = DF \circ D\psi(q)(w) = 0 \quad (21.48)$$

und aus Dimensionsgründen dann

$$T_p M = \text{Ker}(DF(p)) \subset \mathbb{R}^n \quad (21.49)$$

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.

<sup>16</sup>Es macht auch Sinn, Untermannigfaltigkeiten zu betrachten, bei denen  $\varphi^{-1}$  nur stetig, oder ein-, zwei- etc. Mal stetig differenzierbar ist. Häufig ist wie hier bei uns “glatt” synonym mit “unendlich oft differenzierbar”.

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

$$\mathbb{R}^n \ni v \mapsto D(\langle x, x \rangle - 1)(x_0)(v) = 2\langle x_0, v \rangle \quad (21.50)$$

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

$$\begin{aligned} T_{x_0}S^{n-1} &= \text{Ker } D(\langle x, x \rangle - 1)(x_0) = \{v \in \mathbb{R}^n \mid \langle x_0, v \rangle = 0\} \\ &= x_0^\perp \end{aligned} \quad (21.51)$$

(vgl. (21.14)).

**Proposition 21.15.** *Es sei  $M \subset \mathbb{R}^n$  eine  $k$ -dimensionale Untermannigfaltigkeit, und  $f \in C^\infty(\mathbb{R}^n, \mathbb{R})$ . Angenommen,  $f$  hat in  $p_* \in M$  ein lokales Minimum auf  $M$ , d.h. für eine Umgebung  $U \subset \mathbb{R}^n$  von  $p_*$  gilt<sup>17</sup>*

$$f(p) \geq f(p_*) \quad \forall p \in U \cap M \quad (21.52)$$

Dann gilt:

$$T_{p_*}M \subset \text{Ker}(Df(p_*)) \quad (21.53)$$

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 (21.52). Dann hat  $f \circ \gamma : (-\epsilon, \epsilon) \rightarrow \mathbb{R}$  in 0 ein lokales Minimum und es folgt daher aus 9.7

$$0 = \left. \frac{d}{dt} \right|_{t=0} f(\gamma(t)) = Df(p_*)(\dot{\gamma}(0)) = Df(p_*)(v) \quad (21.54)$$

Dies impliziert (21.53). □

*Beweis von 21.13.* Wir wissen bereits, dass die auf  $\mathbb{R}^n$  glatte Funktion

$$x \mapsto \langle x, Ax \rangle \quad (21.55)$$

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

$$\mathbb{R}^n \ni v \mapsto D\langle x, Ax \rangle(x_*)(v) = 2\langle x_*, Av \rangle \quad (21.56)$$

Mit (21.51) bedeutet dann (21.53), dass

$$v \in x_*^\perp \Rightarrow \langle x_*, Av \rangle = 0 \quad (21.57)$$

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

$$\mathbb{R}^n = \mathbb{R} \cdot x_* \oplus x_*^\perp \quad (21.58)$$

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  $\lambda$  mit  $Ax_* = \lambda x_*$ . Es folgt der Rest der Behauptungen. □

---

<sup>17</sup> $p_*$  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  $f(q) < f(p_*)$ .

# KAPITEL 8

## INTEGRATIONSTHEORIE

Die Integralrechnung spielt in der Physik auf mindestens zwei, a priori verschiedene, Weisen eine zentrale Rolle.

(i) Bei der Lösung von gewöhnlichen und partiellen Differentialgleichungen in der klassischen Mechanik und Elektrodynamik ist das Integral mit variablen Grenzen die Umkehrung der Ableitung, aufgefasst als lineare Abbildung zwischen geeigneten Funktionenräumen.

(ii) Das Integral einer Funktion, allenfalls geteilt durch das Volumen des Integrationsgebiets, berechnet als gewichtetes Mittel ihrer Werte Zustandssummen in der statistischen Physik sowie Erwartungswerte und Übergangsamplituden in der Quantenmechanik.

Ich finde die enge mathematische Verbindung zwischen diesen beiden Interpretationen einigermaßen überraschend. (Sie ist aber letztlich wohl einfach genau der Hauptsatz der Differential- und Integralrechnung!)

Der Aufbau der Theorie ist bedauernswerter Weise von einer ganzen Reihe technischer Schwierigkeiten belastet, ganz unabhängig davon, welchen Zugang man wählt (sei es über Masstheorie oder Vervollständigung von Funktionenräumen, mit Treppenfunktionen oder “differentialgeometrisch”). Wir werden hier nur so viele Details behandeln, als sie zum Verständnis und zur Wertschätzung unserer beiden grossen Ziele unumgänglich sind, als da wären:

(A) Die Integralsätze von Gauss und Stokes als “natürlichste” höher-dimensionale Verallgemeinerung des Hauptsatzes, und damit als Abschluss der Vektoranalysis.

(B) Die Identifikation der quadratintegrierbaren Funktionen auf dem  $\mathbb{R}^n$  als vollständigen unitären Raum, intermediär zwischen  $\mathcal{F}^{\text{fin}}(\mathbb{R}^n)$  und seinem Dualraum  $\mathcal{F}(\mathbb{R}^n)$ , und damit als Grundlage zur Verallgemeinerung der Eigenwert- und Darstellungstheorie auf unendlich-dimensionale Vektorräume.

### § 22 Integration im $\mathbb{R}^n$

Zur Erinnerung: Wir hatten im §14 (s. Def. 14.3) 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:

$$\int_a^b f(s)ds = \lim_{n \rightarrow \infty} \sum_{k=0}^{n-1} \left(\frac{b-a}{n}\right) \cdot f\left(a + \frac{k}{n}(b-a)\right) \quad (22.1)$$

Unsere Motivation dazu war die Umkehrung der ein-dimensionalen Ableitung, aufgefasst als lineare Abbildung<sup>18</sup>

$$\mathcal{C}^1(I, V) \rightarrow \mathcal{C}^0(I, V) \quad (22.2)$$

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

$$F(t) := \int_{t_*}^t f(s) ds \quad (22.3)$$

eine Stammfunktion von  $f$ , d.h.  $\dot{F}(t) = f(t)$ , und bis auf eine Konstante ist auch jede Stammfunktion von dieser Form.

Die Konzeption des Integrals als Umkehrung der Ableitung lässt sich nun nicht unmittelbar auf beliebige Funktionen von  $n > 1$  reellen Veränderlichen verallgemeinern, aus mehreren Gründen. Ist z.B.  $U \subset \mathbb{R}^n$  offen, und  $f \in \mathcal{C}^1(U, \mathbb{R})$ , so ist die Ableitung im Sinne des totalen Differentials per Def. 10.3 eine (stetige) Abbildung  $Df : U \rightarrow \text{Hom}(\mathbb{R}^n, \mathbb{R})$ , “lebt also bereits in einem ganz anderen Raum”. Und tatsächlich ist auch nicht jedes solches “Kotangentenvektorfeld”  $\lambda : U \rightarrow (\mathbb{R}^n)^\vee$  die Ableitung einer skalaren Funktion. Ist  $\lambda$  selbst stetig differenzierbar, so folgt eine notwendige lokale Bedingung aus der Symmetrie der zweiten Ableitung von  $f$  (s. Lemma 10.13), in der Form  $\partial_i \lambda_j = \partial_j \lambda_i$ . In den Theorie-Vorlesungen lautet diese Beobachtung (unter stillschweigendem Dazwischenschalten einer euklidischen Metrik) wohl üblicherweise: “Konservative Kraftfelder sind rotationsfrei”. Und global besitzt nur dann jedes  $\lambda$  (welches die lokale Bedingung erfüllt) eine “Stammfunktion”, wenn  $U$  einfach zusammenhängend ist (d.h. “keine Löcher hat”).

Für  $n = 2$  hatten wir ähnliche Zusammenhänge auch bereits in der Funktionentheorie angetroffen, s. §§15–17: Sind  $u$  und  $v$  reellwertige Funktion von  $x$  und  $y$ , und gilt  $\partial_y u = -\partial_x v$ , sowie  $\partial_x u = \partial_y v$ , so existiert lokal eine (reellwertige!) Funktion  $U$  mit  $u = \partial_x U$ ,  $v = -\partial_y U$ , sowie eine weitere reellwertige Funktion  $V$  mit  $u = \partial_y V$ ,  $v = \partial_x V$ . — Denn erfüllen Real- und Imaginärteil der komplexen Funktion  $f = u + iv$  die Cauchy-Riemann-Gleichungen (15.7), so existiert lokal eine holomorphe Stammfunktion  $F = U + iV$ , die in diesem Fall also eine “Funktion der gleichen Sorte” ist, und als Bonus ausserdem noch unendlich oft differenzierbar. Diese “Magie” beruht allerdings ganz wesentlich auf der Existenz einer Körperstruktur auf  $\mathbb{R}^2 \cong \mathbb{C}$ , und daher besteht keine Hoffnung auf eine Verallgemeinerung dieser Aussagen auf höhere Dimensionen, in denen eine solche algebraische Struktur einfach nicht zur Verfügung steht!

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” ihrer Funktionswerte:

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^n f\left(a + \frac{k}{n}(b-a)\right) = \frac{1}{b-a} \int_a^b f(s) ds \quad (22.4)$$

<sup>18</sup>Hier bezeichnet  $\mathcal{C}^0(I, V)$  den Vektorraum der stetigen, und  $\mathcal{C}^1(I, V)$  den der einmal stetig differenzierbaren Abbildungen von  $I$  mit Werten in  $V$ , vgl. Def. 10.14

## § 22. INTEGRATION IM $\mathbb{R}^n$

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

$$\begin{aligned} \mathbb{R}^n &= \{x : N \rightarrow \mathbb{R}\} \rightarrow \mathbb{R} \\ x &= (x^1, \dots, x^n)^T \mapsto \sum_{i=1}^n x^i \end{aligned} \quad (22.5)$$

(vgl. S. 166) der Permutationsdarstellung von  $S_n$  auf ihren invarianten Unterraum eine interessante Sorte Integral in diesem Sinne.) Man beachte allerdings, dass diese Interpretation des “Integrals von  $a$  nach  $b$ ” als “Mittelwert über das dazwischenliegende Intervall” von der Auszeichnung (bzw. dem Vergessen) einer Orientierung von  $\mathbb{R}$  abhängt. Schreiben wir ab jetzt (im Falle  $a < b$ ),

$$\int_{[a,b]} f(x)dx := \int_a^b f(s)ds \quad (22.6)$$

so gilt für  $a > b$ :

$$\int_a^b f(s)ds = - \int_{[b,a]} f(x)dx !!! \quad (22.7)$$

Eine andere (etwas weniger subtile) Charakterisierung dieser Unterscheidung ist die *Monotonie des Integrals* im Falle eines *angeordneten Wertebereichs* (d.h.  $V = \mathbb{R}$ ). Es gilt nämlich offenbar:

$$f(x) \geq 0 \forall x \in [a, b] \Rightarrow \int_{[a,b]} f(x)dx \geq 0 \quad (22.8)$$

Ausserdem haben wir hier den sog. “Mittelwertsatz der Integralrechnung”:

**Lemma 22.1.** *Es seien  $a, b \in \mathbb{R}$  mit  $a \leq b$  und  $f : [a, b] \rightarrow \mathbb{R}$  eine stetige Funktion<sup>19</sup>. Dann existiert ein  $\xi \in [a, b]$  so, dass*

$$\int_{[a,b]} f(x)dx = \int_a^b f(s)ds = f(\xi) \cdot (b - a) = f(\xi) \cdot |b - a| \quad (22.9)$$

*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 3.14 folgenkompakt ist, existieren gemäss Thm. 7.6 Stellen  $x_m$  und  $x_M$  in  $[a, b]$  so, dass

$$\begin{aligned} f(x_m) &= \inf f([a, b]) = \inf\{f(x) \mid x \in [a, b]\} \\ f(x_M) &= \sup f([a, b]) = \sup\{f(x) \mid x \in [a, b]\} \end{aligned} \quad (22.10)$$

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

$$\int_{[a,b]} f(x)dx - f(x_m) \cdot (b - a) = \int_{[a,b]} (f(x) - f(x_m))dx \geq 0 \quad (22.11)$$

<sup>19</sup>Wir lassen als Verallgemeinerung von 1.14 jetzt auch die Notation  $[a, a] := \{a\} \subset \mathbb{R}$  zu.

Ebenso gilt

$$\int_{[a,b]} f(x)dx \leq f(x_M) \cdot (b - a) \quad (22.12)$$

Zusammen

$$f(x_m) \leq \frac{1}{b - a} \cdot \int_{[a,b]} f(x)dx \leq f(x_M) \quad (22.13)$$

Mit Hilfe des Zwischenwertsatzes 7.1 folgt dann die Existenz einer Stelle  $\xi \in [a, b]$  mit

$$f(\xi) = \frac{1}{b - a} \cdot \int_{[a,b]} f(x)dx \quad (22.14)$$

wie gefordert. □

Wir geben also nun als Ziel dieses § aus, das Integral von Funktionen mehrerer Veränderlicher über Teilmengen des  $\mathbb{R}^n$ , interpretiert als gewichtete Summe ihrer Funktionswerte, zu definieren (und zu berechnen). (Und im §23 werden wir dann sehen, wie sich daraus auch in  $n$  Dimensionen (manchmal) Stammfunktionen durch Integration gewinnen lassen.) Wie bei der Funktionentheorie sind wir dabei von den Fragen begleitet, für welche Klasse von Funktionen, und für welche Sorten Teilmengen, die Konstruktion sinnvoll ist. Die Antwort wird zwar eine ganz andere sein, und wir werden auch (wie es sich gehört) unstetige Funktionen über ganz krumme Mengen integrieren können. Als Ausgangspunkt nehmen wir aber wieder stetige Funktionen, als “elementarste Definitionsbereiche” diesmal Produkte von Intervallen.

**Definition 22.2.** Ein *Quader* ist eine Teilmenge  $Q \subset \mathbb{R}^n$  der Form

$$Q = [a_1, b_1] \times \cdots \times [a_n, b_n] = \{x \in \mathbb{R}^n \mid a_i \leq x^i \leq b_i, i = 1, \dots, n\} \quad (22.15)$$

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

$$\text{vol}(Q) := \prod_{i=1}^n |b_i - a_i| \quad (22.16)$$

für das (euklidische!) *Volumen*. Offensichtlich ist  $\text{vol}(Q) = 0 \Leftrightarrow Q$  entartet.

### Topologischer Aufbau

Zur Wiederholung einiger Begriffe und Aussagen aus der HöMa II (s. S. 35 ff., sowie §8), fixieren wir auf dem  $\mathbb{R}^n$  eine beliebige Norm  $\|\cdot\|$ , mit der er zu einem vollständigen metrischen Raum wird. (Insbesondere konvergiert eine Folge  $(x_k) \subset \mathbb{R}^n$  genau dann, wenn sie eine Cauchy-Folge ist.)

## § 22. INTEGRATION IM $\mathbb{R}^n$

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

$$B_r(x) := \{y \in \mathbb{R}^n \mid \|x - y\| < r\} \subset U \quad (22.17)$$

· 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.

· Eine Teilmenge  $K \subset \mathbb{R}^n$  heisst *folgenkompakt* (wir sagen im Folgenden auch einfach “kompakt”), falls jede Folge  $(x_k) \subset K$  eine konvergente Teilfolge besitzt, deren Grenzwert noch in  $K$  liegt. Nach dem Satz 3.14 von Bolzano-Weierstrass ist dies genau dann der Fall, wenn  $K$  abgeschlossen und beschränkt ist. (Und  $K$  ist beschränkt genau dann, wenn ein  $R > 0$  existiert so dass  $K \subset \overline{B_R(0)} := \{y \in \mathbb{R}^n \mid \|y\| \leq R\}$ ).

· 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  $\mathbb{R}^n$ ) konvergenten Folgen mit Werten in  $X$ .

$$\overline{X} = \bigcap_{A \supset X} A = X \cup \left\{ \lim_{n \rightarrow \infty} x_n \mid (x_n)_{n \in \mathbb{N}} \subset X, x_n \text{ ist Cauchy-Folge} \right\} \quad (22.18)$$

$A \supset X$   
 $A$  abgeschlossen

· 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.

$$\overset{\circ}{X} = \bigcup_{\substack{U \subset X \\ U \text{ offen}}} U \quad (22.19)$$

· Der (topologische!) *Rand* von  $X$  ist

$$\partial X = \overline{X} \setminus \overset{\circ}{X} \quad (22.20)$$

Funktionen und Funktionenräume: · Wir interessieren uns letztlich auch wieder für das Integral mit Werten in vollständigen normierten Vektorräumen, die grösste Mühe entsteht aber bereits bei reellen Funktionen, und wir lassen zur Erleichterung daher den Wertebereich aus der Notation normalerweise weg.  $\mathcal{F}(\mathbb{R}^n) := \mathcal{F}(\mathbb{R}^n, \mathbb{R})$  ist dann der Vektorraum aller Funktionen auf  $\mathbb{R}^n$ , und  $\mathcal{C}(U) := \mathcal{C}^0(U)$  der Raum der stetigen Funktionen auf einer offenen Menge  $U$ , etc.

· Für eine Funktion  $f \in \mathbb{R}^n$  heisst

$$\text{supp}(f) := \overline{\{x \in \mathbb{R}^n \mid f(x) \neq 0\}} \quad (22.21)$$

der (topologische) *Träger* (Support) von  $f$ . (Man beachte den Konflikt mit (19.6) in Kapitel 7.)

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

$$\chi_X(x) := \begin{cases} 1 & \text{falls } x \in X \\ 0 & \text{falls } x \notin X \end{cases} \quad (22.22)$$

die *charakteristische Funktion* von  $X$ .

Übungsaufgabe :  $\text{supp}(\chi_X) = \overline{X}$ .

Konvergenz: · Eine Folge von Funktionen  $(f_k) \subset \mathcal{F}(X)$  ( $X \subset \mathbb{R}^n$  beliebig) heisst *punktweise konvergent* gegen  $f \in \mathcal{F}(X)$ , wenn  $\forall x \in X$  die reelle Folge  $(f_k(x)) \subset \mathbb{R}$  gegen  $f(x)$  konvergiert. Sie heisst *gleichmässig konvergent*, wenn

$$\lim_{k \rightarrow \infty} \|f - f_k\|_\infty = 0 \quad (22.23)$$

wobei  $\|f\|_\infty := \sup\{|f(x)| \mid x \in X\} \in \mathbb{R}^+ := \mathbb{R} \cup \{\infty\}$ . (Dies ist nicht eigentlich eine Norm, wir tun aber so. Manchmal schreiben wir auch  $\|\cdot\|_X$ .)

· Die Grenzfunktion einer gleichmässig konvergenten Folge von stetigen Funktionen ist ebenfalls stetig (Theorem 8.5), d.h.  $(\mathcal{C}(X), \|\cdot\|_\infty)$  ist vollständig.

· Auf einer kompakten Menge ist jede stetige Funktion gleichmässig stetig (Theorem 8.9).

### Iterierte Integration

Wir können nun eine erste praktische Definition des  $n$ -dimensionalen Integrals angeben.

**Lemma 22.3.** Sei  $Q = \times_{i=1}^n [a_i, b_i] \subset \mathbb{R}^n$  ein Quader und  $f \in \mathcal{C}(Q)$ . Dann ist die Funktion

$$(x^2, \dots, x^n) \mapsto \int_{[a_1, b_1]} f(x^1, x^2, \dots, x^n) dx^1 \quad (22.24)$$

stetig auf dem Quader  $\times_{i=2}^n [a_i, b_i] \subset \mathbb{R}^{n-1}$ .

*Beweis.* (s. auch Theorem 14.8)  $Q$  ist kompakt. Die gleichmässige Stetigkeit bedeutet, dass für jedes  $\epsilon > 0$  ein  $\delta > 0$  existiert so dass  $|f(x_1) - f(x_2)| < \epsilon$  falls nur  $\|x_1 - x_2\| < \delta$ . Schreiben wir zur Abkürzung  $y = (x^2, \dots, x^n)$  und ist  $\|y_1 - y_2\| < \delta$ , so gilt (zur Not als Folge der Dreiecksungleichung)  $\|(x^1, y_1) - (x^1, y_2)\| < \delta$  für alle  $x^1 \in [a_1, b_1]$ , und daher folgt mit Hilfe der Standardabschätzung des Integrals

$$\begin{aligned} & \left| \int_{[a_1, b_1]} f(x^1, y_1) dx^1 - \int_{[a_1, b_1]} f(x^1, y_2) dx^1 \right| \\ & \leq |b_1 - a_1| \cdot \sup\{|f(x^1, y_1) - f(x^1, y_2)| \mid x^1 \in [a_1, b_1]\} \\ & \leq |b_1 - a_1| \cdot \epsilon \end{aligned} \quad (22.25)$$

□

Dies lässt sich offensichtlich iterieren, und wir setzen

**Definition 22.4.**

$$\int_Q f(x) dx := \int_{[a_n, b_n]} \left( \dots \int_{[a_2, b_2]} \left( \int_{[a_1, b_1]} f(x^1, \dots, x^n) dx^1 \right) dx^2 \dots \right) dx^n \quad (22.26)$$

Diese Definition erfreut sich offensichtlich der folgenden charakteristischen Eigenschaften eines Integrals (vgl. Lemma 14.6):

(i) Linearität und Integration von Konstanten:

$$\int_Q (f_1 + \lambda f_2 + c)(x) dx = \int_Q f_1(x) dx + \lambda \int_Q f_2(x) dx + c \prod_{i=1}^n (b_i - a_i) \quad (22.27)$$

(ii) Monotonie (und *kein* Vorzeichen- bei Orientierungswechsel!):

$$\int_Q f(x) dx \geq 0 \quad \text{falls } f(x) \geq 0 \quad \forall x \in Q \quad (22.28)$$

(iii) Standardabschätzung:

$$\left| \int_Q f(x) dx \right| \leq \|f\|_Q \cdot \underbrace{\prod_i |b_i - a_i|}_{=\text{vol}(Q)} \quad (22.29)$$

Zur Additivität im Integrationsbereich können wir ohne weiteres nur sagen, dass

$$\int_Q f(x) dx = \sum_{K=1}^R \int_{Q_K} f(x) dx \quad (22.30)$$

wenn die  $\mathcal{K} = \{Q_K | K = 1, \dots, R\}$  eine  $n$ -dimensionale *Karierung* von  $Q$  sind: Die Intervalle  $[a_i, b_i]$  sind zerlegt gemäss  $a_i = s_{i,0} < s_{i,1} < \dots < s_{i,r_i} = b_i$  und die  $Q_K$  sind alle die von der Form  $\times_{i=1}^n [s_{i,k_i}, s_{i,k_i+1}]$  ( $k_i = 0, \dots, r_i - 1$ ; es sind dies  $R = \prod_i r_i$  Stück.) Dies folgt sofort aus der Intervalladditivität (und Linearität) des ein-dimensionalen Integrals, und reicht bereits aus um nachzuweisen, dass der Wert des Integrals nicht davon abhängt, in welcher Reihenfolge wir über die verschiedenen  $x^i$  abintegrieren.<sup>20</sup>

**Lemma 22.5.** *Für jedes  $\epsilon > 0$  existiert eine Karierung von  $Q$  in eine geeignete Zahl  $R$  Quader gleichen Volumens, und Zahlen  $c_K$ ,  $K = 1, \dots, R$  derart, dass*

$$\left| \int_Q f(x) dx - \underbrace{\frac{\text{vol}(Q)}{R} \sum_{K=1}^R c_K}_{=\text{vol}(Q_K)} \right| < \epsilon \quad (22.32)$$

*Beweis.* Da  $f$  auf der kompakten Menge  $Q$  gleichmässig stetig ist, existiert ein  $\delta > 0$  so dass  $|f(x_1) - f(x_2)| < \frac{\epsilon}{\text{vol}(Q)}$  falls  $\|x_1 - x_2\| \leq \delta$ . Sei  $\mathcal{K}$  so fein, dass jeder der  $Q_K$  Durchmesser kleiner als  $\delta$  hat (in der gewählten Norm, d.h.  $\sup\{\|x_1 - x_2\| | x_1, x_2 \in Q_K\} < \delta$ ). Wähle beliebige Stützstellen  $x_K \in Q_K$  und setze  $c_K = f(x_K)$ . Dann folgt

<sup>20</sup>Es ist möglich, diese Aussage anzugehen, ohne die Welt der stetigen Funktionen zu verlassen. Skizze: Die Aussage ist offensichtlich richtig für Produkte von Funktionen der einzelnen Variablen, sowie endliche Linearkombinationen hiervon, d.h. für alle Funktionen im algebraischen Tensorprodukt,

$$\mathcal{A}(Q) = \otimes_{i=1}^n \mathcal{C}([a_i, b_i]) \subset \mathcal{C}(Q) \quad (22.31)$$

$\mathcal{A}(Q)$  ist aber ausserdem offensichtlich eine Unteralgebra von  $\mathcal{C}(Q)$ , und punktstetig. Sie erfüllt also die Voraussetzungen des Satzes von Stone-Weierstrass und liegt daher dicht in  $\mathcal{C}(Q)$  (in der Supremumsnorm). Dann argumentiert man wie im Beweis des Lemmas.

für alle  $K$  aus den obigen Eigenschaften des Integrals

$$\begin{aligned} \left| \int_{Q_K} f(x) dx - c_K \operatorname{vol}(Q_K) \right| &= \left| \int_{Q_K} (f(x) - c_K) dx \right| \\ &\leq \|f - f(x_K)\|_{Q_K} \cdot \operatorname{vol}(Q_K) < \epsilon \underbrace{\frac{\operatorname{vol}(Q_K)}{\operatorname{vol}(Q)}}_{=\frac{1}{R}} \end{aligned} \quad (22.33)$$

Summe über  $K$  und (22.30) ergibt die Behauptung.  $\square$

Offenbar geht in den Beweis die Integrationsreihenfolge gar nicht ein (d.h. die Abschätzung gilt mit den gleichen  $Q_K$  und  $c_K$  für beliebige Reihenfolge), und daher folgt:

**Korollar 22.6.** *Für jede Permutation  $\sigma \in S_n$  gilt*

$$\int_Q f(x) dx = \int_{[a_{\sigma(n)}, b_{\sigma(n)}]} \left( \cdots \left( \int_{[a_{\sigma(1)}, b_{\sigma(1)}]} f(x^1, \dots, x^n) dx^{\sigma(1)} \right) \cdots \right) dx^{\sigma(n)} \quad (22.34)$$

### Kompakter Träger oder Integrationsbereich

Das an dieser Stelle vielleicht dringlichste Manko der Definition 22.4 ist die Beschränkung auf Quader als Integrationsbereiche. (Mittelwerte wollen wir ja eigentlich über beliebige Mengen berechnen!) Für  $n = 1$  hatten wir in der HöMa II bereits durch einen Grenzübergang das “uneigentliche” Integral über nicht kompakte Intervalle definiert, s. (14.34), (und mit Hilfe der Funktionentheorie in vielen Fällen berechnet). In höheren Dimensionen kommt eine noch grössere Vielfalt von Integrationsbereichen in Betracht, so z.B. alle kompakten Mengen, Kugeln etc..

Darüber hinaus wird es sich aber als nötig erweisen, auch recht *unstetige* Funktionen integrieren zu können, und zwar spätestens bei der Anwendung zur Lösung von physikalischen Problemen durch Approximation, Iteration, und Ausnutzen der Vollständigkeit (vgl. §15). So ist es etwa bei einer statistischen Interpretation natürlich, die Güte einer Approximation nicht durch die “Supremums-Norm” zu messen, sondern eben genau durch Mittelwertbildung, d.h. Integration. Die Sicherung der Vollständigkeit (s. Beispiel 8.6) erfordert dann die Betrachtung unstetiger Funktionen, jedenfalls in Zwischenschritten.<sup>21</sup>

Eine Verbindung schien bereits beim Beweis von Lemma 22.5 durch: Der Ausdruck  $\sum_K \operatorname{vol}(Q_K) c_K$  lässt sich offensichtlich interpretieren als Integral über  $Q$  der unstetigen Funktion

$$\sum_K c_K \chi_{Q_K} \quad (22.35)$$

(wo  $\chi_{Q_K}$  die charakteristische Funktion von  $Q_K$  ist mit  $\int_Q \chi_{Q_K}(x) dx = \int_{Q_K} 1 dx = \operatorname{vol}(Q_K)$ ), die die stetige Funktion  $f(x)$  gleichmässig approximiert, und die Rechnung ist nicht anderes als die Abschätzung des Mittelwerts durch das Supremum.

<sup>21</sup>Tatsächlich sind die physikalisch relevanten Lösungen häufig wieder stetig (sogar glatt). Die Philosophie ist aber ähnlich wie die Vervollständigung der rationalen zu den reellen Zahlen.

## § 22. INTEGRATION IM $\mathbb{R}^n$

Eine systematische Fortsetzung dieser Idee führt evtl. mit einem Umweg über *Treppenfunktionen* zur *Maßtheorie* (s. Einschübe).

Wir werden (stattdessen) die Strategie verfolgen, das Integral 22.4 mit seinen charakteristischen Eigenschaften Schritt für Schritt auf immer grössere Unterräume von  $\mathcal{F}(\mathbb{R}^n)$  und (Klassen von) Integrationsbereichen fortzusetzen.<sup>22</sup> Am einfachsten ist der folgende:

**Definition 22.7.** Wir schreiben für den *Vektorraum der stetigen Funktionen mit kompaktem Träger*

$$\mathcal{C}_c(\mathbb{R}^n) := \{f \in \mathcal{F}(\mathbb{R}^n) \mid \text{supp}(f) \text{ kompakt}\} \quad (22.36)$$

Da  $\text{supp}(f)$  per Definition abgeschlossen ist, gilt offensichtlich:

$$\begin{aligned} f \in \mathcal{C}_c(\mathbb{R}^n) &\Leftrightarrow f \text{ ist stetig und } f(x) = 0 \text{ für } \|x\| > R \text{ für ein } R > 0 \\ &\Leftrightarrow \text{Es existiert ein Quader } Q \subset \mathbb{R}^n \text{ mit } \text{supp}(f) \subset Q. \end{aligned} \quad (22.37)$$

Für  $f \in \mathcal{C}_c(\mathbb{R}^n)$  setzen wir dann

$$\int_{\mathbb{R}^n} f(x) dx = \int_Q f(x) dx \quad (22.38)$$

für ein beliebiges  $Q \supset \text{supp}(f)$ . Der Wert des Integrals ist offensichtlich unabhängig von der Wahl von  $Q$ . (Ist  $Q' \supset \text{supp}(f)$  ein anderer Quader, so kariert man die Differenz von  $Q$  und  $Q'$ , auf der  $f$  verschwindet, und benutzt (22.30).)

· Ist  $U \subset \mathbb{R}^n$  offen, und  $f \in \mathcal{C}_c(U)$  stetig mit kompaktem Träger in  $U$ , so ist die Fortsetzung

$$\tilde{f}(x) = \begin{cases} f(x) & x \in U \\ 0 & x \in \mathbb{R}^n \setminus U \end{cases} \quad (22.39)$$

stetig auf ganz  $\mathbb{R}^n$  und mit kompaktem Träger und wir setzen

$$\int_U f(x) dx = \int_{\mathbb{R}^n} \tilde{f}(x) dx \quad (22.40)$$

Bei der Vervollständigung der Integrationstheorie (im Sinne eines “vernünftigen linearen Funktional<sup>23</sup> auf einer maximalen Klasse von integrierbaren Funktionen”) im §24 wollen wir insbesondere erreichen, dass für sinnvolle (z.B. kompakte) Mengen  $K \subset \mathbb{R}^n$  die charakteristische Funktion  $\chi_K$  integrierbar ist, und dass wir für (z.B. stetige) Funktionen  $f$  schreiben können

$$\int_K f(x) dx \stackrel{!}{=} \int_{\mathbb{R}^n} \chi_K(x) f(x) dx \quad (22.41)$$

(Dabei werden wir natürlich sicherstellen, dass dies mit dem schon definierten Spezialfall, in dem  $K = Q$  ein Würfel ist, verträglich ist.)

<sup>22</sup>Tatsächlich 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”.

<sup>23</sup>Ein Funktional ist eine Linearform auf einem Vektorraum von Funktionen.

### Treppenfunktionen

\*\*\* werden im Folgenden nicht benötigt \*\*\*

· 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

$$\varphi = \sum_{i=1}^k c_i \chi_{Q_i} \tag{22.42}$$

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  $Q = [a_1, b_1] \times \dots \times [a_n, b_n]$ , und setze für  $i = 1, \dots, n$

$$\begin{aligned} \chi_i &:= \chi_{\mathbb{R}^{i-1} \times [a_i, b_i] \times \mathbb{R}^{n-i}}, & \overset{\circ}{\chi}_i &:= \chi_{\mathbb{R}^{i-1} \times (a_i, b_i) \times \mathbb{R}^{n-i}}, \\ \varphi_i &:= \chi_{\mathbb{R}^{i-1} \times \{a_i\} \times \mathbb{R}^{n-i}}, & \psi_i &:= \chi_{\mathbb{R}^{i-1} \times \{b_i\} \times \mathbb{R}^{n-i}} \end{aligned} \tag{22.43}$$

Dann gilt

$$\begin{aligned} \overset{\circ}{\chi}_i &= \chi_i - \varphi_i - \psi_i \\ \chi_Q &= \prod_{i=1}^n \chi_i & \chi_{\overset{\circ}{Q}} &= \prod_{i=1}^n \overset{\circ}{\chi}_i \end{aligned} \tag{22.44}$$

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, halboffener, 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 (22.42) zerlegen (Übungsaufgabe).

· Man beachte insbesondere, dass die Darstellung einer Treppenfunktion wie in (22.42) nicht eindeutig ist. Für  $n = 1$  gilt beispielsweise

$$\chi_{[a,b]} = \chi_{[a, \frac{a+b}{2}]} + \chi_{[\frac{a+b}{2}, b]} - \chi_{\{\frac{a+b}{2}\}} \tag{22.45}$$

· Aus diesen Betrachtungen folgt etwa auch, dass

$$\mathcal{T}(\mathbb{R}^n) = \text{span}\{\chi_Q \mid Q \subset \mathbb{R}^n \text{ ein Quader}\} \subset \mathcal{F}(\mathbb{R}^n) \tag{22.46}$$

(Charakteristische Funktionen abgeschlossener Quader bilden nur eben keine Basis von  $\mathcal{T}(\mathbb{R}^n)$ , siehe (22.45).)

§ 22. INTEGRATION IM  $\mathbb{R}^n$

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

$$I_{\mathcal{T}}(\varphi) := \sum_{i=1}^k c_i \operatorname{vol}(Q_i) \quad (22.47)$$

tatsächlich *nicht* von der gewählten Darstellung (22.42) abhängt. (Beweis wird nachgereicht), und dass sie ausserdem übereinstimmt mit dem (noch zu definierenden) Integral von  $\varphi$  über  $\mathbb{R}^n$ ,

$$I_{\mathcal{T}}(\varphi) = \int_{\mathbb{R}^n} \varphi(x) dx \quad (22.48)$$

Jedenfalls überprüfen wir leicht:

**Lemma 22.8.** (i) Für alle  $\varphi_1, \varphi_2 \in \mathcal{T}(\mathbb{R}^n)$ ,  $\lambda \in \mathbb{R}$  gilt

$$I_{\mathcal{T}}(\lambda\varphi_1 + \varphi_2) = \lambda I_{\mathcal{T}}(\varphi_1) + I_{\mathcal{T}}(\varphi_2) \quad (22.49)$$

(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  $|I_{\mathcal{T}}(\varphi)| \leq I_{\mathcal{T}}(|\varphi|)$ .<sup>24</sup>

*Beweis.* wird nachgeliefert □

und können auch  $I_{\mathcal{T}}$  auf  $\mathcal{C}_c(\mathbb{R}^n)$  übertragen und überprüfen, dass es dort mit dem Integral 22.7 übereinstimmt:

· Wir benutzen hierzu wieder, dass stetige Funktionen auf folgenkompakten Mengen gleichmässig stetig sind (s. Thm. 8.9). Es folgt daraus, dass sich stetige Funktionen mit kompaktem Träger gleichmässig durch Treppenfunktionen approximieren lassen. Genauer gilt:

**Lemma 22.9.** Es sei  $f \in \mathcal{C}_c(\mathbb{R}^n)$  und  $Q$  ein Quader mit  $Q \supset \operatorname{supp}(f)$ . Dann existiert für jedes  $\epsilon > 0$  ein  $\varphi \in \mathcal{T}(\mathbb{R}^n)$  mit  $\operatorname{supp}(\varphi) \subset Q$  und

$$\|f - \varphi\|_{\infty} \left( := \sup\{|f(x) - \varphi(x)| \mid x \in \mathbb{R}^n\} \right) < \epsilon \quad (22.50)$$

*Beweis.* (Die wesentliche Idee ist die gleiche wie im Beweis von 22.5, allerdings ist wegen der Missachtung der Quaderränder (22.35) *noch nicht* notwendig eine gleichmässige Approximation von  $f$ .) Da  $Q$  kompakt ist und  $f$  stetig, existiert ein  $\delta > 0$  so, dass

$$|f(x) - f(y)| < \epsilon \quad \text{für alle } x, y \in Q \text{ mit } \|x - y\| < \delta \quad (22.51)$$

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

$$\operatorname{diam}(Q'_i) \left( := \sup\{\|x - y\| \mid x, y \in Q'_i\} \right) < \delta \quad (22.52)$$

---

<sup>24</sup> $|\varphi|(x) := |\varphi(x)|$

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

$$\varphi = \sum_{i=1}^k f(x_i) \chi_{Q'_i} \quad (22.53)$$

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  $x \in Q$

$$|f(x) - \varphi(x)| = |f(x) - f(x_{i_x})| < \epsilon \quad (22.54)$$

wegen (22.51) und da  $\|x - x_{i_x}\| \leq \text{diam}(Q'_{i_x}) < \delta$ . Da  $f$  und  $\varphi$  ausserhalb von  $Q$  verschwinden, folgt (22.50).  $\square$

Die folgende Rekonstruktion des Integrals auf  $\mathcal{C}_c(\mathbb{R}^n)$  aus  $I_{\mathcal{T}}$  illustriert die Strategie, die wir im folgenden § (mit etwas anderen Grenzprozessen) noch häufiger durchziehen werden. (Sie ist aber etwas kompliziert und wie der gesamte Einschub logisch auch nicht notwendig.)

**Proposition/Definition 22.10.** *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*

$$I_{\mathcal{C}_c}(f) := \lim_{k \rightarrow \infty} I_{\mathcal{T}}(\varphi_k) \quad (22.55)$$

hängt weder von dem Quader  $Q$  noch von der Folge  $(\varphi_k)$  ab.

*Beweis.* Die Existenz einer Folge mit den geforderten Eigenschaften folgt aus dem Lemma 22.9. Für jedes  $\epsilon > 0$  existiert dann ein  $K$  so, dass für alle  $k, l \geq K$  und für alle  $x \in Q$ :

$$\begin{aligned} |\varphi_k(x) - \varphi_l(x)| &\leq |\varphi_k(x) - f(x)| + |f(x) - \varphi_l(x)| \\ &\leq \|f - \varphi_k\|_{\infty} + \|f - \varphi_l\|_{\infty} < \frac{\epsilon}{\text{vol}(Q)} \end{aligned} \quad (22.56)$$

Für  $x \notin Q$  ist  $\varphi_k(x) - \varphi_l(x) = 0$ . Mit Hilfe von Lemma 22.8 und  $I_{\mathcal{T}}(\chi_Q) = \text{vol}(Q)$  folgt für  $k, l \geq K$

$$\begin{aligned} |I_{\mathcal{T}}(\varphi_k) - I_{\mathcal{T}}(\varphi_l)| &= |I_{\mathcal{T}}(\varphi_k - \varphi_l)| \leq I_{\mathcal{T}}(|\varphi_k - \varphi_l|) \\ &\leq I_{\mathcal{T}}(\|\varphi_k - \varphi_l\|_{\infty} \cdot \chi_Q) < \frac{\epsilon}{\text{vol}(Q)} \cdot \text{vol}(Q) = \epsilon \end{aligned} \quad (22.57)$$

$(I_{\mathcal{T}}(\varphi_k))$  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  $\tilde{Q}$  ein Quader mit  $Q \cap \tilde{Q} \subset \tilde{Q}$ . Dann folgt aus  $\|\varphi_k - \tilde{\varphi}_k\|_{\infty} \rightarrow 0$  mit einer Abschätzung auf  $\tilde{Q}$ , dass  $\lim I_{\mathcal{T}}(\tilde{\varphi}_k) = \lim I_{\mathcal{T}}(\varphi_k)$ .  $\square$

**Proposition 22.11.**  $I_{\mathcal{C}_c} : \mathcal{C}_c(\mathbb{R}^n) \rightarrow \mathbb{R}$  ist linear, monoton, und für  $f \in \mathcal{C}_c(\mathbb{R}^n)$  mit  $\text{supp}(f) \subset Q$  für einen Quader  $Q$  gilt

$$|I_{\mathcal{C}_c}(f)| \leq \|f\|_{\infty} \cdot \text{vol}(Q) \quad (22.58)$$

§ 22. INTEGRATION IM  $\mathbb{R}^n$

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

Wir wollen schliesslich noch zeigen, dass  $I_{c_c} = \int_{\mathbb{R}^n} \cdot dx$  aus Def. 22.7. Im Falle  $n = 1$  überzeugt man sich, modulo der Bemerkungen bei (22.7), leicht davon, dass die Karierungen natürlich nichts anderes sind als die Intervallzerlegungen aus 14.3. Für die Iteration gehen wir vor wie oben.

**Proposition 22.12.** 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

$$\mathbb{R}^n \ni x \mapsto f^{(y)}(x) := f(x, y) \in \mathbb{R} \tag{22.59}$$

stetig mit kompaktem Träger.

(ii) Die Funktion

$$\mathbb{R}^m \ni y \mapsto g(y) := I_{c_c}(f^{(y)}) \tag{22.60}$$

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. Nach dem Beweis lautet die Aussage in der üblichen Notation:

$$\int_{\mathbb{R}^m} \left( \int_{\mathbb{R}^n} f(x, y) dx \right) dy = \int_{\mathbb{R}^{n+m}} f(x, y) d(x, y) \tag{22.61}$$

(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 gelten. 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 (22.46) 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

$$I_{\mathcal{T}}(\chi_{Q \times P}) = \text{vol}(Q \times P) = \text{vol}(Q) \cdot \text{vol}(P) = I_{\mathcal{T}} \left( \underbrace{I_{\mathcal{T}}(\chi_{Q \times P}^{(\cdot)})}_{=\text{vol}(Q) \cdot \chi_P} \right) \tag{22.62}$$

Für stetige Funktionen mit kompaktem Träger sind die Aussage (i) und die Beschränktheit des Träger von  $g$  offensichtlich. Die Stetigkeit von  $g$  folgt wie im Beweis von 22.3 aus der gleichmässigen Stetigkeit und der Standardabschätzung in 22.11.

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 gilt<sup>25</sup>

$$|I_{c_c}(f) - I_{\mathcal{T}}(\varphi)| < \epsilon \cdot \text{vol}(Q \times P) \tag{22.63}$$

<sup>25</sup>Betrachte  $\lim |I_{\mathcal{T}}(\varphi_k) - I_{\mathcal{T}}(\varphi)|$  für  $\|f - \varphi_k\|_{\infty} \rightarrow 0$ .

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

$$\|g - \psi\|_\infty < \epsilon \cdot \text{vol}(Q) \tag{22.64}$$

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

$$|I_{C_c}(g) - I_{\mathcal{T}}(\psi)| < \epsilon \cdot \text{vol}(Q) \cdot \text{vol}(P) \tag{22.65}$$

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

$$|I_{C_c}(g) - I_{C_c}(f)| \leq |I_{C_c}(g) - I_{\mathcal{T}}(\psi)| + |I_{\mathcal{T}}(\varphi) - I_{C_c}(f)| < 2\epsilon \cdot \text{vol}(Q) \cdot \text{vol}(P) \tag{22.66}$$

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

## Maßtheorie

\*\*\* Nur als Anriss gedacht \*\*

· 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” auszuzeichnen<sup>26</sup>, und ihnen ein gewisses “Gewicht” oder “Volumen” (das *Maß*)  $\mu(A) \in \mathbb{R}^+ = [0, \infty]$  für  $A \in \mathcal{A}$  zuzuweisen. Dabei sind  $\mathcal{A}$  und  $\mu$  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

$$\mu(A_1 \cup A_2) = \mu(A_1) + \mu(A_2) - \mu(A_1 \cap A_2) \tag{22.67}$$

Man verlangt auch Verallgemeinerungen dieser Bedingungen auf *abzählbare*<sup>27</sup> 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.

$$\mu(A + x) = \mu(A) \quad \forall A \in \mathcal{A}, x \in \mathbb{R}^n \tag{22.68}$$

und

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

$$\mu([0, 1] \times \cdots \times [0, 1]) = 1 \tag{22.69}$$

<sup>26</sup> $\mathcal{P}(X)$  ist die “Potenzmenge”, d.h. die Menge aller Teilmengen, von  $X$ .

<sup>27</sup>Eine (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. 3.21.

## § 22. INTEGRATION IM $\mathbb{R}^n$

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.

· Unsere Vervollständigungsstrategie im §24 ist moralisch sehr ähnlich, geht aber aus von der (m.E. physikalisch leichter zu motivierenden) Mittelwertinterpretation anstatt von der (u.U. intuitiveren) “Volumenmessung”.

· Ein 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  $\mu(X) = 1$ .

· Für den  $\mathbb{R}^n$  ist das Ergebnis letztlich das gleiche, was wir andeuten, aber nicht vollständig zeigen werden.

---

### Monotone Konvergenz

Mit welchen Grenzprozessen sollen wir nun unsere Integration verfolgen?

· Die Standardabschätzung des Integrals garantiert insbesondere, dass für  $(f_k) \subset C_c(\mathbb{R}^n)$

$$\lim_{k \rightarrow \infty} \|f_k - f\|_\infty = 0 \quad \Rightarrow \quad \lim_{k \rightarrow \infty} \int_{\mathbb{R}^n} f_k(x) dx = \int_{\mathbb{R}^n} f(x) dx \quad (22.70)$$

Die Umkehrung gilt (natürlich) nicht, d.h. also Konvergenz der Integrale impliziert nicht gleichmässige Konvergenz der Funktionenfolge (s. Beispiel 22.13). Das Integral alleine ist aber auch zu schwach. Heidelberg Tradition folgend<sup>28</sup> werden wir das “Prinzip der monotonen Konvergenz” (vgl. Thm. 3.12)

Eine monotone Folge reeller Zahlen ist entweder unbeschränkt oder konvergent.

welches sich bei der Konstruktion der reellen Zahlen mittels Dedekindschen Schnitten (s. Thm. 1.12) schon als nützlich erwiesen hat, von einzelnen Zahlen auf Funktionen ausdehnen, im §24 systematisch, hier zunächst nur für eine spezielle Anwendung. Verabredung: · Für  $f, g \in \mathcal{F}(\mathbb{R}^n, \mathbb{R})$  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  $f_{k+1} \geq f_k \forall k$  ( $f_{k+1} \leq f_k \forall k$ ).

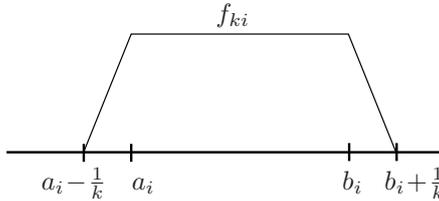
· Wir schreiben  $f_k \uparrow f$  (bzw.  $f_k \downarrow f$ ), falls  $(f_k)$  punktweise monoton wachsend (fallend) gegen  $f$  konvergiert.

Warum dies für die Integration stetiger Funktionen über kompakte Mengen relevant sein könnte (und sein wird), illustriert das

---

<sup>28</sup>Ich lehne mich an das Weissauersche Skript an, frühere Versionen stammen von Prof. Freitag.

**Beispiel 22.13.** Sei  $Q = [a_1, b_1] \times \dots \times [a_n, b_n]$  ein Quader und für  $k = 1, 2, \dots$  und  $i = 1, \dots, n$ ,  $x \in \mathbb{R}$ :

$$\varphi_{ik}(x) := \begin{cases} 0 & x < a_i - \frac{1}{k} \\ k(x - a_i + \frac{1}{k}) & x \in [a_i - \frac{1}{k}, a_i] \\ 1 & x \in [a_i, b_i] \\ 1 - k(x - b_i) & x \in [b_i, b_i + \frac{1}{k}] \\ 0 & x > b_i + \frac{1}{k} \end{cases} \quad (22.71)$$


Dann gilt  $\varphi_{ik} \in \mathcal{C}_c(\mathbb{R})$  und  $\varphi_{ik} \downarrow \chi_{[a_i, b_i]}$  für  $k \rightarrow \infty$  (die Konvergenz ist aber nicht gleichmässig). Daher erfüllen die Funktionen  $\varphi_k$ , für  $x = (x^1, \dots, x^n)^T \in \mathbb{R}^n$  definiert durch,

$$\varphi_k(x) := \prod_{i=1}^n \varphi_{ik}(x^i) \quad (22.72)$$

$\varphi_k \in \mathcal{C}_c(\mathbb{R}^n)$  und  $\varphi_k \downarrow \chi_Q$ . Es gilt

$$\begin{aligned} \int_{\mathbb{R}^n} \varphi_k(x) dx &= \prod_{i=1}^n \left( \int_{\mathbb{R}} \varphi_{ki}(x^i) dx^i \right) \\ &= \prod_{i=1}^n \left( b_i - a_i + \frac{1}{k} \right) \rightarrow \text{vol}(Q) \text{ für } k \rightarrow \infty \end{aligned} \quad (22.73)$$

wobei die letzte Konvergenz monoton fallend ist.

- Ist dann  $f \in \mathcal{C}(\mathbb{R}^n)$  (nicht notwendig kompakter Träger, aber) mit  $f \geq 0$ , so gilt  $f \cdot \varphi_k \downarrow f \chi_Q$  und die Folge  $\int_{\mathbb{R}^n} f(x) \varphi_k(x) dx$  konvergiert monoton fallend gegen  $\int_Q f(x) dx$ . (Übungsaufgabe?)

- Ist hingegen  $f \leq 0$ , so gilt  $f \cdot \varphi_k \uparrow f \chi_Q$ , und die Folge  $\int_{\mathbb{R}^n} f(x) \varphi_k(x) dx$  konvergiert monoton wachsend gegen  $\int_Q f(x) dx$ .

- Ist  $f \in \mathcal{C}(\mathbb{R}^n)$  mit uneinheitlichem Vorzeichen, so zerlegen wir  $f = f_+ + f_-$ , wobei  $f_+ = \max(f, 0) \geq 0$  und  $f_- = \min(f, 0) \leq 0$ .  $f_{\pm}$  sind beide in  $\mathcal{C}(\mathbb{R}^n)$  und  $\int_{\mathbb{R}^n} f(x) \varphi_k(x) dx$  konvergiert gegen  $\int_Q f(x) dx$  (allerdings nicht monoton).

Wie wir sehen werden, liefert die konsequente Umsetzung dieser monotonen Konvergenz eine sinnvolle und flexible Integrationstheorie, zunächst für Wertebereich  $\mathbb{R}$ , und durch lineare Fortsetzung auch für  $\mathbb{C}$  oder andere endlich-dimensionale Vektorräume.

Dass die monotone Konvergenz bei den stetigen Funktionen für sich (für die (22.70) gilt) nichts ändern wird, garantiert der folgende ‘Satz von Dini’

**Proposition 22.14.** *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  $(f_k)$  gleichmässig gegen  $f$ .*

bzw. sein

§ 22. INTEGRATION IM  $\mathbb{R}^n$

**Korollar 22.15.** Konvergiert  $(f_k) \subset \mathcal{C}_c(\mathbb{R}^n)$  punktweise monoton gegen  $f \in \mathcal{C}_c(\mathbb{R}^n)$  so gilt

$$\int_{\mathbb{R}^n} f_k(x) dx \rightarrow \int_{\mathbb{R}^n} f(x) dx \text{ für } k \rightarrow \infty \quad (22.74)$$

*Beweis von 22.14.* 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  $f_k \downarrow 0$ . Wir argumentieren dann indirekt.

· 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 ggfs. 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  $|f_K(x) - f_K(x_*)| < \frac{\epsilon_*}{2}$  falls  $\|x - x_*\| < \delta$  (Stetigkeit von  $f_K$ ). Für  $k \geq K$  gross genug ist dann  $\|x_k - x_*\| < \delta$  und es folgt

$$f_k(x_k) \leq f_K(x_k) \leq f_K(x_*) + \frac{\epsilon_*}{2} < \frac{\epsilon_*}{2} + \frac{\epsilon_*}{2} = \epsilon_* \quad (22.75)$$

$\uparrow$   
 Monotonie!

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

*Beweis von 22.15.* folgt mit der gleichmässigen Konvergenz von  $(f_k)$  sofort aus (22.70). □

**Beispiel 22.16.** Gemäss Übungsaufgabe ist für eine nicht-leere Menge  $X \subset \mathbb{R}^n$  die Funktion  $d(\cdot, X) : \mathbb{R}^n \rightarrow \mathbb{R}$ ,

$$d(x, X) := \inf\{\|x - y\| \mid y \in X\} \quad (22.76)$$

stetig, und es gilt

$$d(x, X) = 0 \Leftrightarrow x \in \overline{X} \quad (22.77)$$

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

$$\varphi_k(x) := (\max(1 - d(x, X), 0))^k \quad (22.78)$$

stetig mit kompaktem Träger. Die Folge  $(\varphi_k)$  fällt monoton und konvergiert für  $k \rightarrow \infty$  punktweise gegen die charakteristische Funktion  $\chi_{\overline{X}}$ , welche i.A. nicht stetig ist.

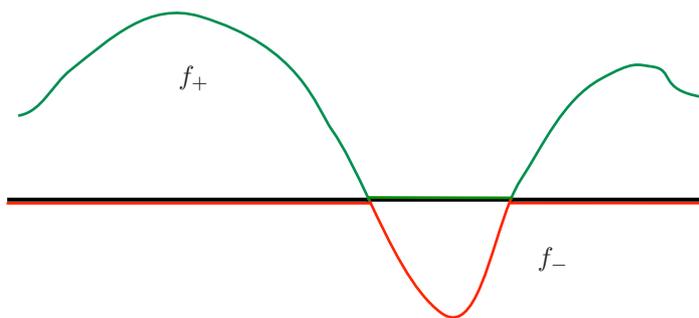
Fazit: Ist  $K \subset \mathbb{R}^n$  kompakt, so existiert eine Folge  $(\varphi_k) \subset \mathcal{C}_c(\mathbb{R}^n)$  mit  $\varphi_k \downarrow \chi_K$ . Dies erlaubt uns die erste, schon recht weitreichende, Fortsetzung des Integrals gemäss (22.41).

**Definition 22.17.** Sei  $f \in \mathcal{C}(\mathbb{R}^n)$ ,  $f \geq 0$ , und  $K \subset \mathbb{R}^n$  kompakt. Wähle eine Folge  $(\varphi_k) \subset \mathcal{C}_c(\mathbb{R}^n)$  mit  $\varphi_k \downarrow \chi_K$ , und setze

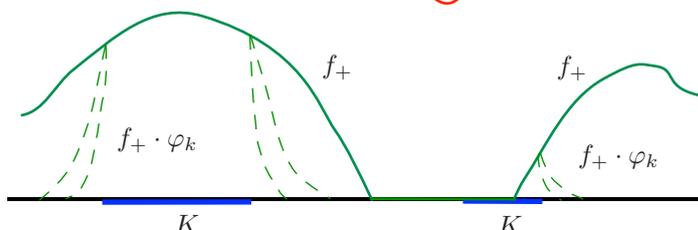
$$\int_K f(x) dx = \lim_{k \rightarrow \infty} \int_{\mathbb{R}^n} f(x) \varphi_k(x) dx \quad (22.79)$$

We im obigen Beispiel verallgemeinern wir dies leicht auf  $f \leq 0$ , und auf allgemeines  $f = f_+ + f_-$  durch Zerlegung in “positiven” und “negativen” Teil, bildlich:

Zerlegung einer stetigen Funktion in positiven und negativen Teil:



Approximation der charakteristischen Funktion des Integrationsbereichs durch punktweise monotone Konvergenz:



Beachte: Offenbar gilt  $f \cdot \varphi_k \downarrow f \cdot \chi_K$ , und die  $f \cdot \varphi_k \in \mathcal{C}_c(\mathbb{R}^n)$ . Daher ist die Folge  $\int_{\mathbb{R}^n} f(x)\varphi_k(x)dx$  monoton fallend (und nach unten beschränkt durch 0), also konvergent. Der Grenzwert hängt nicht von der Wahl der  $\varphi_k$  ab. (Diese Erkenntnis ist wichtig für §24, wir geben hier einen Vorgeschmack.)

**Lemma 22.18.** Ist  $(f_l) \subset \mathcal{C}_c(\mathbb{R}^n)$  eine beliebige Folge mit  $f_l \downarrow f \cdot \chi_K$ , so gilt

$$\lim_{l \rightarrow \infty} \int_{\mathbb{R}^n} f_l(x)dx = \lim_{k \rightarrow \infty} \int_{\mathbb{R}^n} f(x)\varphi_k(x)dx \quad (22.80)$$

(Insbesondere gilt dies für eine Folge der Form  $f_l = f \cdot \psi_l$  für  $\psi_l \downarrow \chi_K$ .)

*Beweis.* Offenbar ist  $\forall k, l \max(f_l, f\varphi_k) \in \mathcal{C}_c(\mathbb{R}^n)$ , und  $\geq f\varphi_k$ . Aus der Monotonie des Integrals folgt (wir unterdrücken das Argument im Integranden)

$$\int_{\mathbb{R}^n} \max(f_l, f\varphi_k)dx \geq \int_{\mathbb{R}^n} f\varphi_k dx \quad (22.81)$$

Für festes  $l$  gilt ausserdem

$$\max(f_l, f\varphi_k) \downarrow f_l \in \mathcal{C}_c(\mathbb{R}^n) \quad \text{für } k \rightarrow \infty \quad (22.82)$$

(Denn  $\forall x$  ist  $f_l(x) \geq f(x)\chi_K(x) = \lim_{k \rightarrow \infty} f(x)\varphi_k(x)$ .) Nach dem Satz von Dini (Kor. 22.15) gilt also

$$\lim_{k \rightarrow \infty} \int_{\mathbb{R}^n} \max(f_l, f\varphi_k)dx = \int_{\mathbb{R}^n} f_l dx \quad (22.83)$$

Aus (22.81) und (22.83) folgt

$$\int_{\mathbb{R}^n} f_l dx \geq \lim_{k \rightarrow \infty} \int_{\mathbb{R}^n} f\varphi_k dx \quad (22.84)$$

und daher auch

$$\lim_{l \rightarrow \infty} \int_{\mathbb{R}^n} f_l dx \geq \lim_{k \rightarrow \infty} \int_{\mathbb{R}^n} f\varphi_k dx \quad (22.85)$$

nach elementaren Rechenregeln für Folgen. Ebenso zeigt man die umgekehrte Ungleichung, zusammen folgt Gleichheit.  $\square$

## § 22. INTEGRATION IM $\mathbb{R}^n$

Mit ähnlichen Betrachtungen (s. auch Details in §24) zeigt man dann die folgenden Eigenschaften des Integrals stetiger Funktionen über kompakte Mengen.

**Proposition 22.19.** (i)  $\int_K f(x)dx$  ist linear in  $f$ .

(ii) Sind  $f, g \in \mathcal{C}(\mathbb{R}^n)$  mit  $f(x) \geq g(x), \forall x \in K \subset \mathbb{R}^n$  kompakt, so gilt

$$\int_K f(x)dx \geq \int_K g(x)dx \quad (22.86)$$

*Beweis.* Wird nachgereicht. □

**Korollar 22.20.** (i) Ist  $f \in \mathcal{C}(\mathbb{R}^n)$  und  $K$  kompakt, so gilt

$$\left| \int_K f(x)dx \right| \leq \int_K |f(x)|dx \leq \|f\|_K \cdot \text{vol}(K) \quad (22.87)$$

wobei  $\text{vol}(K) = \int_K 1 dx$ .

(ii) Ist  $f \geq 0$  und  $K_1, K_2$  kompakt, so gilt

$$\int_{K_1} f(x)dx + \int_{K_2} f(x)dx \geq \int_{K_1 \cup K_2} f(x)dx \quad (22.88)$$

(iii) Für  $f \in \mathcal{C}_c(\mathbb{R}^n)$  ist

$$\int_{\mathbb{R}^n} f(x)dx = \int_{\text{supp}(f)} f(x)dx \quad (22.89)$$

*Beweis.* (i) folgt mit 22.19 aus  $f(x) \leq |f(x)| \leq \|f\|_\infty \forall x \in K$ , (ii) aus  $\chi_{K_1} + \chi_{K_2} \geq \chi_{K_1 \cup K_2}$ . Details werden nachgereicht. □

### Transformationsformel

Eines der wichtigsten Hilfsmittel zur Berechnung von ein-dimensionalen bestimmten Integralen war die *Substitutionsregel*, welche wir in (14.21) 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  $\alpha, \beta \in J$ :

$$\int_{\tau(\alpha)}^{\tau(\beta)} f(s)ds = \int_\alpha^\beta f(\tau(\sigma))\tau'(\sigma)d\sigma \quad (22.90)$$

In den Übungen haben wir gesehen, dass bei Iteration dieser Formel auf das  $n$ -dimensionale Integral 22.4 über Quader  $\tau'(\sigma)$  durch die *Funktionaldeterminante* (Determinante der Jacobi-Matrix) der Substitution zu ersetzen ist. Diese Erkenntnis lässt sich ohne grosse Schwierigkeiten zum Beweis des Stokesschen Integralsatzes ausbauen (wird nachgeliefert). Diese Formulierung der höher-dimensionalen Substitutionsregeln ist allerdings unüblich, und intuitiv auch nicht verträglich mit der Mittelwertinterpretation des Integrals: Wie schon bemerkt gilt die Formel (22.90) auch, wenn  $\tau$  nicht injektiv ist, oder wenn das Bild des Intervalls  $[\alpha, \beta]$  über  $[\tau(\alpha), \tau(\beta)]$  hinausreicht. In solchen Fällen müssen sich die “mehrfach gezählten” Werte aufheben, oder die Berechnung der rechten Seite erfordert die Kenntnis von

Werten von  $f$ , die zum Mittelwert über  $[\tau(\alpha), \tau(\beta)]$  gar nicht beitragen, man mittelt also nicht über die Intervalle “als Mengen”.

· Um noch etwas schärfer zu sehen, was dies mit unserer Entscheidung zu tun hat, die Monotonie (22.8) dem Orientierungswechsel (s. 14.6) vorzuziehen, sei etwa  $f \in \mathcal{C}_c(\mathbb{R})$  mit  $\text{supp}(f) \subset [a, b]$  ( $b > a!$ ). Dann gilt nach 22.13 und (22.6).

$$\int_{\mathbb{R}} f(x) dx = \int_a^b f(s) ds \tag{22.91}$$

Mit der “Substitution”  $s = \lambda\sigma$  für  $\sigma \in \mathbb{R}$  und  $\lambda \in \mathbb{R}$  können wir dies nur dann als

$$\int_a^b f(s) ds = \lambda \int_{\alpha}^{\beta} f(\lambda\sigma) d\sigma \tag{22.92}$$

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

$$\int_a^b f(s) ds = \lambda \int_{a/\lambda}^{b/\lambda} f(\lambda\sigma) d\sigma = \lambda \int_{\mathbb{R}} f(\lambda\xi) d\xi \tag{22.93}$$

Im Fall  $\lambda < 0$  ist andererseits  $\alpha < \beta$  und daher wegen (22.7)

$$\lambda \int_{\alpha}^{\beta} f(\lambda\sigma) d\sigma = -\lambda \int_{\mathbb{R}} f(\lambda\xi) d\xi \tag{22.94}$$

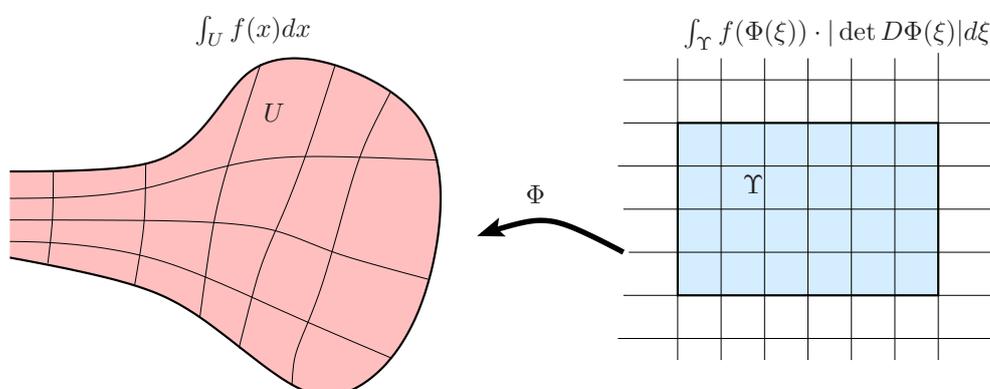
Unser Integral transformiert sich also unter der *invertierbaren* Substitution  $x = \lambda\xi$  in beiden Fällen gemäss

$$\int_{\mathbb{R}} f(x) dx = |\lambda| \int_{\mathbb{R}} f(\lambda\xi) d\xi \tag{22.95}$$

(Anstatt mit dem Hauptsatz der ein-dimensionalen Differential- und Integralrechnung könnte man diese Formel auch direkt über den Grenzprozess 22.10 aus der Definition (22.16) herleiten.) In  $n$  Dimensionen gilt

**Theorem 22.21.** *Es sei  $\Phi : \Upsilon \rightarrow U$  ein Diffeomorphismus zwischen offenen Teilmengen des  $\mathbb{R}^n$ , d.h.  $\Phi$  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  $f \in \mathcal{C}_c(U, \mathbb{R})$  (s. Def. 22.7). Dann ist  $f \circ \Phi \in \mathcal{C}_c(\Upsilon)$  und es gilt*

$$\int_U f(x) dx = \int_{\Upsilon} f(\Phi(\xi)) |\det D\Phi(\xi)| d\xi \tag{22.96}$$



§ 22. INTEGRATION IM  $\mathbb{R}^n$

Bemerkung: Im Unterschied zu der Formel, die wir unter (22.90) angedeutet haben, muss hier  $\Phi$  ein Diffeomorphismus sein, da der Absolutbetrag über der Funktionaldeterminante die Aufhebung der nicht injektiven Beiträge verhindert. Dafür gilt die Formel aber ohne Einschränkung an das Verhalten von  $\Phi$  am Rand oder ausserhalb des Integrationsbereichs.

Interpretation: Der Diffeomorphismus  $\Phi$  macht die  $\xi$  zu “krummlinigen Koordinaten”, wie wir sie im Anschluss an den Satz über die Umkehrabbildung diskutiert hatten (s. Thm. 13.1 und Beispiel 13.3). Stellt man sich  $\int_U f(x)dx$  vor als Mittelwert  $\sum_{K=1}^R f(x_K) \text{vol}(Q_K)$  der mit dem euklidischen Volumen von Quadern gewichteten Funktionswerte im Limes  $R \rightarrow \infty$  (s. Lemma 22.5), so berechnet die Transformationsformel dieses Integral durch Überdeckung von  $U$  mit Bildern von Quadern im  $\mathbb{R}_\xi^n$  unter  $\Phi$ , bzw. dessen affin-lineare Approximation  $D\Phi(\xi)$ , und der Faktor  $|\det D\Phi(\xi)|$  gibt das Volumen eines solchen “infinitesimalen  $\Upsilon$ -Quaders”, gemessen im “euklidischen”  $\mathbb{R}_x^n$ .

Beweisstrategie:<sup>29</sup> Mit Hilfe von (22.95) und anderen bekannten Eigenschaften des ein-dimensionalen Integrals begründen wir zunächst die Transformation unter affin-linearen Substitutionen des gesamten  $\mathbb{R}^n$ . Dann schätzen wir die Güte der affin-linearen Approximation an  $\Phi$  lokal ab. Die allgemeine Aussage führen wir durch Karrierung (in  $\Upsilon!$ ) ähnlich wie im Beweis von 22.5 auf den lokalen Fall zurück.

Vorbereitungen: Zur Abkürzung schreiben wir den Integranden auf der rechten Seite von (22.96) als

$$f^\Phi(\xi) := f(\Phi(\xi)) \cdot |\det D\Phi(\xi)| \quad (22.97)$$

Offensichtlich ist  $\text{supp}(f^\Phi) = \Phi^{-1}(\text{supp}(f))$ , und daher wegen der Stetigkeit von  $\Phi^{-1}$  kompakt (stetige Bilder von kompakten Mengen sind kompakt, s. Übungen). Gemäss Def. 22.7, Gl. (22.40) setzen wir dann  $f^\Phi$  zu einer stetigen Funktion auf ganz  $\mathbb{R}_\xi^n$  fort. (Und zwar ohne die Notation zu ändern, und auch wenn  $\Phi$  selbst ist nur auf  $\Upsilon$  definiert / invertierbar ist.)

· Wir halten weiterhin 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 Determinante auch für  $\Phi \circ \Psi$  gilt: Aus

$$|\det D(\Phi \circ \Psi)(\mathfrak{x})| = |\det D\Phi(\Psi(\mathfrak{x}))| \cdot |\det D\Psi(\mathfrak{x})| \quad (22.98)$$

folgt  $(f^\Phi)^\Psi = f^{\Phi \circ \Psi}$  für alle  $f$ . Wir benutzen dies vor allem im 1. Beweisschritt (affin-lineare Substitutionen), aber auch für  $\Psi = \Phi^{-1}$ , für die

$$(f^\Phi)^{\Phi^{-1}} = f \quad (22.99)$$

· Wie in 22.13 und bei der Def. 22.17 genügt es, die Aussage für  $f \geq 0$  zu zeigen.

1. Schritt: Wir zeigen die Aussage für  $U = \mathbb{R}^n$ ,  $\Upsilon = \mathbb{R}_\xi^n$ , und  $\Phi$  eine affine Abbildung, d.h.  $f \in \mathcal{C}_c(\mathbb{R}^n)$  und mit einer invertierbaren Matrix  $A \in GL(n, \mathbb{R})$  und  $b \in \mathbb{R}^n$  gilt

$$\Phi(\xi) = A\xi + b, \quad (22.100)$$

Beachte, dass in diesem Fall  $D\Phi = A$ .

---

<sup>29</sup>Nicht unähnlich zu Lemma 16.1.

· Bekanntlich (s. Übungen) bilden die affinen Abbildungen des  $\mathbb{R}^n$  eine Gruppe, die erzeugt wird von den Translationen und den invertierbaren linearen Abbildungen. Letztere wiederum lassen sich schreiben als endliches Produkt sog. Elementarmatrizen. Wegen 22.98 genügt es also, die folgenden Fälle getrennt voneinander zu untersuchen:

- (i)  $A = \text{id}$ , beliebiges  $b$  (Translation)
- (ii)  $b = 0$ ,  $A$  eine Transposition (Vertauschen zweier Koordinaten).
- (iii)  $b = 0$ ,  $A$  ist diagonal mit genau einem Eintrag  $\neq 1$  (Skalieren einer Koordinate).
- (iv)  $b = 0$ ,  $A - \text{id}$  hat genau einen nicht-verschwindenden, ausser-diagonalen Eintrag (Addieren einer Koordinate zu einer anderen).

Beachte, dass die Determinante bei den Typen (i) und (iv) gleich 1 ist, und bei (ii) gleich  $-1$ , so dass also das gesamte Gewicht von (iii) stammt, der Skalierung der Koordinaten (den ‘Eigenwerte von  $D\Phi$ ’). Zur Begründung ist es nützlich festzuhalten, dass sich für stetige Funktionen mit kompaktem Träger das Integral über  $\mathbb{R}^n$  (wie in (22.7) definiert) auch rekursiv schreiben lässt und 22.6 sinngemäss weiter gilt.

· Nun folgt (i) sofort aus der Translationsinvarianz des ein-dimensionalen Integrals über ganz  $\mathbb{R}$ , d.h.  $\int_{\mathbb{R}} f(x + b)dx = \int_{\mathbb{R}} f(x) \forall b \in \mathbb{R}$ . (ii) ist nichts anderes als das Vertauschen zweier Koordinaten in der Integrationsreihenfolge. (iii) folgt aus (22.95).

· Zum Nachweis von (iv) bringen wir mit Hilfe von (ii) die betroffene Koordinate an erste Stelle und benutzen dann die Translationsinvarianz des ein-dimensionalen Integrals. Beispiel ( $n = 2$ ):  $(x^1, x^2) = \Phi(\xi^1, \xi^2) = (\xi^1 + 2\xi^2, \xi^2)$ ,

$$\begin{aligned} \int_{\mathbb{R}^2} f(\xi^1 + 2\xi^2, \xi^2) d\xi^1 d\xi^2 &= \int_{\mathbb{R}} \left( \int_{\mathbb{R}} f(\xi^1 + 2\xi^2, \xi^2) d\xi^1 \right) d\xi^2 \\ &= \int_{\mathbb{R}} \left( \int_{\mathbb{R}} f(\xi^1, \xi^2) d\xi^1 \right) d\xi^2 = \int_{\mathbb{R}^2} f(x^1, x^2) dx \end{aligned} \tag{22.101}$$

□<sub>1. Schritt</sub>

2. Schritt: Sei weiterhin  $U = \mathbb{R}^n$ ,  $\Upsilon = \mathbb{R}_{\xi}^n$ , und  $\Phi$  eine affine Abbildung, ausserdem  $f \in \mathcal{C}(\mathbb{R}^n)$  und  $K \subset \mathbb{R}^n$  kompakt. Dann gilt

$$\int_K f(x) dx = \int_{\Phi^{-1}(K)} f^{\Phi}(\xi) d\xi \tag{22.102}$$

Beweis: Ist  $(\varphi_k) \subset \mathcal{C}_c(\mathbb{R}^n)$  mit  $\varphi_k \downarrow \chi_K$  so gilt  $\varphi_k \circ \Phi \downarrow \chi_{\Phi^{-1}(K)}$  und daher per Definition 22.17 und Lemma 22.18

$$\begin{aligned} \int_K f(x) dx &= \lim_{k \rightarrow \infty} \int_{\mathbb{R}^n} f(x) \varphi_k(x) dx \\ \text{(1. Schritt)} &= \lim_{k \rightarrow \infty} \int_{\mathbb{R}_{\xi}^n} f^{\Phi}(\xi) \varphi_k(\Phi(\xi)) d\xi = \int_{\Phi^{-1}(K)} f^{\Phi}(\xi) d\xi \end{aligned} \tag{22.103}$$

□<sub>2. Schritt</sub>

3. Schritt: Sei  $\Phi : \Upsilon \rightarrow U$  ein beliebiger Diffeomorphismus, und  $f \in \mathcal{C}_c(U)$ ,  $f \geq 0$ . Dann gilt:

$$\int_{\Upsilon} f^{\Phi}(\xi) d\xi \geq \int_U f(x) dx \tag{22.104}$$

§ 22. INTEGRATION IM  $\mathbb{R}^n$

Beweis von (22.104): (indirekt, und lang) Angenommen, die linke Seite ist kleiner als die rechte. Dann existiert ein  $\kappa > 0$  so dass (sogar noch)

$$\underbrace{\int_{Q_0} f^\Phi(\xi) d\xi}_{=\int_{\Gamma} f^\Phi(\xi) d\xi} < \underbrace{\int_{\Phi(Q_0 \cap \text{supp}(f^\Phi))} f(x) dx}_{=\int_U f(x) dx} - \kappa \quad (22.105)$$

für einen geeigneten Quader  $Q_0 \supset \text{supp}(f^\Phi)$  (s. Def. 22.7). (Auf der rechten Seite ist natürlich  $Q_0 \cap \text{supp}(f^\Phi) = \text{supp}(f^\Phi)$ . Der Schnitt mit  $\text{supp}(f^\Phi)$  ist notwendig, damit  $\Phi$  wohldefiniert ist, wie sich gleich noch auszahlen wird.) Zur Vereinfachung und ohne Einschränkung nehmen wir  $Q_0$  als ‘‘Würfel’’, d.h. die Kantenlängen sind alle gleich. Karieren wir  $Q_0$  durch Seitenhalbierung in  $2^n$  Würfel  $Q_0^{(j)}$ ,  $j = 1, \dots, 2^n$ , so gilt einerseits wegen (22.30)

$$\int_{Q_0} f^\Phi(\xi) d\xi = \sum_{j=1}^{2^n} \int_{Q_0^{(j)}} f^\Phi(\xi) d\xi \quad (22.106)$$

Andererseits sind die  $\Phi(Q_0^{(j)} \cap \text{supp}(f^\Phi))$  alle kompakt, und es gilt

$$\bigcup_{j=1}^{2^n} \Phi(Q_0^{(j)} \cap \text{supp}(f^\Phi)) \supset \text{supp}(f) \quad (22.107)$$

Daher folgt mit 22.20<sup>30</sup>

$$\int_{\Phi(Q_0 \cap \text{supp}(f^\Phi))} f(x) dx \leq \sum_{j=1}^{2^n} \int_{\Phi(Q_0^{(j)} \cap \text{supp}(f^\Phi))} f(x) dx \quad (22.109)$$

Es muss dann ein  $j$  geben, sodass mit  $Q_1 := Q_0^{(j)}$ :

$$\int_{Q_1} f^\Phi(\xi) d\xi < \int_{\Phi(Q_1 \cap \text{supp}(f^\Phi))} f(x) dx - \frac{\kappa}{2^n} \quad (22.110)$$

(sonst wäre in der Summe (22.105) wieder verletzt). ( $Q_1 \cap \text{supp}(f^\Phi)$  ist nicht-leer, sonst wäre das Integral Null, und die Ungleichung verletzt. Beachte auch  $f \geq 0$ .)

Durch Iteration erhalten wir dann eine Folge  $(Q_k)$  von nicht-leeren Würfeln mit  $Q_k \supset Q_{k+1}$ ,  $\text{diam}(Q_k) = \text{diam}(Q_0)/2^k$  und  $\text{vol}(Q_k) = \text{vol}(Q_0)/2^{nk}$  sowie

$$\int_{Q_k} f^\Phi(\xi) d\xi < \int_{\Phi(Q_k \cap \text{supp}(f^\Phi))} f(x) dx - \frac{\kappa}{2^{nk}} \quad (22.111)$$

---

<sup>30</sup>De facto gilt hier natürlich Gleichheit, weil auch die Bilder der  $Q_0^{(j)}$  unter  $\Phi$  sich nur ‘‘am Rand’’ schneiden. An dieser Stelle ist es aber schwer zu zeigen, dass das ‘‘Integral über die Ränder’’, welches auch über 22.17 definiert ist, nicht beiträgt. Die Topologie reicht nicht aus: Es gibt offene Mengen, deren Rand keine Nullmenge ist. Beispiel:

$$\bigcup_{k=1}^{\infty} (r_k - \frac{\delta}{2^k}, r_k + \frac{\delta}{2^k}) \subset (0, 1) \quad (22.108)$$

wobei  $(r_k)$  eine (geeignete) Abzählung von  $\mathbb{Q} \cap (0, 1)$  ist.  $\bar{U} = [0, 1]$ ,  $\text{vol}(U) \leq \delta \Rightarrow \text{vol}(\partial U) \geq 1 - \delta$ .

für alle  $k$ . Aus dem Schachtelungsprinzip 4.24 folgt die Existenz einer eindeutigen Stelle  $\xi_*$  mit

$$\bigcap_{k=0}^{\infty} Q_k = \{\xi_*\} \quad (22.112)$$

Es muss  $\xi_* \in \text{supp}(f^\Phi) \subset \Upsilon$  gelten (sonst lägen für grosses  $k$  die  $Q_k$  ausserhalb des Trägers, und die Integrale in (22.111) wären eh 0). Dann liegen für grosses  $k$  die  $Q_k$  in  $\Upsilon$  (und wir können in (22.111) das lästige  $\cdot \cap \text{supp}(f^\Phi)$  weglassen).

• Die Idee ist nun, auf den immer kleiner werdenden Würfeln  $\Phi$  durch seine affine Approximation zu ersetzen und dies zum Widerspruch mit (22.102) zu führen. Es bezeichne hierzu:

- für einen (beliebigen) Würfel  $Q$   $\eta$  dessen Mittelpunkt, und  $2l$  seine Kantenlänge, d.h.  $Q = \{\xi \mid \|\xi - \eta\|_\infty \leq l\}$ .
- für eine Zahl  $q > 1$   $Q^{(q)}$  den von  $\eta$  aus um den Faktor  $q$  gestreckten Würfel, in Formeln  $Q^{(q)} = \{\xi \mid \|\xi - \eta\|_\infty \leq ql\}$
- $\Phi_\eta$  die affin-lineare Approximation von  $\Phi$  um  $\eta$ , d.h.

$$\Phi_\eta(\xi) := \Phi(\eta) + D\Phi(\eta)(\xi - \eta) \quad (22.113)$$

**Lemma 22.22.** *Sei  $q > 1$ . Dann existiert ein  $r > 0$  so dass für jeden Würfel  $Q \subset B_r(\xi_*) \subset \Upsilon$ :*

$$\Phi(Q) \subset \Phi_\eta(Q^{(q)}) \quad (22.114)$$

(In Worten: Der Würfel wird unter  $\Phi$  nicht weiter abgebildet als der um  $q$  gestreckte Würfel unter  $\Phi_\eta$ .)

*Beweis.*  $B_r(\xi_*)$  ist die Kugel vom Radius  $r$  um  $\xi_*$ , in der beliebig gewählten Norm  $\|\cdot\|$  (die natürlich auch die  $\infty$ -Norm sein könnte). Die Existenz von  $r$  so dass  $B_r(\xi_*) \subset \Upsilon$  folgt also einfach aus der Offenheit von  $\Upsilon$ .

• Da  $\Phi_\eta$  invertierbar ist, ist die Behauptung äquivalent zu  $\Phi_\eta^{-1}(\Phi(Q)) \subset Q^{(q)}$ , m.a.W.

$$\|\xi - \eta\|_\infty \leq l \Rightarrow \|\Phi_\eta^{-1}(\Phi(\xi)) - \eta\|_\infty \leq ql \quad (22.115)$$

Nun ist aber (für festes  $\eta$ )  $\Phi_\eta^{-1} \circ \Phi$  stetig differenzierbar auf  $\Upsilon$ . Als Funktion von  $(\eta, \xi) \in B_r(\xi_*) \times B_r(\xi_*)$  ist die Ableitung (nur nach  $\xi$ !)  $D(\Phi_\eta^{-1}\Phi)(\xi)$  stetig, und gleich  $\text{id}_{\mathbb{R}^n}$  auf der Diagonalen  $\xi = \eta$ , d.h. insbesondere bei  $\xi = \eta = \xi_*$ . Daher ist für  $r$  klein genug

$$\|D(\Phi_\eta^{-1}\Phi)(\xi)\| < q \quad \forall \xi, \eta \in B_r(\xi_*) \quad (22.116)$$

(Operatornorm bezüglich  $\|\cdot\|_\infty$ !) Aus der Konvexität von  $Q \subset B_r(\xi_*)$  folgt mit Hilfe des Schrankensatzes 10.11, dass

$$\|\Phi_\eta^{-1}(\Phi(\xi)) - \eta\|_\infty = \|\Phi_\eta^{-1}(\Phi(\xi)) - \Phi_\eta^{-1}(\Phi(\eta))\|_\infty \leq q \cdot \|\xi - \eta\|_\infty \quad (22.117)$$

Dies zeigt (22.115). □<sub>Lemma 22.22</sub>

Mit Hilfe des Lemmas 22.22 folgt dann wegen  $f \geq 0$  und 22.20 aus (22.111) für  $k$  gross genug:

$$\int_{Q_k} f^\Phi(\xi) d\xi < \int_{\Phi_{\eta_k}(Q_k^{(q)})} f(x) dx - \frac{\kappa}{2^{nk}} = \int_{Q_k^{(q)}} f^{\Phi_{\eta_k}}(\xi) d\xi - \frac{\kappa}{2^{nk}} \quad (22.118)$$

(Wobei  $\eta_k$  der Mittelpunkt von  $Q_k$  ist, und wir benutzt haben, dass  $Q_k \subset B_r(\xi_*)$  für  $k$  gross genug, und im letzten Schritt (22.102) aus dem 2. Schritt.)

## § 23. INTEGRALSÄTZE

**Lemma 22.23.**

$$\begin{aligned} \lim_{k \rightarrow \infty} 2^{nk} \int_{Q_k} f^\Phi(\xi) d\xi &= f^\Phi(\xi_*) \operatorname{vol}(Q_0) \\ \lim_{k \rightarrow \infty} 2^{nk} \int_{Q_k^{(q)}} f^{\Phi \circ \eta_k}(\xi) d\xi &= q^n f^\Phi(\xi_*) \operatorname{vol}(Q_0) \end{aligned} \quad (22.119)$$

*Beweis.* Wie oben bemerkt ist  $2^{nk} \operatorname{vol}(Q_k) = \operatorname{vol}(Q_0)$  und daher folgt mit den elementaren Eigenschaften der Def. 22.4

$$\begin{aligned} \left| 2^{nk} \int_{Q_k} f^\Phi(\xi) d\xi - f^\Phi(\xi_*) \operatorname{vol}(Q_0) \right| &= \left| 2^{nk} \int_{Q_k} (f^\Phi(\xi) - f^\Phi(\xi_*)) d\xi \right| \\ &\leq 2^{nk} \operatorname{vol}(Q_k) \cdot \|f^\Phi - f^\Phi(\xi_*)\|_{Q_k} \\ &= \operatorname{vol}(Q_0) \cdot \|f^\Phi - f^\Phi(\xi_*)\|_{Q_k} \end{aligned} \quad (22.120)$$

Aus der Stetigkeit von  $f^\Phi$  und  $\xi_* \in Q_k \forall k$  mit  $\operatorname{diam}(Q_k) \rightarrow 0$  für  $k \rightarrow \infty$  folgt  $\lim_{k \rightarrow \infty} \|f^\Phi - f^\Phi(\xi_*)\|_{Q_k} = 0$  und damit die erste Behauptung. Für den zweiten Grenzwert benutzt man  $\operatorname{vol} Q^{(q)} = q^n \operatorname{vol} Q$  und die Stetigkeit von  $f^{\Phi \circ \eta}(\xi)$  als Funktion von  $\eta$  und  $\xi$ . □<sub>Lemma 22.23</sub>

Setzen wir die Grenzwerte des Lemmas in (22.118) ein, so folgt

$$f^\Phi(\xi_*) \cdot \operatorname{vol}(Q_0) \leq q^n f^\Phi(\xi_*) \cdot \operatorname{vol}(Q_0) - \kappa \quad (22.121)$$

für alle  $q > 1$  und das  $\kappa > 0$  aus (der Widerspruchsvoraussetzung) (22.105). Für  $q - 1$  klein genug ist das offenbar Unsinn. □<sub>3. Schritt</sub>

Wenden wir (22.104) an auf  $\Phi^{-1} : U \rightarrow \Upsilon$  so folgt mit (22.99) die Ungleichungskette

$$\int_U f(x) dx = \int_U f^\Phi \circ \Phi^{-1}(x) dx \geq \int_\Upsilon f^\Phi(\xi) d\xi \geq \int_U f(x) dx \quad (22.122)$$

was nur bei Gleichheit möglich ist. □<sub>22.21</sub>

Wie im Beweis des 2. Schritts zeigt man dann leicht, was vielleicht die praktischste Version der Transformationsformel ist.

**Korollar 22.24.** Sei  $\Phi : \Upsilon \rightarrow U$  ein Diffeomorphismus zwischen offenen Teilmengen des  $\mathbb{R}^n$ . Dann gilt für alle  $f \in C(U)$  und alle kompakten Teilmengen  $K \subset U$

$$\int_K f(x) dx = \int_{\Phi^{-1}(K)} f(\Phi(\xi)) |\det D\Phi(\xi)| d\xi \quad (22.123)$$

□

## § 23 Integralsätze

Das bis hierher entwickelte Integral (stetiger Funktionen mit kompaktem Träger in offenen Mengen bzw. stetiger Funktionen über kompakte Mengen) reicht aus zur Aufstellung und Begründung der (aus den Physik-Vorlesungen in Einzelfällen bereits

vertrauten) Integralsätze. Diese verallgemeinern den ein-dimensionalen Hauptsatz auf höhere Dimensionen und erlauben (unter gewissen Voraussetzungen) die Auswertung eines Integrals über eine Teilmenge des  $\mathbb{R}^n$  durch ein (anderes) Integral über ihren "Rand". Zu ihrer allgemeinen Formulierung müssen wir

- (i) uns an die Orientierung erinnern, die wir in (22.6) (fast) vergessen hatten.
- (ii) klären, was unter einer "Stammfunktion" des Integranden zu verstehen ist, und wie wir das Integral über den Rand auswerten.
- (iii) hierzu eine Reihe von differentialgeometrischen Begriffen einführen.

Wir beginnen wieder mit der einfachsten Situation (auch um die Notation festzulegen), und benutzen als Schlüssel für die Motivation von Differentialformen etc. vor allem die Transformationsformel für das  $n$ -dimensionale Integral.

### Der Gaussche Integralsatz

· Aus der iterativen Definition 22.4 folgt in Verbindung mit dem ein-dimensionalen Hauptsatz sofort:

**Proposition 23.1.** Sei  $Q = \times_{j=1}^n [a_j, b_j] \subset \mathbb{R}^n$  und  $f \in C(Q)$ . Ist  $F$  eine "Stammfunktion bezüglich der ersten Koordinate", d.h.  $F$  ist in einer Umgebung von  $Q$  partiell nach  $x^1$  differenzierbar und  $\partial_1 F(x) = \frac{\partial}{\partial x^1} F(x) = f(x)$ , so gilt

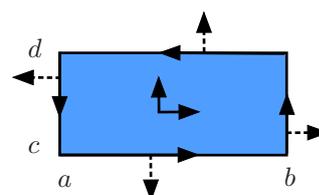
$$\int_Q f(x) d^n x = \int_{Q \cap \{x^1=b_1\}} F(b_1, x^2, \dots, x^n) d^{n-1} x - \int_{Q \cap \{x^1=a_1\}} F(a_1, x^2, \dots, x^n) d^{n-1} x \tag{23.1}$$

Analog falls  $f = \partial_i F$ , und allgemein für  $n$  Funktionen  $F^{(i)}$  mit  $\sum_i \partial_i F^{(i)} = f$

$$\int_Q f(x) d^n x = \sum_{i=1}^n \left( \int_{Q \cap \{x^i=b_i\}} F^{(i)} d^{n-1} x - \int_{Q \cap \{x^i=a_i\}} F^{(i)} d^{n-1} x \right) \tag{23.2}$$

Interpretation/Vereinbarung: Die rechte Seite von (23.2) ist offenbar das  $n - 1$ -dimensionale Integral über die  $2n$  Seitenflächen des Quaders (welche in diesem Fall ebenfalls Quader sind). Wir zeigen die Dimensionalität (aber nicht systematisch) im Differential  $d^n x := dx = dx^1 \cdots dx^n$  etc. an.

Die Vorzeichen zeigen an, dass die Orientierung nicht immer diejenige ist, die das Integral über diese Teile des Randes als Menge definiert. Betrachte speziell  $n = 2$ :



$$\begin{aligned} \int_Q f(x) dx &= \\ &= \int_{[c,d]} F^{(1)}(b, x^2) dx^2 - \int_{[c,d]} F^{(1)}(a, x^2) dx^2 + \int_{[a,b]} F^{(2)}(x^1, d) dx^1 - \int_{[a,b]} F^{(2)}(x^1, c) dx^1 \\ &= \int_c^d F^{(1)}(b, x^2) dx^2 - \int_c^d F^{(1)}(a, x^2) dx^2 + \int_a^b F^{(2)}(x^1, d) dx^1 - \int_a^b F^{(2)}(x^1, c) dx^1 \\ &= \int_a^b (-F^{(2)}(x^1, c)) dx^1 + \int_c^d F^{(1)}(b, x^2) dx^2 + \int_b^a (-F^{(2)}(x^1, d)) dx^1 + \int_d^c F^{(1)}(a, x^2) dx^2 \end{aligned} \tag{23.3}$$

### § 23. INTEGRALSÄTZE

· Als nächstes betrachten wir eine Situation, die wir mit Hilfe der Transformationsformel auf diesen Fall zurückführen können: Sei  $\Phi : \Upsilon \rightarrow U$  ein Diffeomorphismus zwischen offenen Mengen des  $\mathbb{R}^n$ , und  $K \subset U$  eine kompakte Menge mit der Eigenschaft, dass  $\Phi^{-1}(K) = Q = \times_{\alpha=1}^n [\gamma_\alpha, \delta_\alpha]$  ein Quader in  $\mathbb{R}_\xi^n$  ist. Angenommen  $f \in \mathcal{C}^1(U)$ , und  $F$  ist eine "Stammfunktion bezüglich  $x^1$ ", d.h.  $f = \partial_1 F$ . (Die Verallgemeinerung für  $f = \partial_i F^{(i)}$  wird offensichtlich sein.) Dann können wir die Berechnung des Integrals von  $f$  über  $K$  wie folgt auf ein Integral über den Rand zurückführen.

· Wir schreiben gemäss 22.24:

$$\int_K f(x) dx = \int_Q f(\Phi(\xi)) |\det D\Phi(\xi)| d\xi \quad (23.4)$$

· Wir stellen fest, dass das Vorzeichen der Determinante auf ganz  $Q$  konstant sein muss. Das hinreichende Kriterium hierfür ist, dass  $Q$  *zusammenhängend* ist. Zur Not folgt dies mit dem Zwischenwertsatz 7.1 daraus, dass  $Q$  konvex ist. (Wäre  $\det$  an einer Stelle positiv, an einer anderen negativ, so müsste sie an einer Stelle auf der Verbindungsstrecke verschwinden.)

· Wir entwickeln die Determinante nach der ersten Zeile,

$$\det D\Phi = \det(\partial_\beta \Phi^j)_{\substack{j=1,\dots,n \\ \beta=1,\dots,n}} = \sum_{\alpha=1}^n (-1)^{\alpha-1} \partial_\alpha \Phi^1 \cdot \det(\partial_\beta \Phi^j)_{\substack{j \neq 1 \\ \beta \neq \alpha}} \quad (23.5)$$

· Wir multiplizieren mit  $f = \partial_1 F$  und benutzen die Produktregel ( $f \in \mathcal{C}^1 \Rightarrow F$  ist partiell differenzierbar):

$$\begin{aligned} f(\Phi(\xi)) |\det D\Phi(\xi)| &= (\partial_1 F)(\Phi(\xi)) \sum_{\alpha} (-1)^{\alpha-1} \partial_\alpha \Phi^1 \det(\partial_\beta \Phi^j)_{\substack{j \neq 1 \\ \beta \neq \alpha}} \\ &= \sum_{\alpha} (-1)^{\alpha-1} \partial_\alpha [F(\Phi(\xi)) \det(\partial_\beta \Phi^j)_{\substack{j \neq 1 \\ \beta \neq \alpha}}] \\ &\quad - \sum_{\alpha} (-1)^{\alpha-1} \left[ \sum_{i=2}^n \partial_i F(\Phi(\xi)) \partial_\alpha \Phi^i (\det(\partial_\beta \Phi^j)_{\substack{j \neq 1 \\ \beta \neq \alpha}}) + F(\Phi(\xi)) \partial_\alpha (\det(\partial_\beta \Phi^j)_{\substack{j \neq 1 \\ \beta \neq \alpha}}) \right] \end{aligned} \quad (23.6)$$

· Wir verifizieren, dass der zweite Term verschwindet:

Beh.: Ist  $\Phi : \Upsilon \rightarrow U$  zwei Mal stetig differenzierbar, so gilt

$$\sum_{\alpha} (-1)^{\alpha-1} \partial_\alpha (\det(\partial_\beta \Phi^j)_{\substack{j \neq 1 \\ \beta \neq \alpha}}) = 0 \quad (23.7)$$

Bew.: Der Vergleich der Leibniz-Formel für die Determinante einer  $n \times n$ -Matrix  $(A_i^j)$  mit ihrer Laplace-Entwicklung nach der ersten Zeile,

$$\det(A_\alpha^j) = \sum_{\sigma \in S_n} \text{sgn}(\sigma) \prod_j A_{\sigma(j)}^j = \sum_{i=1}^n (-1)^{\alpha-1} A_\alpha^1 \det(A_\beta^j)_{\substack{j \neq 1 \\ \beta \neq \alpha}} \quad (23.8)$$

zeigt, zusammen mit der Produktregel

$$\sum_{\alpha} (-1)^{\alpha-1} \partial_\alpha (\det(\partial_\beta \Phi^j)_{\substack{j \neq 1 \\ \beta \neq \alpha}}) = \sum_{i=2}^n \sum_{\sigma \in S_n} \text{sgn}(\sigma) \partial_{\sigma(1)\sigma(i)} \Phi^i \prod_{j \neq i} \partial_{\sigma(j)} \Phi^j = 0 \quad (23.9)$$

wegen der Symmetrie der zweiten Ableitung und der Antisymmetrie der Determinante. (Ersetze  $\sigma$  durch  $\sigma \circ$  (Vertauschen von 1 und  $i$ ).)  $\square$

· Aus ähnlichen Gründen verschwindet auch  $\sum_{\alpha} (-1)^{\alpha-1} \partial_{\alpha} \Phi^i (\det(\partial_{\beta} \Phi^j)_{\substack{j \neq 1 \\ \beta \neq \alpha}})$  falls  $i \geq 2$ . Damit folgt

**Proposition 23.2.** *Unter den obigen Voraussetzungen gilt*

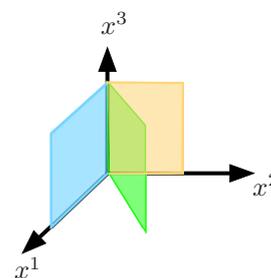
$$\int_K f(x) dx = \sum_{\alpha} (-1)^{\alpha-1} \left( \int_{Q \cap \xi^{\alpha} = \delta_{\alpha}} - \int_{Q \cap \xi^{\alpha} = \gamma_{\alpha}} \right) F(\Phi(\xi)) \det(\partial_{\beta} \Phi^j)_{\substack{j \neq 1 \\ \beta \neq \alpha}} d^{n-1} \xi \quad (23.10)$$

Interpretation: Wir hatten im vorigen § festgehalten, dass die Funktionaldeterminante in (23.4) das  $n$ -dimensionale euklidische Volumen eines infinitesimalen Quaders im  $\mathbb{R}_{\xi}^n$  angibt, welcher mit Hilfe von  $D\Phi$  in den  $\mathbb{R}_x^n$  abgebildet, und dort gemessen wird. In ähnlicher Weise können wir die Determinante  $(\partial_{\beta} \Phi^j)_{\substack{j \neq 1 \\ \beta \neq \alpha}}$  identifizieren als “orientiertes Volumen” eines  $n - 1$ -dimensionalen Quaders in  $\mathbb{R}_{\alpha}^{n-1} = \mathbb{R}_{\xi}^n \cap \{\xi^{\alpha} = \text{const.}\}$ , welcher mit  $D\Phi$  in den  $\mathbb{R}_x^n$  abgebildet wird, und dort auf die Hyperebene  $x^1 = \text{const.}$  projiziert wird. (Im Allgemeinen ist dies *nicht* gleich dem euklidischen Hyperflächeninhalt! “Projektion” bedeutet, dass man auf der Ebene der Basisvektoren  $(\partial_{\beta} \Phi^j)$  den ersten Eintrag  $j = 1$  vergisst.) Unsere “Stammfunktion”  $F$  wird also mit diesem Volumen gewichtet über den Rand von  $K$  (dem Bild des Randes von  $Q$  unter  $\Phi$ ) integriert. Die Vorzeichen in (23.10) reflektieren hierbei die natürliche Orientierung der jeweiligen Randstücke von  $K$ .

Im Falle  $f = \partial_i F^{(i)}$  haben wir allgemeiner

$$\sum_i F^{(i)} (-1)^{i-1} \det(\partial_{\beta} \Phi^j)_{\substack{j \neq i \\ \beta \neq \alpha}} \quad (23.11)$$

über den Rand von  $Q$  zu integrieren. Das heisst, die  $F^{(i)}$  gewichten (mit einem weiteren Vorzeichen) die *Projektion des Quaders auf die Hyperfläche  $x^i = \text{const.}$*



Im physikalisch einleuchtendsten Fall ist  $f$  die zeitliche Ableitung einer Dichte,  $F^{(i)}$  die lokale Flussdichte durch die Hyperebene  $x^i = \text{const.}$ , und der Gaussische Integralsatz ein Erhaltungssatz.

Fazit: Als Stammfunktion von  $f$  ist anzusehen eine Funktion, die an jeder Stelle  $x \in \mathbb{R}^n$  jedem infinitesimalen  $n - 1$ -dimensionalen Hyperflächenstück eine Zahl zuordnet. Ein solches Objekt heisst Differentialform (vom Grad  $n - 1$ ). Das Integral von  $F$  über den Rand von  $K$  ist (jedenfalls zunächst) definiert, und prinzipiell berechenbar, über die durch  $\Phi$  gegebene Parametrisierung.

· Wir stellen nun die für die allgemeine Formulierung des Stokesschen Satzes benötigten (differential-)geometrischen Begriffe zusammen.

### Alternierende Multilinearformen ...

... dienen zur Buchhaltung der Vorzeichen und der Determinanten wie sie in (23.4), (23.10), und (23.11) auftauchen. Wir wiedergeben ihre Definition und wichtigsten Eigenschaften von S. 168 ff.

§ 23. INTEGRALSÄTZE

**Definition 23.3.** Sei  $V$  ein  $n$ -dimensionaler Vektorraum über  $\mathbb{K}$  ( $= \mathbb{R}$  oder  $\mathbb{C}$ ). Eine *alternierende Multilinearform vom Grad  $k$*  ( $k = 0, 1, \dots$ ), kurz “ $k$ -Form” ist eine  $k$ -lineare Abbildung

$$\omega : \underbrace{V \times V \times \dots \times V}_{k \text{ Mal}} \rightarrow \mathbb{K} \quad (23.12)$$

mit der Eigenschaft, dass  $\forall \sigma \in S_k$  (der symmetrischen Gruppe, s. 20.2)

$$\omega(v_{\sigma(1)}, \dots, v_{\sigma(k)}) = \text{sgn}(\sigma) \omega(v_1, \dots, v_k) \quad (23.13)$$

**Beispiel 23.4.** Sind  $\lambda_1, \dots, \lambda_k \in V^\vee$ , so ist

$$\lambda_1 \wedge \dots \wedge \lambda_k := \sum_{\sigma \in S_k} \text{sgn}(\sigma) \lambda_{\sigma(1)} \otimes \dots \otimes \lambda_{\sigma(k)} \quad (23.14)$$

eine  $k$ -Form, wobei wir die Identifikation

$$L(V, \dots, V, \mathbb{K}) = (V \otimes \dots \otimes V)^\vee = V^\vee \otimes \dots \otimes V^\vee \quad (23.15)$$

benutzt haben, s. Gl. (19.64)). (Mit anderen Worten ist für  $v_1, \dots, v_k \in V$ :

$$\begin{aligned} (\lambda_1 \wedge \dots \wedge \lambda_k)(v_1, \dots, v_k) &:= \sum_{\sigma \in S_k} \text{sgn}(\sigma) \lambda_1(v_{\sigma(1)}) \cdot \lambda_2(v_{\sigma(2)}) \cdot \dots \cdot \lambda_k(v_{\sigma(k)}) \\ &= \det(\lambda_i(v_j))_{\substack{i=1, \dots, k \\ j=1, \dots, k}} \end{aligned} \quad (23.16)$$

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

**Fakten 23.5.** Die Menge  $\wedge^k V^\vee$  der  $k$ -Formen ist ein  $\binom{n}{k}$ -dimensionaler Vektorraum. Ist  $(e_1, \dots, e_n)$  eine Basis von  $V$ , und  $(\epsilon^1, \dots, \epsilon^n)$  die zugehörige Dualbasis von  $V^\vee$ , so ist

$$\{\epsilon^{i_1} \wedge \dots \wedge \epsilon^{i_k} \mid 1 \leq i_1 < \dots < i_k \leq n\} \quad (23.17)$$

eine Basis von  $\wedge^k V^\vee$ . Insbesondere sind  $\wedge^0 V^\vee$  und  $\wedge^n V^\vee$  ein-dimensional und  $\wedge^k V^\vee = \{0\}$  für  $k > n$ .

**Definition 23.6.** Für beliebige  $k, l \geq 1$  ist das *Dachprodukt* (gesprochen “wedge”)

$$\begin{aligned} \cdot \wedge \cdot : \wedge^k V^\vee \times \wedge^l V^\vee &\rightarrow \wedge^{k+l} V^\vee \\ (\omega, \alpha) &\mapsto \omega \wedge \alpha \end{aligned} \quad (23.18)$$

die lineare Fortsetzung von

$$(\lambda_1 \wedge \dots \wedge \lambda_k) \wedge (\mu_1 \wedge \dots \wedge \mu_l) := \lambda_1 \wedge \dots \wedge \lambda_k \wedge \mu_1 \wedge \dots \wedge \mu_l \quad (23.19)$$

für faktorisierbare  $k$ - und  $l$ -Formen. Für  $k$  oder  $l = 0$  setzt man noch  $\omega \wedge a = a \wedge \omega := a \cdot \omega$ . Die Formel (20.43) wird in der Praxis nicht benötigt.

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

$$\begin{aligned} (\omega_1 \wedge \omega_2) \wedge \omega_3 &= \omega_1 \wedge (\omega_2 \wedge \omega_3) \\ \omega_1 \wedge \omega_2 &= (-1)^{k_1 k_2} \omega_2 \wedge \omega_1 \end{aligned} \quad (23.20)$$

Geometrische Interpretation: Für einen  $n$ -dimensionalen reellen Vektorraum  $V$  ist wegen (20.40)  $\wedge^n V^\vee$  ein-dimensional. Man nennt ein nicht-verschwindendes Element  $\omega \in \wedge^n V^\vee$ ,  $\omega \neq 0$  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  $\alpha \in \wedge^n W^\vee$ , und  $A : W \rightarrow V$  eine lineare Abbildung, so wird durch

$$A^\vee(\omega)(c_1, \dots, c_n) := \omega(Ac_1, \dots, Ac_n) \tag{23.21}$$

eine alternierende  $n$ -Form  $A^\vee(\omega)$  definiert. (Übungsaufgabe). Wegen  $\dim \wedge^n W^\vee = 1$  gilt  $A^\vee(\omega) = d \cdot \alpha$  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 Parallelotope in einem  $n$ -dimensionalen umgebenden Raum, s. insbesondere die Formel (23.16). Im Unterschied zu der Situation für  $k = n$  sind  $k$ -Formen nicht mehr bis auf Skalierung eindeutig, und können auf verschiedene Arten entarten.

### Differentialformen ...

... sind diejenige Klasse von Funktionen auf dem  $\mathbb{R}^n$ , die an den  $n$ -dimensionalen Integralsätzen partizipieren.

**Definition 23.7.** 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

$$\omega : U \rightarrow \wedge^k(\mathbb{R}^n)^\vee \tag{23.22}$$

· Eine Differentialform  $\omega$  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

$$\omega(x_0)(v_1, \dots, v_k) \in \mathbb{R} \tag{23.23}$$

zuzuweisen. Diese Zahl stellt eine Sorte “ $k$ -dimensionales Volumen” des von  $(v_1, \dots, v_k)$  aufgespannten Parallelotops dar, physikalisch üblicherweise interpretiert als “Fluss(dichte)”, in der Elektrodynamik etwa als Magnetfeld.

· Formal gesehen handelt es sich bei dem  $(\mathbb{R}^n)^\vee$  auf der rechten Seite von (23.22) um den Kotangentenraum  $T_{x_0}^\vee U$  von  $U$  an  $x_0$ , und bei den Vektoren in (23.23) um Tangentialvektoren in  $T_{x_0} U$ . Das Parallelotop ist also als *infinitesimal und orientiert* zu denken.

· In diesem Zusammenhang ist es üblich (und sehr sinnvoll, s. (23.28)), die Basis von  $(\mathbb{R}^n)^\vee$ , 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  $\omega$  existieren dann gemäss (20.39)  $\binom{n}{k}$  (stetige, differenzierbare, etc.) Funktionen  $\omega_{i_1, i_2, \dots, i_k} : U \rightarrow \mathbb{R}$  mit

$$\omega(x) = \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} \omega_{i_1, \dots, i_k}(x) \cdot dx^{i_1} \wedge \dots \wedge dx^{i_k} \quad \forall x \in U \tag{23.24}$$

§ 23. INTEGRALSÄTZE

$\omega_{i_1, \dots, i_k}(x)$  ist also die “Flussdichte” durch die von  $(e_{i_1}, \dots, e_{i_k})$  aufgespannte Hyperfläche.

**Beispiel 23.8.** Das euklidische Volumen in der sog. Standardorientierung (s. Def. 23.14) des  $\mathbb{R}^n$  wird von der  $n$ -Form

$$\alpha := dx^1 \wedge \dots \wedge dx^n \tag{23.25}$$

gemessen. Ist  $\mathbb{R}^{n-1} \hookrightarrow \mathbb{R}^n$  ein  $n - 1$ -dimensionaler Unterraum mit einem ausgezeichneten Vektor<sup>31</sup>  $\mathbf{n} \in \mathbb{R}^n \setminus \mathbb{R}^{n-1}$ , so induziert man durch

$$\alpha_{\mathbf{n}}(v_1, \dots, v_{n-1}) = \alpha(\mathbf{n}, v_1, \dots, v_{n-1}) \tag{23.26}$$

eine Volumenform auf  $\mathbb{R}^{n-1}$ . Speziell für die Hyperebenen  $\mathbb{R}_i^{n-1} = \{x^i = 0\}$ ,  $\mathbf{n} = e_i$ ,

$$\alpha_i := (-1)^{i-1} dx^1 \wedge \dots \wedge \widehat{dx^i} \wedge \dots \wedge dx^n \tag{23.27}$$

wobei  $\widehat{\phantom{x}}$  “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  $df := Df : U \rightarrow (\mathbb{R}^n)^\vee$  schreibt, dann bereits ein erstes Beispiel für eine 1-Form. Die Entwicklung (23.24) vereinfacht sich dabei zu:

$$df = \sum_{i=1}^n \frac{\partial f}{\partial x^i} dx^i \tag{23.28}$$

(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 (23.28) 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 (23.24) ist punktweise endlich.)

**Definition 23.9.** Die Cartan- (oder äussere) Ableitung einer  $k$ -Form  $\omega \in \Omega^k(U)$  ist die  $k + 1$ -Form

$$d\left(\sum_{i_1 < \dots < i_k} \omega_{i_1, \dots, i_k} \cdot dx^{i_1} \wedge \dots \wedge dx^{i_k}\right) := \sum_{i_1 < \dots < i_k} d(\omega_{i_1, \dots, i_k}) \wedge dx^{i_1} \wedge \dots \wedge dx^{i_k} \tag{23.29}$$

wobei auf der rechten Seite das punktweise Dachprodukt (23.19) steht. (Es gelten auch die Rechenregeln (23.20). Die Formel (20.43) muss man glücklicherweise praktisch nie benutzen.) Speziell für  $k = 1$ :

$$d\left(\sum_{i=1}^n \omega_i dx^i\right) = \sum_{i=1}^n \sum_{j=1}^n \left(\frac{\partial \omega_i}{\partial x^j} dx^j\right) \wedge dx^i = \sum_{i_1 < i_2} \left(\frac{\partial \omega_{i_2}}{\partial x^{i_1}} - \frac{\partial \omega_{i_1}}{\partial x^{i_2}}\right) dx^{i_1} \wedge dx^{i_2} \tag{23.30}$$

<sup>31</sup> $\mathbf{n}$  ist nicht “normal” im Sinne einer euklidischen Metrik auf  $\mathbb{R}^n$ .

**Lemma 23.10.** (i) Für  $\omega_1 \in \Omega^{k_1}(U)$ ,  $\omega_2 \in \Omega^{k_2}(U)$  gilt

$$\omega_1 \wedge \omega_2 = (-1)^{k_1 k_2} \omega_2 \wedge \omega_1 \quad (23.31)$$

(ii) und die Leibniz-Regel

$$d(\omega_1 \wedge \omega_2) = (d\omega_1) \wedge \omega_2 + (-1)^{k_1} \omega_1 \wedge (d\omega_2) \quad (23.32)$$

(iii)  $d^2 = 0$ , d.h. für alle  $k$ -Formen  $\omega \in \Omega^k(U)$  ist

$$d(d(\omega)) = 0 \in \Omega^{k+2}(U) \quad (23.33)$$

*Beweis.* (i) folgt durch Anwendung der gewöhnlichen Produktregel

$$d(f_1 \cdot f_2) = df_1 \cdot f_2 + f_1 \cdot df_2 \quad (23.34)$$

(für  $k_1 = k_2 = 0$  gilt  $\wedge = \cdot$ ) auf die Koeffizientenfunktionen:

$$\begin{aligned} \omega_1 \wedge \omega_2 &= \\ &= \sum_{i_1, \dots, i_{k_1}} \sum_{j_1, \dots, j_{k_2}} (\omega_1)_{i_1, \dots, i_{k_1}} \cdot (\omega_2)_{j_1, \dots, j_{k_2}} \cdot dx^{i_1} \wedge \dots \wedge dx^{i_{k_1}} \wedge dx^{j_1} \wedge \dots \wedge dx^{j_{k_2}} \end{aligned} \quad (23.35)$$

und die Rechenregeln (23.20).

• Auch (ii) zeigen wir zunächst für  $k = 0$ . Mit (23.28) und (23.30) erhalten wir

$$d(df) = \sum_{i_1 < i_2} \left( \frac{\partial^2 f}{\partial x^{i_1} \partial x^{i_2}} - \frac{\partial^2 f}{\partial x^{i_2} \partial x^{i_1}} \right) dx^{i_1} \wedge dx^{i_2} = 0 \quad (23.36)$$

wegen der Symmetrie der zweiten Ableitungen 10.13. Insbesondere ist auch  $d(dx^i) = 0$ , so dass die Aussage für  $k > 0$  durch Anwenden von (i) auf (23.29) folgt.  $\square$

### Untermannigfaltigkeiten (mit und ohne Rand) ...

... sind diejenigen Teilmengen des  $\mathbb{R}^n$ , über die wir Differentialformen integrieren können. Sie spielen in den höher-dimensionalen Integralsätzen die Rolle der kompakten Intervalle mit ihren Randpunkten. Wir verallgemeinern gleich die Definition 21.14.

Ich ändere hier die Notation und bezeichne lineare Koordinaten im Parameterbereich mit  $x$  (anstatt wie bisher mit  $\xi$ ), lokale Parametrisierungen mit  $\psi$ , und deren Umkehrung (Karten) mit  $\varphi$ .

**Definition 23.11.** Der *linke Halbraum* des  $\mathbb{R}^n$  ist die Teilmenge

$$\mathbb{R}_-^n := \{x \in \mathbb{R}^n \mid x^1 \leq 0\} \quad (23.37)$$

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

$$\varphi : U \cap M \rightarrow \mathbb{R}^k \quad (23.38)$$

## § 23. INTEGRALSÄTZE

eine  $k$ -dimensionale (glatte) Karte mit Rand von  $M$  um  $p$  (kurz: Karte) falls

(i)  $\varphi$  injektiv und stetig ist, eine offene Teilmenge  $V \subset \mathbb{R}^k$  existiert so, dass  $\varphi(U \cap M) = V \cap \mathbb{R}_-^k$ ,<sup>32</sup> sowie

(ii) die Umkehrabbildung  $\psi := \varphi^{-1} : V \cap \mathbb{R}_-^k \rightarrow \mathbb{R}^n$  (unendlich oft) differenzierbar ist<sup>33</sup>, mit  $\text{rang}(D\psi(x)) = k \ \forall x \in V \cap \mathbb{R}_-^k$ .

· 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  $\varphi(p)^1 = 0$ , m.a.W. falls  $\varphi(p) \in \partial\mathbb{R}_-^k = \mathbb{R}_1^{k-1}$ , dem Rand des linken Halbraumes. Beachte aber, dass  $p$  im Allgemeinen kein Randpunkt im  $n$ -dimensionalen Sinne von (22.20) 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 13.1 bzw. 13.4 kann man umgekehrt zeigen, dass für zwei Karten  $(U_1, \varphi_1)$ ,  $(U_2, \varphi_2)$  der “Kartenwechsel”

$$\varphi_2 \circ \varphi_1^{-1} : \varphi_1(U_1 \cap U_2 \cap M) \rightarrow \varphi_2(U_1 \cap U_2 \cap M) \quad (23.39)$$

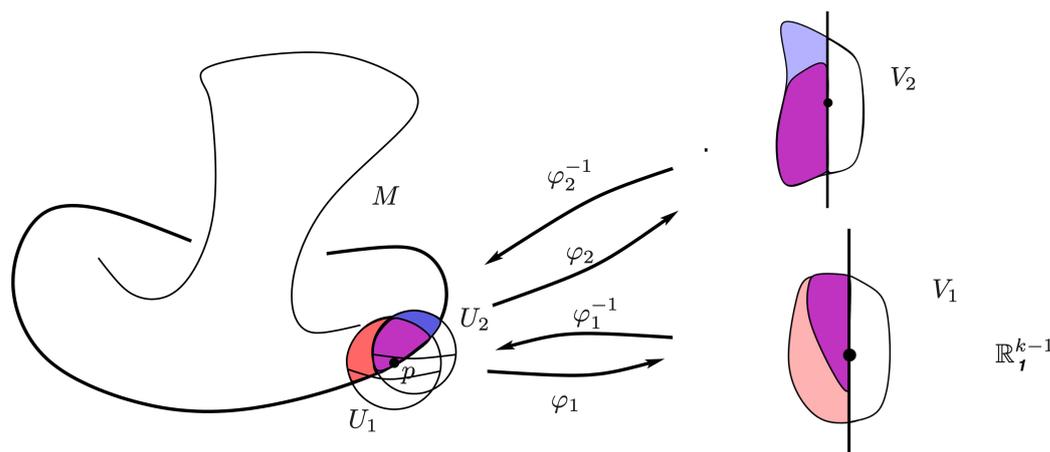
ein Diffeomorphismus von Teilmengen des  $\mathbb{R}^k$  der Form  $(\text{offene Menge}) \cap \mathbb{R}_-^k$  ist. Nun sind Diffeomorphismen insbesondere in beiden Richtungen stetig, und ein Punkt  $x \in \mathbb{R}_-^k$  liegt in  $\mathbb{R}_1^{k-1}$  genau dann, wenn jede Umgebung von  $x$  sowohl Punkte in  $\mathbb{R}_-^k$  als auch Punkte ausserhalb von  $\mathbb{R}_-^k$  enthält. Daraus folgt, dass der Kartenwechsel (23.39)  $\varphi_1(U_1 \cap U_2 \cap M) \cap \mathbb{R}_1^{k-1}$  bijektiv auf  $\varphi_2(U_1 \cap U_2 \cap M) \cap \mathbb{R}_1^{k-1}$  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 die Block-Gestalt

$$D(\varphi_2 \circ \varphi_1^{-1})|_{\varphi_1(U_1 \cap U_2 \cap M) \cap \mathbb{R}_1^{k-1}} = \begin{pmatrix} \frac{\partial(\varphi_2 \circ \varphi_1^{-1})^1}{\partial x^1} & 0 \\ * & \left( \frac{\partial(\varphi_2 \circ \varphi_1^{-1})^j}{\partial x^i} \right)_{i,j=2,\dots,k} \end{pmatrix} \quad (23.40)$$

hat. (“Die 1-te Komponente von  $\varphi_2$  ist für  $x^1 = 0$  stets 0 und hängt daher nicht von  $(x^2, \dots, x^k)$  ab.”)

<sup>32</sup>Für  $p$  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.

<sup>33</sup>Nach unserer Vereinbarung aus 10 bedeutet dies, dass  $\psi$  die Einschränkung einer differenzierbaren Abbildung in einer offenen Umgebung von  $V \cap \mathbb{R}_-^k$  ist.



· Da  $D(\varphi_2 \circ \varphi_1^{-1})$  überall Rang  $k$  hat, gilt dann auf dem Rand

$$\frac{\partial(\varphi_2 \circ \varphi_1^{-1})^1}{\partial x^1} \Big|_{\mathbb{R}_1^{k-1}} \neq 0 \quad \text{und} \quad \text{rang} \left( \frac{\partial(\varphi_2 \circ \varphi_1^{-1})^j}{\partial x^i} \right)_{i,j=2,\dots,k} \Big|_{\mathbb{R}_1^{k-1}} = k - 1 \quad (23.41)$$

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  $\partial M$  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 21.14 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”  $F^{-1}(\{x = (x^1, 0, \dots, 0) \in \mathbb{R}^{n-k+1} \mid x^k \leq 0\})$ . (Wir denken uns hier die Komponenten von  $x \in \mathbb{R}^{n-k+1}$  als  $(x^1, x^{k+1}, \dots, x^n)$  durchnummeriert.)

· Für jeden Punkt  $p \in M$  können wir den Tangentialraum von  $M$  in  $p$  wie in (21.47) definieren. (Unsere lokalen Parametrisierungen stehen ja an den Randpunkten ein wenig in den oberen Halbraum über.) ( $\psi = \varphi^{-1}$ )

$$T_p M = D\psi(\varphi(p))(\mathbb{R}^k) = \{D\psi(\varphi(p))(w) \mid w \in \mathbb{R}^k\} \quad (23.42)$$

Ist  $p \in \partial M$ , so teilt der Tangentialraum an den Rand,

$$T_p \partial M = D\psi(\varphi(p))(\mathbb{R}_k^{k-1}) \quad (23.43)$$

den Tangentialraum an dem Rand in zwei Halbräume von “inneren” und “äusseren” Normalenvektoren, Richtungsvektoren von Kurven, die “nach  $M$  hinein” bzw. “aus  $M$  hinaus” laufen. Weil diese Charakterisierung ebenfalls nicht von der Karte abhängt, gilt in (23.41) tatsächlich

$$\frac{\partial(\varphi_2 \circ \varphi_1^{-1})^1}{\partial x^1} \Big|_{\mathbb{R}_1^{k-1}} > 0 \quad (23.44)$$

## § 23. INTEGRALSÄTZE

Beispiele: Die  $n$ -dimensionale Kugel

$$K^n = \{x \in \mathbb{R}^n \mid \langle x, x \rangle \leq 1\} \quad (23.45)$$

ist eine glatte Untermannigfaltigkeit mit Rand  $S^{n-1}$ .

· Die Halbkugel

$$\{x \in \mathbb{R}^n \mid \langle x, x \rangle = 1, x^n \geq 0\} \quad (23.46)$$

ist eine  $n - 1$ -dimensionale Untermannigfaltigkeit mit Rand  $S^{n-2}$ .

### Atlanten und Zerlegung der Eins

Wir wollen nun betrachten

- eine  $k$ -dimensionale Untermannigfaltigkeit  $M \subset \mathbb{R}^n$ , möglicherweise mit  $(k - 1)$ -dimensionalem Rand, sowie
- eine Differentialform  $\omega \in \Omega^k(\mathfrak{U})$ , welche in einer offenen ( $n$ -dimensionalen) Umgebung von  $M$  definiert (und differenzierbar) ist, mit dem Ziel,  $\omega$  über  $M$  zu integrieren. Für die Definition sind lokale Karten/Koordinaten unumgänglich (wie wir beispielhaft schon in 23.2 gesehen hatten). Das Integral selbst wird aber unabhängig von der *Wahl* der Koordinaten sein, und als solches intrinsisch.

Zur Durchführung brauchen wir noch zwei weitere Voraussetzungen: Kompaktheit und Orientierbarkeit, und zwei technische Hilfsmittel: Eine sog. Teilung der Eins und den Rückzug von Differentialformen.

**Definition 23.12.** Ein *Atlas* auf einer Untermannigfaltigkeit  $M \subset \mathbb{R}^n$  ist eine Menge  $\mathfrak{A} = \{(U_\iota, \varphi_\iota)\}$  von Karten i.S.v. Def. 23.11 mit der Eigenschaft, dass  $M \subset \cup_\iota U_\iota$ .

- Per Definition besitzt jede Untermannigfaltigkeit einen Atlas.
- Durch Diffeomorphismen, Aufteilen von Kartengebieten oder anderweitiges Hinzufügen von Karten lassen sich aus einem gegebenen Atlas andere konstruieren.
- Für die *Definition* ist es bequem, "glatt" vorzugehen, d.h. durch Aufteilung auf offene und überlappende Kartengebiete. Für die *Berechnung* benutzt man normalerweise wieder praktische kompakte Teilmengen im  $\mathbb{R}^k$ .

**Proposition 23.13.** Sei  $M$  eine **kompakte** Untermannigfaltigkeit des  $\mathbb{R}^n$  (mit Rand). Dann existiert ein **endlicher Atlas**  $\mathfrak{A} = \{(U_\iota, \varphi_\iota)\}$  und beliebig oft differenzierbare Funktionen  $h_\iota : U_\iota \rightarrow \mathbb{R}$  mit der Eigenschaft, dass

(i)  $1 \geq h_\iota \geq 0$

(ii)  $\text{supp}(h_\iota)$  ist kompakt in  $U_\iota$

(iii)  $\sum_\iota h_\iota(y) = 1 \quad \forall y \in M$

Ist  $M \subset \mathfrak{A}$  mit  $\mathfrak{A}$  offen in  $\mathbb{R}^n$ , so kann o.E. der Atlas so gewählt werden, dass  $U_\iota \subset \mathfrak{A}$ .

- Die Existenz eines endlichen Atlases (ohne weitere Eigenschaften) folgt direkt aus der Kompaktheit (s. Beweis).
- Zur Konstruktion benutzen wir den euklidischen Abstand  $\|\cdot\|_2$ , was bequem, aber nicht zwingend erforderlich ist.  $B_r(p) = \{y \in \mathbb{R}^n \mid \|y - p\|_2 < r\}$ .

*Beweis.* · Aus (23.38) erhält man durch Verkleinern von  $U$  für jedes  $p \in M$  eine Karte der Form

$$\varphi_p : B_{2r_p}(p) \cap M \rightarrow V_p \cap \mathbb{R}_-^k \quad (23.47)$$

für einen geeignet gewählten Radius  $r_p > 0$  so, dass noch  $B_{2r_p}(p) \subset \mathfrak{U}$ .

· Nun überdecken die offenen Mengen

$$\{B_{r_p}(p) \mid p \in M\} \quad (23.48)$$

trivialerweise ganz  $M$  (d.h. jedes  $p \in M$  ist in einer dieser Mengen enthalten). Aus der Kompaktheit von  $M$  folgt<sup>34</sup>, 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

$$M \subset \bigcup_{\iota=1}^{\varsigma} B_{r_\iota}(p_\iota) \quad (23.49)$$

· Es sei nun  $h : [0, \infty) \rightarrow \mathbb{R}$  eine glatte Funktion mit der Eigenschaft, dass

- (i)  $1 \geq h(t) \geq 0 \forall t \in [0, \infty)$
- (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

$$h(t) := 1 - \frac{\int_0^t v(s) ds}{\int_0^{3/2} v(s) ds}, \quad (23.50)$$

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 15.

---

· Damit ist für jedes  $\iota = 1, \dots, \varsigma$  die Funktion (glatter Hut)

$$\begin{aligned} \hat{h}_\iota &: \mathbb{R}^n \rightarrow \mathbb{R}, \\ \hat{h}_\iota(y) &:= h\left(\frac{\|y - p_\iota\|_2}{r_\iota}\right) \end{aligned} \quad (23.51)$$

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)  $\text{supp}(\hat{h}_\iota) \subset B_{2r_\iota}(p_\iota)$ .

---

<sup>34</sup>Dies ist eine Konsequenz des Satzes von Heine-Borel, vgl. auch den Begriff der “Überdeckungskompaktheit”, den wir evtl. in den Übungen kennengelernt haben.

### § 23. INTEGRALSÄTZE

Wegen  $\hat{h}_\iota > 0$  auf  $B_{r_\iota}(p_\iota)$  verschwindet  $\sum_{\iota=1}^s \hat{h}_\iota$  nirgends auf  $\mathfrak{B} := \bigcup_{\iota=1}^s B_{r_\iota}(p_\iota) \subset \mathfrak{U}$ .  
Daher gilt für die auf  $\mathfrak{B}$  definierten Funktionen

$$h_\iota := \frac{\hat{h}_\iota}{\sum_{\iota=1}^s \hat{h}_\iota} \quad (23.52)$$

$\sum_\iota h_\iota(y) = 1$ , und  $\text{supp}(h_\iota) \subset B_{2r_\iota}(p_\iota)$  ist kompakt, wie gefordert.  $\square$

· Man nennt so eine Sammlung von Funktionen eine “der Überdeckung (23.49) untergeordnete Teilung der Eins”.

Bemerkung: 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  $(h_\iota)$  einen endlichen Atlas mit zugehöriger Teilung der Eins auch auf dem Rand von  $M$ .

### Orientierung, Rückzug, und Integration

Wir “erinnern” kurz:

**Definition 23.14.** · Zwei Basen  $B_1, B_2$  eines endlich-dimensionalen reellen Vektorraums  $V$  heißen *gleich orientiert*, wenn der Basiswechsel  $B_2^{-1}B_1 : \mathbb{R}^n \rightarrow \mathbb{R}^n$  positive Determinante hat. Eine Orientierung von  $V$  ist die Wahl einer (der zwei) Äquivalenzklassen gleich orientierter Basen.

· Die Standardorientierung des  $\mathbb{R}^n$  ist die Klasse von  $(e_1, \dots, e_n)$ .

· Ist  $M$  eine Untermannigfaltigkeit und  $p \in M$ , so ist der Tangentialraum  $T_pM$  ein  $k$ -dimensionaler Vektorraum. Die Wahl einer Karte  $(U, \varphi = \psi^{-1})$  “induziert” eine Orientierung von  $T_pM \forall p \in U \cap M$  als die Klasse des Bildes der Standardbasis von  $\mathbb{R}^k$  unter  $D\psi$ . Sie hängt aber im Allgemeinen von der Wahl der Karte ab.

**Definition 23.15.** · Zwei Karten  $(U_1, \varphi_1)$  und  $(U_2, \varphi_2)$  heißen orientierungsverträglich, falls

$$\det D(\varphi_2 \circ \varphi_1^{-1})(x) > 0 \quad \forall x \in \varphi_1(U_1 \cap U_2 \cap M) \quad (23.53)$$

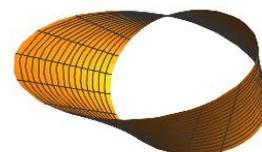
· Eine Untermannigfaltigkeit  $M$  heisst *orientierbar*, falls ein Atlas  $\mathfrak{A} = \{(U_\iota, \varphi_\iota)\}$  existiert, so dass jedes Paar von zwei Karten aus  $\mathfrak{A}$  orientierungsverträglich sind. Ein solcher Atlas heisst *orientiert*.

· Eine orientierte Untermannigfaltigkeit ist eine orientierbare Untermannigfaltigkeit zusammen mit der Wahl einer Äquivalenzklasse von orientierten Atlanten. Wir schreiben im Folgenden  $M^\uparrow$  für eine Untermannigfaltigkeit mit ausgezeichnete Orientierung.

Bemerkung: Im Unterschied zu Vektorräumen ist nicht jede Untermannigfaltigkeit orientierbar und im Allgemeinen besitzt eine orientierbare Untermannigfaltigkeit mehr als zwei verschiedene Orientierungen (z.B. unzusammenhängende Untermannigfaltigkeiten). Auf keinen Fall ist es möglich zu sagen, welche der Orientierungen “positiv” bzw. “negativ” ist.

Beispiel: Der Möbius-Streifen, Bild der Abbildung

$$[0, 2\pi] \times [-1/2, 1/2] \ni (\vartheta, s) \mapsto \begin{pmatrix} \cos \vartheta (1 - s \sin \frac{\vartheta}{2}) \\ \sin \vartheta (1 - s \sin \frac{\vartheta}{2}) \\ s \cos \frac{\vartheta}{2} \end{pmatrix} \quad (23.54)$$



ist eine nicht-orientierbare kompakte 2-dimensionale Untermannigfaltigkeit des  $\mathbb{R}^3$  mit Rand.

**Lemma 23.16.** *Ist  $M^\uparrow$  eine orientierte Untermannigfaltigkeit mit Rand, so erhält man durch Einschränken des orientierten Atlases auf den Rand dort wieder einen orientierten Atlas. Wir schreiben  $\partial M^\uparrow$  für den Rand mit der so induzierten Orientierung.*

*Beweis.* Aus (23.40) und (23.44) folgt

$$\begin{aligned} \det \left( \frac{\partial(\varphi_\iota \circ \varphi_\kappa^{-1})^j}{\partial x^i} \right)_{i,j=1,\dots,k} &> 0 \quad \forall x \in \varphi_\kappa(U_\iota \cap U_\kappa \cap M) \\ \Rightarrow \det \left( \frac{\partial(\varphi_\iota \circ \varphi_\kappa^{-1})^j}{\partial x^i} \right)_{i,j=2,\dots,k} &\forall x \in \varphi_\kappa(U_\iota \cap U_\kappa \cap \partial M) \end{aligned} \quad (23.55)$$

□

**Definition 23.17.** Es sei  $M$  eine  $k$ -dimensionale Untermannigfaltigkeit von  $\mathbb{R}^n$ ,  $M \subset \mathfrak{U}$  offen, und  $\omega \in \Omega^l(\mathfrak{U})$  (gleich interessieren uns vor allem  $l = k$  und  $k - 1$ ). Dann ist für jede Karte  $\varphi : U \cap M \xrightarrow{\cong} V \cap \mathbb{R}_+^k$  von  $M$  der Rückzug von  $\omega$  entlang  $\psi = \varphi^{-1}$  die  $l$ -Form auf  $V \cap \mathbb{R}_+^k$  definiert 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  $D\psi(x)(v_i) \in T_{\psi(x)}M \subset T_{\psi(x)}\mathbb{R}^n$  und wir setzen

$$(\psi^*(\omega)(x))(v_1, \dots, v_l) := \omega(\psi(x))(D\psi(x)(v_1), \dots, D\psi(x)(v_l)) \quad (23.56)$$

(Dies ist klarerweise eine alternierende  $l$ -Form auf  $T_x V$ , die Differenzierbarkeit folgt aus Kettenregeln, d.h.  $\psi^*(\omega) \in \Omega^l(V \cap \mathbb{R}_+^k)$ .) Anschaulich: Wir messen das ‘‘Volumen’’ des durch  $v_1, \dots, v_l$  aufgespannten Parallelotops durch Auswerten von  $\omega$  auf ihrem Bild unter  $D\psi$ .

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 ( $v \in T_x V$ ):

$$\begin{aligned} (\psi^*(df)(x))(v) &= \sum_{j=1}^n \frac{\partial f}{\partial y^j}(\psi(x)) \overbrace{\sum_{i=1}^k \frac{\partial \psi^j}{\partial x^i}(x) v^i}^{(D\psi(x)(v))^j} \\ &= \sum_{i=1}^k \left( \sum_{j=1}^n \frac{\partial f}{\partial y^j}(\psi(x)) \frac{\partial \psi^j}{\partial x^i}(x) \right) v^i \\ \text{(Kettenregel)} &= \frac{\partial(f \circ \psi)}{\partial x^i}(x) v^i \\ &= d(\psi^*(f))(v) \end{aligned} \quad (23.57)$$

## § 23. INTEGRALSÄTZE

was man unschwer verallgemeinert:

**Lemma 23.18.** *Es gilt*

- (i)  $\psi^*(\omega)$  ist linear in  $\omega$
- (ii)  $\psi^*(\omega_1 \wedge \omega_2) = \psi^*(\omega_1) \wedge \psi^*(\omega_2)$
- (iii)  $d(\psi^*(\omega)) = \psi^*(d(\omega))$

*Beweis.* klar/wird nachgeliefert. □

Beachte: In (iii) operiert das  $d$  auf der linken Seite im  $\mathbb{R}^k$ , das auf der rechten Seite im  $\mathbb{R}^n$ , es sind aber nur die Ableitungen in den Tangentialrichtungen von  $M^\uparrow$  relevant. Diese Verträglichkeit mit dem Rückzug ist (neben  $d^2 = 0$ ) eine weitere wichtige Eigenschaft der äusseren Ableitung.

Wir betrachten nun speziell  $k = l$ , und eine orientierte Untermannigfaltigkeit  $M^\uparrow$  mit endlichem orientierten Atlas  $\mathfrak{A} = \{(U_\iota, \varphi_\iota)\}$  und untergeordneter Teilung der Eins ( $h_\iota$ ) (die Orientierung liefert offenbar keine weitere Einschränkungen an die Teilung). Dann ist für alle  $\iota$  mit  $\psi_\iota = \varphi_\iota^{-1}$ :

$$\psi_\iota^*(h_\iota \cdot \omega) \in \Omega_c^k(V_\iota \cap \mathbb{R}_-^k) \quad (23.58)$$

eine  $k$ -Form mit kompaktem Träger. Wegen  $\dim \wedge^k(\mathbb{R}^k)^\vee = 1$  existiert eine Funktion  $\omega_\iota \in \mathcal{C}_c^\infty(V_\iota \cap \mathbb{R}_-^k)$ , so dass

$$\psi_\iota^*(h_\iota \cdot \omega) = \omega_\iota(x) \cdot dx^1 \wedge \dots \wedge dx^k =: \omega_\iota(x) \cdot \alpha \quad (23.59)$$

Mit anderen Worten ist

$$\omega_\iota(x) = h_\iota(\psi_\iota(x)) \cdot \omega(\psi_\iota(x)) (D\psi_\iota(x)(e_1), \dots, D\psi_\iota(x)(e_k)) \quad (23.60)$$

wobei  $(e_1, \dots, e_k)$  die Standardbasis des  $\mathbb{R}^k$  ist. Wir definieren dann das *Integral der  $k$ -Form  $\omega \in \Omega^k(\mathfrak{A})$  über die orientierte kompakte  $k$ -dimensionale Untermannigfaltigkeit mit Rand  $M^\uparrow$*  durch die Formel:

$$\int_{M^\uparrow} \omega := \sum_\iota \int_{\mathbb{R}_-^k} \omega_\iota(x) d^k x = \sum_\iota \int_{\mathbb{R}^k} \chi_{\mathbb{R}_-^k}(x) \omega_\iota(x) d^k x \quad (23.61)$$

wobei auf der rechten Seite das Integral über kompakte Mengen/kompakten Träger aus dem letzten § steht. Diese Definition ist sinnvoll wegen des folgenden Lemmas:

**Lemma 23.19.** *Der Wert des Integrals (23.61) 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  $\tilde{U}_\kappa \subset \mathfrak{A}$  und der gleichen Orientierung wie  $\mathfrak{A}$ , d.h.

$$\det D(\varphi_\iota \circ \tilde{\varphi}_\kappa^{-1}) > 0 \quad \text{auf } \tilde{\varphi}_\kappa(\tilde{U}_\kappa \cap U_\iota \cap M^\uparrow) \quad \forall \iota, \kappa, \quad (23.62)$$

und  $(\tilde{h}_\kappa)$  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  $M^\dagger$ .  
 Dann gilt wegen der Linearität des Integrals

$$\begin{aligned} \sum_l \int_{\mathbb{R}_-^k} \omega_l(x) dx &= \sum_l \int_{\mathbb{R}_-^k} \sum_\kappa \tilde{h}_\kappa(\psi_l(x)) \cdot \omega_l(x) dx \\ &= \sum_l \sum_\kappa \int_{\varphi_l(U_l \cap \tilde{U}_\kappa \cap M^\dagger)} \tilde{h}_\kappa(\psi_l(x)) \cdot \omega_l(x) dx \end{aligned}$$

Durch die Substitution  $x = (\varphi_l \circ \tilde{\varphi}_\kappa^{-1})(\tilde{x})$  (d.h.  $\psi_l(x) = \tilde{\psi}_\kappa(\tilde{x}) \in M^\dagger$ ) wird daraus mit Hilfe der Transformationsformel 22.21 und Gl. (23.62) einerseits

$$= \sum_\kappa \sum_l \int_{\tilde{\varphi}_\kappa(\tilde{U}_\kappa \cap U_l \cap M^\dagger)} \tilde{h}_\kappa(\tilde{\psi}_\kappa(\tilde{x})) \cdot \omega_l(x) \cdot \det D(\varphi_l \circ \tilde{\varphi}_\kappa^{-1})(\tilde{x}) d\tilde{x} \quad (23.63)$$

Wegen

$$D\psi_l(x) = D(\tilde{\psi}_\kappa \circ \tilde{\varphi}_\kappa \circ \varphi_l)(x) = D(\tilde{\psi}_\kappa)(\tilde{x}) \cdot D(\tilde{\varphi}_\kappa \circ \varphi_l^{-1})(x) \quad (23.64)$$

folgt aus (23.60) andererseits

$$\begin{aligned} \tilde{h}_\kappa(\tilde{\psi}_\kappa(\tilde{x})) \cdot \omega_l(x) &= \tilde{h}_\kappa(\tilde{\psi}_\kappa(\tilde{x})) \cdot h_l(\psi_l(x)) \cdot \omega(\psi_l(x))(D\psi_l(x)(e_1), \dots, D\psi_l(x)(e_k)) \\ &= h_l \cdot \underbrace{\tilde{h}_\kappa \cdot \omega(\tilde{\psi}_\kappa(\tilde{x}))(D\tilde{\psi}_\kappa(\tilde{x})(e_1), \dots, D\tilde{\psi}_\kappa(\tilde{x})(e_k))}_{=\tilde{\omega}_\kappa(\tilde{x})} \cdot \det D(\tilde{\varphi}_\kappa \circ \varphi_l^{-1})(x) \end{aligned} \quad (23.65)$$

wobei wir die Transformation (23.21) einer Volumenform unter einem Endomorphismus des  $\mathbb{R}^k$  benutzt haben.<sup>35</sup>

• 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 (23.63)

$$\begin{aligned} \sum_l \int_{\mathbb{R}_-^k} \omega_l(x) dx &= \sum_\kappa \sum_l \int_{\tilde{\varphi}_\kappa(\tilde{U}_\kappa \cap U_l \cap M^\dagger)} h_l(\tilde{\psi}_\kappa(\tilde{x})) \cdot \tilde{\omega}_\kappa(\tilde{x}) d\tilde{x} \\ &= \sum_\kappa \int_{\mathbb{R}_-^k} \underbrace{\sum_l h_l(\tilde{x}) \cdot \tilde{\omega}_\kappa(\tilde{x})}_{=1} d\tilde{x} = \sum_\kappa \int_{\mathbb{R}_-^k} \tilde{\omega}_\kappa(\tilde{x}) d\tilde{x} \end{aligned} \quad (23.67)$$

wie behauptet. □

<sup>35</sup>Um 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 (23.64)

$$\omega(D\psi_l(e_1), \dots, D\psi_l(e_k)) = \sum_{j_1, \dots, j_k=1}^k \omega(D\tilde{\psi}_\kappa(e_{j_1})A_1^{j_1}, \dots, D\tilde{\psi}_\kappa(e_{j_k})A_k^{j_k})$$

Da  $\omega$  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

$$= \omega(D\tilde{\psi}_\kappa(e_1), \dots, D\tilde{\psi}_\kappa(e_k)) \cdot \underbrace{\sum_{\sigma \in S_k} \text{sgn}(\sigma) \prod_{j=1}^k A_{\sigma^{-1}(j)}^j}_{=\det A} \quad (23.66)$$

nach der Leibniz-Formel für die Entwicklung der Determinante.

## § 23. INTEGRALSÄTZE

· Wir schreiben ab jetzt auch

$$\int_{\mathbb{R}^k} \psi_\iota^*(h_\iota \cdot \omega) \quad \text{für} \quad \int_{\mathbb{R}^k} \omega_\iota(x) dx \quad (23.68)$$

· Für  $n = 1$  lässt sich nun (etwas umständlich) das bestimmte Integral

$$\int_a^b f(s) ds = \int_{[a,b]} f(x) dx \quad (23.69)$$

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.

· Der Gaussche Satz 23.2 lässt sich wie folgt wiedergeben. Auf der linken Seite ist  $\int_K f(x) dx$  das Integral der Differentialform  $\omega = f(x) dx^1 \wedge \cdots \wedge dx^n$  über  $K = \Phi(Q)$  in der Standardorientierung. Es kann berechnet werden durch Rückzug via  $\Phi$  als Integral über  $Q$ :

$$\int_K f(x) d^n x = \int_{\Phi(Q)} \omega = \int_Q \Phi^*(\omega) = \int_Q f(\Phi(\xi)) \det(\partial_\beta \Phi^j)_{\beta,j=1,\dots,n} d^n \xi \quad (23.70)$$

Mit  $\eta = \sum_i F^{(i)} (-1)^{i-1} dx^1 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^n$  ist  $d\eta = \omega$ , und die rechte Seite ist das Integral von  $\eta$  über den Rand von  $K$ , ausgewertet durch Rückzug via  $\Phi$  auf die Seiten von  $Q$  in der induzierten Orientierung.

$$\int_{\partial K} \eta = \sum_{\partial Q} \Phi^*(\eta) = \sum_{i,\alpha} (-1)^{i-1} (-1)^{\alpha-1} \int_{\xi^\alpha = \delta_\alpha} F^{(i)} \det(\partial_\beta \Phi^j)_{\beta \neq \alpha, j \neq i} d^{n-1} \xi \quad (23.71)$$

Der Gaussche Integralsatz in diesem Fall ist die Identität:

$$\int_K \omega = \int_{\partial K} \eta \quad (23.72)$$

### Der Stokessche Satz

Allgemein ist dann für eine Differentialform  $\omega$  vom Grad  $k$  eine "Stammform" eine  $k - 1$ -Form  $\eta$  mit  $d\eta = \omega$ , und die eine Richtung des Hauptsatzes lautet:

**Theorem 23.20.** *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  $\eta \in \Omega^{k-1}(\mathfrak{U})$*

$$\int_{M^\uparrow} d\eta = \int_{\partial M^\uparrow} \eta \quad (23.73)$$

*Beweis.* Es sei  $\{U_\iota, \varphi_\iota = \psi_\iota^{-1}\}$  ein orientierter Atlas der obigen Art und  $(h_\iota)$  eine untergeordnete Teilung der Eins. Dann erhält man durch geeignetes Einschränken Gleiches auf  $\partial M^\uparrow$ . (Wobei hier die  $p_\iota$  nicht auf  $\partial M^\uparrow$  liegen müssen, und es auch mal sein kann, dass  $\text{supp}(h_\iota) \cap \partial M^\uparrow = \emptyset$ .) Da  $\partial M^\uparrow$  leeren Rand hat, lautet die Behauptung zunächst

$$\sum_\iota \int_{\mathbb{R}_-^k} \psi_\iota^*(h_\iota \cdot d\eta) = \sum_\iota \int_{\mathbb{R}_1^{k-1}} \psi_\iota^*(h_\iota \cdot \eta) \quad (23.74)$$

Wegen der Produktregel 23.10 und der Verträglichkeit (23.18) gilt auf der linken Seite

$$\begin{aligned} \sum_\iota \int_{\mathbb{R}_-^k} \psi_\iota^*(h_\iota \cdot d\eta) &= \sum_\iota \int_{\mathbb{R}_-^k} \psi_\iota^*(d(h_\iota \cdot \eta)) - \sum_\iota \int_{\mathbb{R}_-^k} \psi_\iota^*(dh_\iota \wedge \eta) \\ &= \sum_\iota \int_{\mathbb{R}_-^k} d(\psi_\iota^*(h_\iota \cdot \eta)) - \underbrace{\int_{M^\uparrow} d\left(\underbrace{\sum_\iota h_\iota}_{=1}\right) \wedge \eta}_{=0} \end{aligned} \quad (23.75)$$

Es genügt daher zu zeigen, dass in jeder Karte

$$\int_{\mathbb{R}_-^k} d(\psi_\iota^*(h_\iota \cdot \eta)) = \int_{\mathbb{R}_1^{k-1}} \psi_\iota^*(h_\iota \cdot \eta) \quad (23.76)$$

Für die  $k - 1$ -Formen existieren Funktionen  $\eta_\iota^i \in C_c^\infty(V_\iota \cap \mathbb{R}_-^k)$  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:

$$\psi_\iota^*(h_\iota \cdot \eta) = \sum_{i=1}^k \eta_\iota^i \alpha_i \quad (23.77)$$

Es gilt dann

$$d(\psi_\iota^*(h_\iota \cdot \eta)) = \sum_{i=1}^k (d\eta_\iota^i) \wedge \alpha_i = \sum_{i=1}^k \frac{\partial \eta_\iota^i}{\partial x^i} \cdot \alpha_i \quad (23.78)$$

so dass mit Hilfe von Prop. 22.12

$$\begin{aligned} \int_{\mathbb{R}_-^k} d(\psi_\iota^*(h_\iota \cdot \eta)) &= \sum_{i>1} \int_{\mathbb{R}_-^{k-1}} \overbrace{\left( \int_{-\infty}^{\infty} \frac{\partial \eta_\iota^i}{\partial x^i}(x^1, \dots, x^k) dx^i \right)}{=0, \text{ da } \text{supp}(\eta_\iota^i) \text{ kompakt}} d^{k-1}x \\ &\quad + \int_{\mathbb{R}_1^{k-1}} \left( \int_{-\infty}^0 \frac{\partial \eta_\iota^1}{\partial x^1} dx^k \right) d^{k-1}x \quad (23.79) \\ &= \int_{\mathbb{R}_1^{k-1}} \eta_\iota^1(0, x^2, \dots, x^k) d^{k-1}x \end{aligned}$$

nach dem ein-dimensionalen Hauptsatz, angewendet auf  $\eta_\iota^1(\cdot, x^2, \dots, x^k) \in C_c^\infty((-\infty, 0])$ . Andererseits ist wegen  $x^1 = 0 = \text{konst.}$  auf  $\mathbb{R}_1^{k-1}$ ,  $\alpha_i|_{\mathbb{R}_1^{k-1}} = 0$  ausser für  $i = 1$ . Daher gilt auch

$$\int_{\mathbb{R}_1^{k-1}} \psi_\iota^*(h_\iota \cdot \eta) = \int_{\mathbb{R}_1^{k-1}} \eta_\iota^1(0, x^2, \dots, x^k) d^{k-1}x \quad (23.80)$$

Zusammen ist dies die Behauptung. □

**Das Poincarésche Lemma**

Für die andere Richtung des Hauptsatzes (Konstruktion einer Stammform durch Integration) benötigen wir eine Voraussetzung an die Topologie des Definitionsbereichs. Wie in der Funktionentheorie 16 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 23.10) 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 23.21.** Die Kontraktion (oder Paarung) von alternierenden Multilinearformen mit Vektoren ist die bilineare Abbildung

$$\begin{aligned} \iota_v : \mathbb{R}^n \times \wedge^k(\mathbb{R}^n)^\vee &\rightarrow \wedge^{k-1}(\mathbb{R}^n)^\vee \\ (\iota_v(\omega))(v_1, \dots, v_{k-1}) &:= \omega(v, v_1, \dots, v_{k-1}) \end{aligned} \tag{23.81}$$

(vgl. (23.26)). In den Standardbasen  $(e_1, \dots, e_n)$  und  $\{dx^{i_1} \wedge \dots \wedge dx^{i_k} \mid i_1 < \dots < i_n\}$  ist

$$\iota_{e_i}(dx^{i_1} \wedge \dots \wedge dx^{i_k}) = \begin{cases} 0 & \text{falls } i \notin \{i_1, \dots, i_k\} \\ (-1)^{r-1} dx^{i_1} \wedge \dots \wedge \widehat{dx^{i_r}} \wedge \dots \wedge dx^{i_k} & \text{für } i = i_r \end{cases} \tag{23.82}$$

Es gilt für alle  $\omega \in \wedge^k(\mathbb{R}^n)^\vee$ :

$$\iota_{e_i}(dx^j \wedge \omega) + dx^j \wedge \iota_{e_i}(\omega) = \delta_i^j \omega = \begin{cases} \omega & j = i \\ 0 & \text{sonst} \end{cases} \tag{23.83}$$

sowie

$$\sum_{i=1}^n dx^i \wedge \iota_{e_i} \omega = k \cdot \omega \tag{23.84}$$

*Beweis.* Übungsaufgabe (Blatt 9). Wichtig ist noch die offensichtlichen punktweise Verallgemeinerung dieser Formeln für Vektorfelder und Differentialformen.  $\square$

**Theorem 23.22.** 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  $d\eta = \omega$ .

**Definition 23.23.** 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  $\eta$  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.

*Beweis von Thm. 23.22.* 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

$$(I_{x_0}(\omega))(x)(v_1, \dots, v_{k-1}) := \int_0^1 t^{k-1} \sum_{i=1}^n (x^i - x_0^i) \omega(x_0 + t(x - x_0))(e_i, v_1, \dots, v_{k-1}) dt \quad (23.85)$$

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  $\omega$  und des Vertauschungssatzes 25.13). Im Sinne des Integrals vektorwertiger Funktionen ist dies auch

$$I_{x_0}(\omega)(x) = \int_0^1 t^{k-1} \iota_{\sum_i (x-x_0)^i e_i} (\omega(x_0 + t(x-x_0))) dt \quad (23.86)$$

Beh.: Es gilt für alle  $\omega \in \Omega^k(U)$

$$d(I_{x_0}(\omega)) + I_{x_0}(d\omega) = \omega \quad (23.87)$$

Ist insbesondere  $d\omega = 0$ , so erfüllt  $\eta := I_{x_0}(\omega)$  die Bedingung  $d\eta = \omega$ .

Bew.: Der Einfachheit halber transferieren wir die Behauptung nach  $x_0 = 0$ , und benutzen die Linearität in  $\omega$ , 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:

$$\begin{aligned} & d\left(\int_0^1 t^{k-1} \sum_{i=1}^n x^i \iota_{e_i} \omega(tx) dt\right) + \int_0^1 t^k \sum_{i=1}^n x^i \iota_{e_i} (d\omega)(tx) dt = \\ & \int_0^1 \left[ t^{k-1} f(tx) \sum_{i=1}^n dx^i \wedge \iota_{e_i} (dx^{i_1, \dots, i_k}) + t^k \sum_{i,j=1}^n x^i \frac{\partial f}{\partial x^j} (tx) dx^j \wedge \iota_{e_i} (dx^{i_1, \dots, i_k}) \right. \\ & \quad \left. + t^k \sum_{i,j=1}^n x^i \iota_{e_i} \left( \frac{\partial f}{\partial x^j} (tx) dx^j \wedge dx^{i_1, \dots, i_k} \right) \right] dt \\ (\text{mit Lemma 23.21}) & = \int_0^1 \left[ t^{k-1} f(tx) \cdot k \cdot dx^{i_1, \dots, i_k} + t^k \underbrace{\sum_{i=1}^n x^i \frac{\partial f}{\partial x^i} (tx)}_{= \frac{df(tx)}{dt}} \cdot dx^{i_1, \dots, i_k} \right] dt \\ & = \int_0^1 \frac{d}{dt} (t^k f(tx)) dt \cdot dx^{i_1, \dots, i_k} = f(x) \cdot dx^{i_1, \dots, i_k} \end{aligned} \quad (23.88)$$

wegen  $k > 0$  und dem gewöhnlichen Hauptsatz in einer Dimension. □

### Vektoranalysis

Die Rechenregeln für die vertrauten Vektordifferentialoperatoren im  $\mathbb{R}^3$  lassen sich recht elegant in die Sprache der Differentialformen umformulieren. Der entscheidende Hinweis ist die Ähnlichkeit zwischen (23.30) und der expliziten Formel (10.58)

§ 23. INTEGRALSÄTZE

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  $\cdot^b : T_{x_0}U \rightarrow T_{x_0}^\vee U$  und  $\cdot^\sharp : T_{x_0}^\vee U \rightarrow T_{x_0}U$  als Spezialfälle von (21.5) bezüglich dem euklidischen Standardprodukt (21.7) auf  $\mathbb{R}^3$ .

· Allgemein ist das für  $\omega_1, \omega_2 \in \wedge^k(\mathbb{R}^n)^\vee$  durch

$$\langle \omega_1, \omega_2 \rangle := \sum_{i_1 < \dots < i_k} (\omega_1)_{i_1, \dots, i_k} (\omega_2)_{i_1, \dots, i_k} \quad (23.89)$$

definierte innere Produkt symmetrisch und positiv definit, induziert also einen Isomorphismus

$$\wedge^k(\mathbb{R}^n)^\vee \cong (\wedge^k(\mathbb{R}^n)^*)^\vee \quad (23.90)$$

· Andererseits wird das Dachprodukt

$$\begin{aligned} \wedge^k(\mathbb{R}^n)^\vee \times \wedge^{n-k}(\mathbb{R}^n)^\vee &\rightarrow \wedge^n(\mathbb{R}^n)^\vee \\ (\omega, \tilde{\omega}) &\mapsto \omega \wedge \tilde{\omega} \end{aligned} \quad (23.91)$$

durch die Identifikation

$$\begin{aligned} \wedge^n(\mathbb{R}^n)^\vee &\rightarrow \mathbb{R} \\ \alpha = dx^1 \wedge \dots \wedge dx^n &\mapsto 1 \end{aligned} \quad (23.92)$$

zu einer Bilinearform

$$\wedge^k(\mathbb{R}^n)^\vee \times \wedge^{n-k}(\mathbb{R}^n)^\vee \rightarrow \mathbb{R} \quad (23.93)$$

In den Basen (20.39) ist explizit

$$\begin{aligned} &\left( (dx^{i_1} \wedge \dots \wedge dx^{i_k}) \wedge (dx^{j_1} \wedge \dots \wedge dx^{j_{n-k}}) \right) / dx^1 \wedge \dots \wedge dx^n \\ &= \begin{cases} 0 & \text{falls } \{i_1, \dots, i_k\} \cup \{j_1, \dots, j_{n-k}\} \neq \{1, \dots, n\} \\ \pm 1 & \text{je nach Parität der Permutation} \\ & (i_1, \dots, i_k, j_1, \dots, j_{n-k}) = (\sigma(1), \dots, \sigma(n)) \text{ sonst} \end{cases} \end{aligned} \quad (23.94)$$

Offenbar ist also (23.93) nicht-entartet, und liefert daher einen Isomorphismus

$$(\wedge^k(\mathbb{R}^n)^*)^\vee \cong \wedge^{n-k}(\mathbb{R}^n)^\vee \quad (23.95)$$

Durch Zusammenschalten mit (23.90) erhalten wir den sogenannten Hodge-Stern-Operator

$$\star : \wedge^k(\mathbb{R}^n)^\vee \xrightarrow{\cong} \wedge^{n-k}(\mathbb{R}^n)^\vee \quad (23.96)$$

Zwei Formeln dafür:

$$\omega_1 \wedge (\star \omega_2) = \langle \omega_1, \omega_2 \rangle \cdot dx^1 \wedge \dots \wedge dx^n \quad (23.97)$$

und

$$\begin{aligned} \star(dx^{i_1} \wedge \cdots \wedge dx^{i_k}) &= \pm 1 \cdot dx^{j_1} \wedge \cdots \wedge dx^{j_{n-k}} \\ \text{je nach Parität der Permutation } (i_1, \dots, i_k, j_1, \dots, j_{n-k}) &= (\sigma(1), \dots, \sigma(n)) \end{aligned} \quad (23.98)$$

· Zurück im  $\mathbb{R}^3$  überprüft man nun, dass die abstrakte Formel (10.57) nichts anderes ist als

$$\text{rot}(Y) = \underbrace{\underbrace{\star(d \underbrace{Y^b}_{\in \wedge^1})}_{\in \wedge^2}}_{\in \wedge^1} \# \in TU \quad (23.99)$$

Für (10.55) kann man schreiben

$$\text{div } Y = \star \underbrace{\underbrace{\underbrace{d \star Y^b}_{\in \wedge^1}}_{\in \wedge^2}}_{\in \wedge^3} \in \wedge^0(\mathbb{R}^3)^* = \mathbb{R} \quad (23.100)$$

Ausserdem ist (per Definition!)

$$(\text{grad } f)^b = df \quad (23.101)$$

Aus diesen drei Formeln folgt mit 23.10 (ii)

$$\begin{aligned} \text{rot} \circ \text{grad } f &= (\star d(df))^\# = 0 \\ \text{div} \circ \text{rot } Y &= \star d(\star \star dY^b) = 0 \end{aligned} \quad (23.102)$$

wobei wir noch  $\star(\star(\omega)) = (-1)^{k(n-k)}\omega$  benutzt haben wie man am leichtesten aus (23.98) folgert.

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 21.13 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

### § 23. INTEGRALSÄTZE

$b_1, \dots, b_k \in \mathbb{R}^n$  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 21 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.

$$\tilde{b}_j = \sum_i R_j^i b_i \quad (23.103)$$

mit  $\det(R_j^i) = 1$ , und

$$\text{vol}^{(k)}(b_1, \dots, b_k) = \prod_{j=1}^k \|\tilde{b}_j\|_2 = \sqrt{\det(\text{diag}(\langle \tilde{b}_j, \tilde{b}_j \rangle))} = \sqrt{\det g} \quad (23.104)$$

wobei wir die Multiplikativität der Determinante benutzt haben,  $\det R = 1$ , und die sog. Gramsche Matrix gegeben ist durch

$$g = (g_{ij})_{i,j=1,\dots,k} = (\langle b_i, b_j \rangle)_{i,j=1,\dots,k} \quad (23.105)$$

Das Volumen (23.104) erfüllt nun tatsächlich die obigen Bedingungen, ist aber im Gegensatz zu (23.23) 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

$$\tilde{b}_j = \sum_i R_j^i b_i \quad (23.106)$$

transformiert sich die Gramsche Determinante gemäss

$$\sqrt{\det \tilde{g}} = |\det R| \cdot \sqrt{\det g} \quad (23.107)$$

· 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

$$g_{ij}(x) = \langle D\psi(x)(e_i), D\psi(x)(e_j) \rangle = \sum_{r=1}^n B_i^r(x) B_j^r(x) \quad (23.108)$$

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  $(h_\iota)$ , und definiert

$$\text{vol}^{(k)}(M) := \sum_\iota \int_{V_\iota} h_\iota(\psi_\iota(x)) \cdot \sqrt{\det g(x)} dx \quad (23.109)$$

· Wegen (23.107) 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

$$\begin{aligned} \int_V h \cdot \tilde{h} \cdot \sqrt{\det g(x)} dx &= \int_{\tilde{V}} h \cdot \tilde{h} \cdot \sqrt{\det g(x)} \cdot |\det D(\varphi \circ \tilde{\varphi}^{-1})(\tilde{x})| d\tilde{x} \\ &= \int_{\tilde{V}} h \cdot \tilde{h} \sqrt{\det \tilde{g}(\tilde{x})} d\tilde{x} \end{aligned} \quad (23.110)$$

Denn der Basiswechsel in  $T_{\psi(x)}M = T_{\tilde{\psi}(\tilde{x})}M$  ist bestimmt durch

$$D\tilde{\psi}(\tilde{x}) = D\psi(x) \cdot D(\varphi \circ \tilde{\varphi}^{-1})(\tilde{x}) \quad (23.111)$$

oder anders gesagt gilt:

$$\tilde{g}_{ij}(\tilde{x}) = \sum_{r,s=1}^k \frac{\partial \varphi^r}{\partial \tilde{x}^i}(\tilde{x}) \frac{\partial \varphi^s}{\partial \tilde{x}^j}(\tilde{x}) \cdot g_{rs}(x) \quad (23.112)$$

Beispiel: Für den Möbius-Streifen in der Parametrisierung (23.54) ist

$$B = \begin{pmatrix} -\sin \vartheta (1 - s \sin \frac{\vartheta}{2}) - \frac{s}{2} \cos \vartheta \cos \frac{\vartheta}{2} & -\cos \vartheta \sin \frac{\vartheta}{2} \\ \cos \vartheta (1 - s \sin \frac{\vartheta}{2}) - \frac{s}{2} \sin \vartheta \cos \frac{\vartheta}{2} & -\sin \vartheta \sin \frac{\vartheta}{2} \\ -\frac{s}{2} \sin \frac{\vartheta}{2} & \cos \frac{\vartheta}{2} \end{pmatrix} \quad (23.113)$$

$$g = \begin{pmatrix} 1 + \frac{3s^2}{4} - \frac{s^2}{2} \cos \vartheta - 2s \sin \frac{\vartheta}{2} & 0 \\ 0 & 1 \end{pmatrix} \quad (23.114)$$

und der Flächeninhalt

$$\int_0^{2\pi} \int_{-1/2}^{1/2} \sqrt{1 + \frac{3s^2}{4} - \frac{s^2}{2} \cos \vartheta - 2s \sin \frac{\vartheta}{2}} ds d\vartheta \approx 6,353271 \dots \quad (23.115)$$

· Der Laplace-Operator auf einer Untermannigfaltigkeit

$$\Delta = \frac{1}{\sqrt{\det g}} \sum_{i,j=1}^k \frac{\partial}{\partial x^i} \left( g^{ij} \sqrt{\det g} \frac{\partial}{\partial x^j} \right) \quad (23.116)$$

2. Nach unserer Definition (23.42) 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

## § 24. VERVOLLSTÄNDIGUNG

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 abhängt: 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

$$\tilde{w}_i = D(\tilde{\varphi} \circ \varphi^{-1})(\varphi(p_i))(w_i) \quad (23.117)$$

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*  $\gamma$ , falls sie in erster Ordnung nur in den Richtungen orthogonal zu  $M$  variiert, d.h.

$$\mathbb{R}^n \ni \dot{v}(t) \perp T_{\gamma(t)}M \subset \mathbb{R}^n \quad \forall t \in [t_1, t_2] \quad (23.118)$$

Für vorgegebenes  $\gamma$  ist dies eine gewöhnliche Differentialgleichung, deren Lösung mit  $v(t_1) = v_1$  den “Paralleltransport von  $v_1 \in T_{p_1}M$  entlang  $\gamma$  nach  $p_2$ ” definiert. (Das Ergebnis hängt allerdings i.A. von  $\gamma$  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.

## § 24 Vervollständigung

Die physikalisch interessantesten Aussagen aus der linearen Algebra, die wir im Kapitel 7 für endliche Dimension bewiesen haben, übertragen sich auf den (physikalisch noch interessanteren) Fall unendlich-dimensionaler Vektorräume, aber *nicht ohne weitere Anstrengungen* und zusätzliche Voraussetzungen. Die Schwierigkeiten entstehen dadurch, dass einerseits die metrische Vollständigkeit normierter Räume (im Sinne von 4.9) 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.

Es existieren daher auch verschiedene Konvergenzbegriffe und wir müssen uns darüber Rechenschaft ablegen, welcher davon für die jeweilige Anwendung der “richtige” ist.<sup>36</sup>

---

<sup>36</sup>Es ist intuitiv nicht ganz einfach zu sagen, welche physikalischen Prinzipien bei der Auswahl

Ein verwandtes Problem tritt bei der Lösung von linearen und nicht-linearen Differentialgleichungen auf. Während wir für gewöhnliche Differentialgleichungen im §15 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.

### Monotone Hüllen

Die grundlegende Idee hinter den folgenden Konstruktionen ist ähnlich wie beim Übergang von den rationalen zu den reellen Zahlen die *Vervollständigung* von Funktionenräumen unter geeigneten Konvergenzbegriffen. Wie bereits im Zusammenhang mit der Integration stetiger Funktionen über kompakte Teilmengen angedeutet, stellen wir dabei die *monotone Konvergenz* in den Vordergrund. Dies ist analog zur Konstruktion der reellen Zahlen via Dedekindscher Schnitte.

· Unser Ausgangspunkt diesmal ist das im §22 eingeführte Integral für stetige Funktionen mit kompaktem Träger, welches wir jetzt auffassen als lineare Abbildung (Funktional)

$$J : \mathcal{C}_c(\mathbb{R}^n) \rightarrow \mathbb{R}, \quad J(f) := \int_{\mathbb{R}^n} f(x) d^n x \quad (24.1)$$

Wir schreiben wieder  $f_k \uparrow f$  (bzw.  $f_k \downarrow f$ ) für punktweise monotone Konvergenz (von Funktionen etc.). Die in 22.15 festgestellte Eigenschaft

$$J(f_k) \uparrow J(f) \text{ für } (f_k) \subset \mathcal{C}_c(\mathbb{R}^n), f_k \uparrow f \text{ und } f \in \mathcal{C}_c(\mathbb{R}^n), \quad (24.2)$$

auch geschrieben als  $\sup(J(f_k)) = J(\sup f_k)$ , bezeichnen wir als *Ordnungstetigkeit* von  $J$ .

**Definition 24.1.**  $\forall f, g \in \mathcal{C}_c(\mathbb{R}^n), \lambda \in \mathbb{R}$  sind auch  $\lambda f + g, |f|, |g|$  (punktweise definiert) und daher auch

$$\begin{aligned} \max(f, g) &= \frac{f + g + |f - g|}{2} \\ \min(f, g) &= \frac{f + g - |f - g|}{2} \end{aligned} \quad (24.3)$$

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*).<sup>37</sup> Wir schreiben ab jetzt  $\mathcal{V}$  für den Verband  $\mathcal{C}_c(\mathbb{R}^n)$ .

---

helfen sollen. Ein wichtiges Rationalitätskriterium ist sicherlich die Möglichkeit der “beliebig genauen Approximation durch endlich viele Schritte”.

<sup>37</sup>Eine 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.

§24. VERVOLLSTÄNDIGUNG

· Eine reellwertige Funktion (man sagt auch Funktional) auf einem Verband, welche (wie  $J$ ) linear, monoton, und ordnungsstetig ist, heisst *Daniell-Integral*.

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  $(x_k)$

$$\lim_{k \rightarrow \infty} x_k = \sup(x_k) = \infty \quad \left( \lim_{k \rightarrow \infty} x_k = \inf(x_k) = -\infty \right) \quad (24.4)$$

schreiben zu können, erweitern wir (zunächst als Mengen)

$$\mathbb{R}^+ := \mathbb{R} \cup \{\infty\} \quad (\mathbb{R}^- := \mathbb{R} \cup \{-\infty\}) \quad (24.5)$$

(Achtung: Wir erweitern zunächst noch nicht gleichzeitig mit  $\infty$  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 (24.3) zum Absolutbetrag setzt sich aber nicht auf  $\mathbb{R}^+$  fort.

**Definition 24.2.** Unter Benutzung dieses erweiterten Konvergenzbegriffs führen wir die *monotonen Hüllen* des Verbandes  $\mathcal{V}$  ein:

$$\begin{aligned} \mathcal{V}^+ &:= \{f : \mathbb{R}^n \rightarrow \mathbb{R}^+ := \mathbb{R} \cup \{\infty\} \mid \exists (f_k) \subset \mathcal{V} \mid f_k \uparrow f\} \\ \mathcal{V}^- &:= \{f : \mathbb{R}^n \rightarrow \mathbb{R}^- := \mathbb{R} \cup \{-\infty\} \mid \exists (f_k) \subset \mathcal{V} \mid f_k \downarrow f\} \end{aligned} \quad (24.6)$$

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.<sup>38</sup>

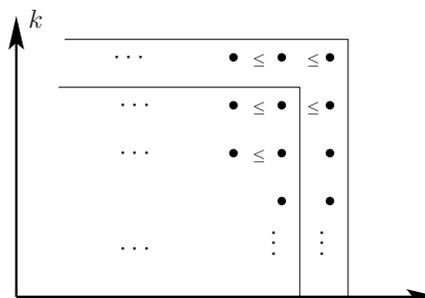
Beh.: Die monotone Hülle  $\mathcal{V}^+$  ist monoton abgeschlossen, d.h.: Ist  $(f_k) \subset \mathcal{V}^+$  eine monoton wachsende Folge von Funktionen in  $\mathcal{V}^+$ , so ist  $f := \lim_{k \rightarrow \infty} f_k \in \mathcal{V}^+$  (und sinngemäss für  $\mathcal{V}^-$ ).

*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  $l \in \mathbb{N}$

$$g_l := \max\{f_{kl} \mid k \leq l\} \quad (24.7)$$

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

$$\begin{aligned} g_{l+1} &= \max\{f_{k,l+1} \mid k \leq l+1\} \\ &\geq \max\{f_{kl} \mid k \leq l+1\} \\ &\geq \max\{f_{kl} \mid k \leq l\} \\ &= g_l \end{aligned} \quad (24.8)$$



<sup>38</sup>Wir könnten auch "Riesz-Kegel" sagen.

Ausserdem ist

$$\begin{aligned} \sup(g_l) &= \sup\{f_{kl} \mid k \leq l, l \in \mathbb{N}\} \\ &= \sup\{f_{kl} \mid k, l \in \mathbb{N}\} = \sup\{f_k \mid k \in \mathbb{N}\} = f \end{aligned} \quad (24.9)$$

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

**Proposition 24.3.** 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 Kegelverband  $\mathcal{V}^+$ , 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 24.4.** 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  $J(g) \leq \sup_k J(f_k)$ .  $\square$

*Beweis von 24.3.* 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 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

$$\lim_{k \rightarrow \infty} J(f_k) = \lim_{k \rightarrow \infty} J^+(f_k) = J^+(f) \quad (24.10)$$

sein. Um andererseits  $J^+$  durch die Formel (24.10) 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  $\forall k$

$$\tilde{f}_k \leq f = \sup_l (f_l) \xrightarrow{24.4} J(f_k) \leq \sup_l J(f_l) \quad (24.11)$$

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 (24.10) die geforderten Eigenschaften nachzuweisen.

(i) Kegellinearität folgt aus der Linearität von  $J$ :

$$J^+(\lambda f + g) = \sup J(\lambda f_k + g_k) = \lambda \sup J(f_k) + \sup J(g_k) = \lambda J^+(f) + J^+(g) \quad (24.12)$$

## § 24. VERVOLLSTÄNDIGUNG

(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 24.4

$$J(g_l) \leq \sup J(f_k) \Rightarrow J^+(g) \leq J^+(f) \quad (24.13)$$

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 24.2)  $g_l \uparrow f$  und daher gilt

$$\sup J^+(f_k) = \sup J(f_{kl}) = \sup J(g_l) = J^+(f) \quad (24.14)$$

□

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.

Z.B. liegen charakteristische Funktionen von Quadern und von kompakten Mengen in  $\mathcal{V}^-$ , etc. wie wir in Beispiel 22.16 gesehen haben.

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

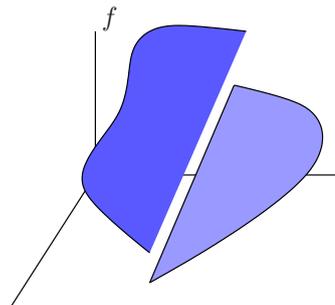
### Integrierbare Funktionen

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 24.5.** 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

$$f(y) > C \quad \text{falls } \|x - y\| < \delta \quad (24.15)$$

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

$$f(y) > f(x) - \epsilon \quad \text{falls } \|x - y\| < \delta \quad (24.16)$$

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

$$|f_K(x) - f_K(y)| < \frac{\epsilon}{2} \quad \text{falls } \|x - y\| < \delta \quad (24.17)$$

d.h. insbesondere

$$f_K(y) > f_K(x) - \frac{\epsilon}{2} \quad (24.18)$$

Zusammen folgt

$$f(y) \geq f_K(y) > f_K(x) - \frac{\epsilon}{2} > f(x) - \epsilon \quad \text{falls } \|x - y\| < \delta \quad (24.19)$$

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

$$f(y) > N \quad \text{falls } \|x - y\| < \delta \quad (24.20)$$

und der Beweis ist sehr ähnlich. □

· 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

$$f_k(x) := \inf\{f(y) + k\|x - y\| \mid y \in \mathbb{R}^n\} \quad (24.21)$$

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

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

$$f \in \mathcal{V}^+ \Leftrightarrow f = \sup\{g \mid g \in \mathcal{C}_c(\mathbb{R}^n), g \leq f\} \quad (24.22)$$

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

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

· Weiterhin gilt:

$$f \in \mathcal{V}^+ \Leftrightarrow -f \in \mathcal{V}^- \quad (24.23)$$

(mit geeigneter Deutung  $-1 \cdot \infty = -\infty$ , s. (24.27)), und für  $f \in \mathcal{V}^+$  ist

$$J^-(-f) = -J^+(f) \quad (24.24)$$

· 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

$$f(0) \cdot \chi_{B_k(0)} \uparrow f(0) \quad \text{für } k \rightarrow \infty, \quad (24.25)$$

$\chi_{B_k(0)} \in \mathcal{V}^+$ , 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  $\infty$ . (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. 1.12) 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

## § 24. VERVOLLSTÄNDIGUNG

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

$$\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\} \quad (24.26)$$

Mit den Verabredungen:

$$\begin{aligned} \infty + x &= \infty \quad \forall x \in \mathbb{R}^+, & -\infty + x &= -\infty \quad \forall x \in \mathbb{R}^- \\ x \cdot \infty &= \infty \text{ und } x \cdot (-\infty) = -\infty \quad \forall x \in \mathbb{R}, x > 0 & \text{ sowie } -1 \cdot \infty &= -\infty \\ \infty \cdot \infty &= \infty, \quad \infty \cdot (-\infty) = -\infty, & (-\infty) \cdot (-\infty) &= \infty \end{aligned} \quad (24.27)$$

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,

$$0 \cdot \infty = 0, \quad \infty - \infty = -\frac{1}{12} \quad (24.28)$$

**Definition 24.6.** 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  $f \in \overline{\mathcal{F}}(\mathbb{R}^n) := \mathcal{F}(\mathbb{R}^n, \overline{\mathbb{R}})$ <sup>39</sup> heisst dann

$$\bar{J}^+(f) := \inf\{J^+(g) \mid g \in \mathcal{V}^+, g \geq f\} \quad (24.29)$$

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

$$\bar{J}^-(f) := \sup\{J^-(h) \mid h \in \mathcal{V}^-, h \leq f\} \quad (24.30)$$

das *Unterintegral* von  $f$ .<sup>40</sup>

• 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

$$\bar{J}(f) := \bar{J}^+(f) = \bar{J}^-(f) \in \mathbb{R} \quad (24.32)$$

ihr *Integral*. Wir schreiben  $\overline{\mathcal{V}}$  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  $\overline{\mathcal{L}}(\mathbb{R}^n)$  statt  $\overline{\mathcal{V}}$  sowie

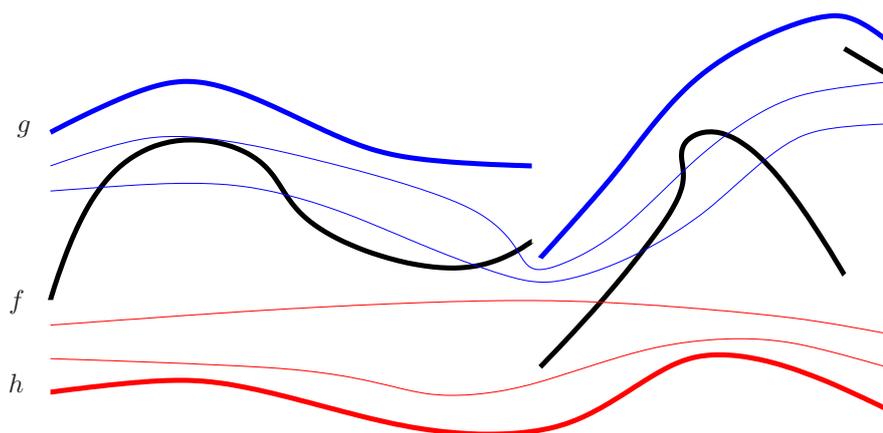
$$\int_{\mathbb{R}^n} f(x) dx \quad \left( \text{manchmal auch } \int_{\mathbb{R}^n} f(x) d^n x \right) \quad \text{statt } \bar{J}(f) \quad (24.33)$$

<sup>39</sup>Beachte, dass ohne (24.28)  $\overline{\mathcal{F}}(\mathbb{R}^n)$  kein Vektorraum ist.

<sup>40</sup>Wegen  $\pm\infty \in \mathcal{V}^\pm$ , s. (24.25), 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  $\infty > x > -\infty \quad \forall x \in \mathbb{R}$

$$\sup(\emptyset) = \text{“kleinste obere Schranke”} = -\infty \quad \text{und} \quad \inf(\emptyset) = \infty \quad (24.31)$$

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



**Lemma 24.7.** (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

$$J^+(g_\epsilon) - J^-(h_\epsilon) < \epsilon \tag{24.34}$$

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

Bew.: Gemäss (24.24) ist  $-h \in \mathcal{V}^+$  und  $-J^-(h) = J^+(-h)$ . Mit der Kegellinearität und der Monotonie von  $J^+$ , s. 24.3, folgt

$$J^+(g) - J^-(h) = J^+(g) + J^+(-h) = J^+(\underbrace{g - h}_{\geq 0}) \geq 0 \tag{24.35}$$

und daraus die Behauptung.

(i) Aus dem Hilfslemma folgt

$$\{J^+(g) \mid g \in \mathcal{V}^+, g \geq f\} \geq \{J^-(h) \mid h \in \mathcal{V}^-, h \leq f\} \tag{24.36}$$

und durch inf/sup-Bildung daraus die Behauptung.

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

$$\bar{J}^+(f) = J^+(f) \quad \text{und} \quad \bar{J}^-(f) \geq J^+(f) \tag{24.37}$$

Die erste Aussage folgt aus der Ungleichungskette

$$J^+(f) \leq \inf\{J^+(g) \mid g \in \mathcal{V}^+, g \geq f\} \leq J^+(f) \tag{24.38}$$

$\uparrow$   $\uparrow$   
 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

$$\bar{J}^-(f) = \sup\{J^-(h) \mid h \in \mathcal{V}^-, h \leq f\} \geq \sup_k J^-(f_k) = \sup_k J(f_k) = J^+(f) \tag{24.39}$$

## § 25. EIGENSCHAFTEN DES LEBESGUE-INTEGRALS

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

$$\bar{J}^+(f) = \inf\{J^+(g) \mid g \in \mathcal{V}^+, g \geq f\} \leq J^+(g_\epsilon) < J^-(h_\epsilon) + \epsilon < \infty \quad (24.40)$$

Zusammen mit

$$\bar{J}^-(f) \geq J^-(h_\epsilon) \quad (24.41)$$

folgt

$$\bar{J}^+(f) \leq \bar{J}^-(f) + \epsilon \quad (24.42)$$

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 □

## § 25 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, und machen ein paar Bemerkungen zu seiner Einzigartigkeit. Zunächst halten wir fest, in welchem Sinne wir mit dem Funktional  $\bar{J}$  auf  $\overline{\mathcal{V}}$  unser Ziel erreicht haben.

**Proposition/Definition 25.1.** *Der Raum*

$$\mathcal{L}^1(\mathbb{R}^n) := \{f \in \overline{\mathcal{L}}(\mathbb{R}^n) \mid f(x) \in \mathbb{R} \forall x \in \mathbb{R}^n\} \quad (25.1)$$

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 monotonies Funktional, welches wir als

$$\mathcal{L}^1(\mathbb{R}^n) \ni f \mapsto \int_{\mathbb{R}^n} f(x) dx := \bar{J}(f) \quad (25.2)$$

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

$$\begin{aligned} \bar{J}^+(f_1) + \bar{J}^+(f_2) &= \inf\{J^+(g_1) \mid g_1 \in \mathcal{V}^+, g_1 \geq f_1\} + \inf\{J^+(g_2) \mid g_2 \in \mathcal{V}^+, g_2 \geq f_2\} \\ &= \inf\{J^+(g_1 + g_2) \mid g_1, g_2 \in \mathcal{V}^+, g_1 + g_2 \geq f_1 + f_2\} \\ &\geq \inf\{J^+(g) \mid g \in \mathcal{V}^+, g \geq f_1 + f_2\} = \bar{J}^+(f_1 + f_2) \end{aligned} \quad (25.3)$$

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

$$\bar{J}^-(f_1 + f_2) \geq \bar{J}(f_1) + \bar{J}(f_2) \geq \bar{J}^+(f_1 + f_2) \quad (25.4)$$

Mit Lemma 24.7 (i) folgt  $f_1 + f_2 \in \overline{\mathcal{V}}$  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 (24.24)

$$\begin{aligned} \bar{J}^+(f) &= \inf\{J^+(g) \mid g \in \mathcal{V}^+, g \geq f\} = -\sup\{J^-(-g) \mid g \in \mathcal{V}^+, g \geq f\} \\ &= -\sup\{J^-(h) \mid h \in \mathcal{V}^-, h \leq -f\} = -\bar{J}^-(-f) \end{aligned} \quad (25.5)$$

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 (24.3),

$$|f| = \max(f, 0) - \min(f, 0) \quad (25.6)$$

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 24.7 (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 Kegelvebandsstruktur von  $\mathcal{V}^+$ , dass  $g_+ := \max(g, 0)$ ,  $g_- := \min(g, 0) \in \mathcal{V}^+$  und ebenso  $h_\pm \in \mathcal{V}^-$ . Wegen der Kegelinearität und (24.24) folgt

$$\begin{aligned} \epsilon > J^+(g) - J^-(h) &= J^+(g_+) + J^+(g_-) + J^+(-h_+) + J^+(-h_-) \\ &= J^+(g_+ - h_+) + J^+(g_- - h_-) \end{aligned} \quad (25.7)$$

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 (25.7) nicht-negativ und daher *getrennt*  $< \epsilon$  sind. Damit gilt (wieder mit (24.24))

$$J^+(g_+) - J^-(h_+) < \epsilon \quad (25.8)$$

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

*Bemerkungen.* Die Abschätzung (25.3) und (25.4) gelten offenbar weiterhin für integrierbare Funktionen  $f_1, f_2 \in \bar{\mathcal{L}}(\mathbb{R}^n)$  mit Werten in den erweiterten reellen Zahlen, wenn  $f_1 + f_2$  an den Unendlichkeitsstellen mit den Regeln (24.27) bzw. für  $\infty - \infty$  beliebig definiert ist. Jede solche Funktion ist also integrierbar und hat das gleiche Integral. In einem ähnlichen Sinne ist  $\bar{\mathcal{L}}(\mathbb{R}^n)$  auch unter reeller Skalarmultiplikation und unter  $\min / \max$  abgeschlossen.

**Korollar 25.2.** 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

$$\int_{\mathbb{R}^n} \varphi(x) dx = I_{\mathcal{T}}(\varphi) \quad \text{und} \quad \int_{\mathbb{R}^n} f(x) dx = I_{\mathcal{C}_c}(f) \quad (25.9)$$

*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 24.7. Für  $\mathcal{T}(\mathbb{R}^n)$  folgen die Aussagen aus dem Beispiel 22.13 und Gl. (22.46). □

## § 25. EIGENSCHAFTEN DES LEBESGUE-INTEGRALS

*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 (14.34), falls Letzteres nur bedingt und nicht absolut konvergiert. Beispiel:

$$\int_{-\infty}^{\infty} \frac{\sin x}{x} dx = \lim_{\alpha \rightarrow \infty} \int_0^{\alpha} + \lim_{\beta \rightarrow -\infty} \int_{\beta}^0 = \pi \quad (25.10)$$

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

$$\int_0^{N\pi} \left| \frac{\sin x}{x} \right| dx \geq \sum_{k=0}^{N-1} \frac{1}{k+1} \int_{k\pi}^{(k+1)\pi} |\sin x| dx = \sum_{k=0}^{N-1} \frac{2}{k+1} \rightarrow \infty \text{ für } N \rightarrow \infty \quad (25.11)$$

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

### Vertauschungssätze I und Anwendungen

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 25.3.** *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. 25.1) 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

$$\begin{aligned} \bar{J}^-(f) &= \sup \{ J^-(h) \mid h \in \mathcal{V}^-, h \leq f \} \\ &\geq \sup \{ J^-(h) \mid h \in \mathcal{V}^-, h \leq f_k \text{ für ein } k \in \mathbb{N} \} \\ &= \sup_k \bar{J}^-(f_k) = \sup_k \bar{J}(f_k) = S \end{aligned} \quad (25.12)$$

(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

$$\delta_k \geq 0 \forall k > 1, \quad \sum_{k=1}^{\infty} \delta_k = f \quad \text{und} \quad \sum_{k=1}^{\infty} \bar{J}(\delta_k) = S. \quad (25.13)$$

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

$$\bar{J}^+(f) \leq J^+\left(\sum_{k=1}^{\infty} g_k\right) = \sum_{k=1}^{\infty} J^+(g_k) \leq \sum_{k=1}^{\infty} \left(\bar{J}(\delta_k) + \frac{\epsilon}{2^k}\right) = S + \epsilon \quad (25.14)$$

(geometrische Reihe!). Da dies für jedes  $\epsilon > 0$  gilt, folgt die behauptete Ungleichung. Zusammen mit 24.7 (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. 48).

**Theorem 25.4.** 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  $|f_k| \leq F \ \forall k$ .<sup>41</sup> Dann ist die punktweise gebildete Grenzfunktion  $f(x) := \lim_{k \rightarrow \infty} f_k(x)$  integrierbar mit

$$\int_{\mathbb{R}^n} f(x) dx = \lim_{k \rightarrow \infty} \int_{\mathbb{R}^n} f_k(x) dx \quad (25.15)$$

*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  $h_k \in \bar{\mathcal{L}}(\mathbb{R}^n)$ .

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  $h_{kj} \in \bar{\mathcal{L}}(\mathbb{R}^n)$ . ( $\bar{\mathcal{L}}(\mathbb{R}^n)$  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 25.3 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  $\lim_{k \rightarrow \infty} \bar{J}(f_k) = \bar{J}(f)$ . Dies folgt aber aus  $h_k \leq f_k \leq g_k$  wegen der Monotonie von  $\bar{J}$  und des Sandwich-Lemmas 3.8.  $\square$

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 14.34 “uneigentlich” definiert.) Wir werden weiterhin nicht viel mehr brauchen als die Aussage, dass für  $f \in \mathcal{L}^1(\mathbb{R}^n)$  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

$$\int_U f(x) dx := \int_{\mathbb{R}^n} f(x) \cdot \chi_U(x) dx \quad \text{und} \quad \int_A f(x) dx := \int_{\mathbb{R}^n} f(x) \cdot \chi_A(x) dx. \quad (25.16)$$

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. 234). 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  $f \cdot \gamma_k \in \mathcal{L}^1(\mathbb{R}^n)$ ,  $f \cdot \gamma_k \uparrow f \cdot \chi_U$ , 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 (25.16) 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  $x \notin U$ .

Hier noch einmal festhalten:

<sup>41</sup>Es genügt offenbar  $F \in \bar{\mathcal{F}}(\mathbb{R}^n)$  mit  $|f_k| \leq F$  und  $\bar{J}^+(F) < \infty$ .

§ 25. EIGENSCHAFTEN DES LEBESGUE-INTEGRALS

**Theorem 25.5.** *Es sei  $\Phi : \Upsilon \rightarrow U$  ein Diffeomorphismus zwischen offenen Teilmengen des  $\mathbb{R}^n$ , d.h.  $\Phi$  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  $\Upsilon$  integrierbar und es gilt*

$$\int_U f(x)dx = \int_{\Upsilon} f(\Phi(\xi)) |\det D\Phi(\xi)| d\xi \tag{25.17}$$

**Mass und Integral**

\*\*\* Benötigt wird nur der Begriff der Nullmenge \*\*\*

- 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*, wie sie auf S. 194 angedeutet wurde.
- Zur Definition eines Integrals auf einer Teilmenge von  $\mathcal{F}(\mathbb{R}^n, \mathbb{R})$  ordnet man dabei 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 24.6.)
- 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 25.6.** 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

$$\text{vol}(X) := \int_{\mathbb{R}^n} \chi_X(x)dx \tag{25.18}$$

- Dass beschränkte abgeschlossene und offene Mengen endlich messbar sind, hatten wir bereits festgehalten. Die Eigenschaft (22.67) folgt aus der Linearität des Integrals und der Identität

$$\chi_{A_1} + \chi_{A_2} = \chi_{A_1 \cap A_2} + \chi_{A_1 \cup A_2} \tag{25.19}$$

bzw. verbandstheoretisch aus (24.3) zusammen mit Beziehungen der Art

$$\begin{aligned} \chi_{A_1 \cup A_2} &= \max(\chi_{A_1}, \chi_{A_2}), & \chi_{A_1 \cap A_2} &= \min(\chi_{A_1}, \chi_{A_2}) = \chi_{A_1} \cdot \chi_{A_2} \\ |\chi_{A_1} - \chi_{A_2}| &= \chi_{A_1 \setminus A_2} + \chi_{A_2 \setminus A_1} \end{aligned} \tag{25.20}$$

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 22.21).

### Nullfunktionen und Nullmengen

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

**Definition 25.7.** (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  $\chi_N$  eine Nullfunktion ist.

(Wegen  $\chi_N = |\chi_N|$  und  $0 \leq \bar{J}^-(\chi_N) \leq \bar{J}^+(\chi_N)$  ist die charakteristische Funktion  $\chi_N$  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 25.8.** *Es sei  $f \in \overline{\mathcal{L}}(\mathbb{R}^n)$ .*

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

$$N_\infty := \{x \in \mathbb{R}^n \mid f(x) = \infty \text{ oder } -\infty\} \quad (25.21)$$

*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 (24.28) ist  $\chi_{N_\infty} = -12(f + (-f))$  und nach den Bemerkungen zu Theorem 25.1 bzw. aus der Abschätzung (25.3) 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 25.3 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

$$\bar{J}^+(g) \leq \bar{J}^+(f) = \bar{J}^-(f) \leq \bar{J}^-(g) \quad (25.22)$$

und mit 24.7 (i) die Behauptung. □

**Beispiel 25.9.** · 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

§ 25. EIGENSCHAFTEN DES LEBESGUE-INTEGRALS

Funktion der Koordinatenebene  $\{x_1 = 0\} \subset \mathbb{R}^n$ , so dass aus 25.3 folgt, dass das  $n$ -dimensionale Volumen von  $\mathbb{R}^{n-1}$  verschwindet.

$$\int_{\mathbb{R}^n} \chi_{\{x_1=0\}} d^n x = 0 \tag{25.23}$$

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

$$\max\{\chi_{N_l} \mid l \leq k\} \tag{25.24}$$

monoton gegen  $\chi_{\cup_k N_k}$ , so dass wieder aus 25.3 folgt, dass

$$\cup_{k=1}^{\infty} N_k = N_1 \cup N_2 \cup N_3 \cup \dots \tag{25.25}$$

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 wählen<sup>42</sup> 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

$$\mathbb{R} = \bigcup_{x \in X} (x + \mathbb{Q}) = \bigcup_{q \in \mathbb{Q}} (q + X) \tag{25.26}$$

Beh.:  $\chi_X$  ist nicht integrierbar.

Bew.: Wäre  $\chi_X$  integrierbar, so folgte aus der Translationsinvarianz des Integrals/Maßes, dass auch alle  $\chi_{q+X}$  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

$$\bigcup_{q \in \mathbb{Q} \cap [0, 1]} (q + X) \subset [0, 2) \tag{25.27}$$

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.

**Ein Banach-Raum**

Fazit aus 25.8: Man darf eine integrierbare Funktion auf einer Nullmenge beliebig abändern, ohne an der Integrierbarkeit oder dem Wert des Integrals etwas zu ändern.

· Ist insbesondere  $f \in \overline{\mathcal{L}}(\mathbb{R}^n)$ , so existiert ein  $h \in \mathcal{L}^1(\mathbb{R}^n)$  mit  $h = f$  fast überall (und insbesondere  $\int h(x) dx = \int f(x) dx$ ).

· Ist  $\tilde{h}$  eine andere Funktion mit  $\tilde{h} = f$  f.ü., dann ist  $h - \tilde{h}$  eine Nullfunktion, m.a.W.,  $h \sim f \Leftrightarrow h = f$  f.ü. ist eine Äquivalenzrelation und die Menge der Nullfunktionen ist ein Untervektorraum von  $\mathcal{L}^1(\mathbb{R}^n)$ .

---

<sup>42</sup>mit dem Auswahlaxiom

**Definition 25.10.**

$$L^1(\mathbb{R}^n) := \mathcal{L}^1(\mathbb{R}^n)/\text{Nullfunktionen} = \{[f] = \{h \in \mathcal{L}^1(\mathbb{R}^n) | h = f \text{ f.ü.}\}\}$$

**Proposition 25.11.** ·  $L^1(\mathbb{R}^n)$  ist ein Vektorverband, insbesondere mit  $\|[f]\| = \|[f]\|$ .

·  $\int [f] := \int h$  für ein  $h \in [f]$  ist wohldefiniert.

· Die Abbildung  $[f] \rightarrow \int \|[f]\| =: \|f\|_1$  ist eine Norm auf  $L^1(\mathbb{R}^n)$ .

·  $(L^1(\mathbb{R}^n), \|\cdot\|_1)$  ist Cauchy-vollständig, m.a.W. ein Banach-Raum (s. Def. 27.1).

*Beweis.* Wird nachgereicht □

*Bemerkungen.* · Ist  $f \in \mathcal{C}(\mathbb{R}^n) \cap \mathcal{L}^1(\mathbb{R}^n)$ , so ist  $f$  der einzige stetige Repräsentant in  $[f] \in L^1(\mathbb{R}^n)$ . In diesem Sinne ist insbesondere  $\mathcal{C}_c(\mathbb{R}^n) \hookrightarrow L^1(\mathbb{R}^n)$ .

·  $\mathcal{C}_c(\mathbb{R}^n)$  ist dicht in  $L^1(\mathbb{R}^n)$ :  $\forall [f] \in L^1(\mathbb{R}^n), \forall \epsilon > 0$  existiert ein  $h \in \mathcal{C}(\mathbb{R}^n)$  so dass

$$\|h - [f]\|_1 = \int_{\mathbb{R}^n} |h - f| < \epsilon \tag{25.28}$$

· Man schreibt Elemente von  $L^1(\mathbb{R}^n)$  auch sonst nur als  $f$  (und denkt sich die Äquivalenzklasse dazu).

**Vertauschungssätze II**

Für die Berechnung des Integrals bleibt die Verallgemeinerung von 22.12 ein wichtiges Hilfsmittel.

**Theorem 25.12.** 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 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

$$\int_{\mathbb{R}^m} \left( \int_{\mathbb{R}^n} f(x, y) dx \right) dy = \int_{\mathbb{R}^{n+m}} f(x, y) d(x, y) = \int_{\mathbb{R}^n} \left( \int_{\mathbb{R}^m} f(x, y) dy \right) dx \tag{25.29}$$

Ausserdem gilt in Verallgemeinerung von 14.8:

**Theorem 25.13.** 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

$$y \mapsto \int_{\mathbb{R}^n} f(x, y) dx \tag{25.30}$$

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 letztinge  $\forall y$  integrierbar, die Funktion in (25.30) ist stetig differenzierbar in Richtung  $w$  und es gilt

$$D_w \left( \int_{\mathbb{R}^n} f(x, y) dx \right) = \int_{\mathbb{R}^n} D_w f(x, y) dx \tag{25.31}$$

## § 25. EIGENSCHAFTEN DES LEBESGUE-INTEGRALS

*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 14.3 mit 22.10, 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.  $\overline{\mathbb{C}} := \mathbb{C} \cup \{\infty\}$  benutzen. Dann ist whatever Definition äquivalent zur Aussage:

· Eine erweitert komplexwertige Funktion  $f \in \mathcal{F}(\mathbb{R}^n, \overline{\mathbb{C}})$  ist genau dann integrierbar, wenn ihre (Unendlichkeitsstellenmenge eine Nullmenge ist und die entsprechende modifizierten) Real- und Imaginärteile im Sinne von 24.6 integrierbar sind, und es gilt

$$\int_{\mathbb{R}^n} f(x) dx = \int_{\mathbb{R}^n} \operatorname{Re}(f)(x) dx + i \int_{\mathbb{R}^n} \operatorname{Im}(f)(x) dx \quad (25.32)$$

Wir schreiben  $\mathcal{L}^1(\mathbb{R}^n, \mathbb{C})$  für die komplexwertigen integrierbaren Funktionen. Die Vertauschungssätze (ausser natürlich Beppo Levi) gelten sinngemäss.

*KAPITEL 8. INTEGRATIONSTHEORIE*

# 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.

### § 26 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 21.13.

Eine sparsame Kombination dieser Konzepte ergibt sich aus dem bereits in ?? festgehaltenen Zusammenhang zwischen innerem Produkt und Norm:

**Lemma 26.1.** *Es sei  $V$  ein komplexer Vektorraum.*

(i) *Ist  $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{C}$  ein hermitesches inneres Produkt, so definiert*

$$\|\cdot\| : V \rightarrow \mathbb{R}, \quad v \mapsto \|v\| := \sqrt{\langle v, v \rangle} \quad (26.1)$$

*eine Norm auf  $V$ .*

(ii) *Inneres Produkt und Norm erfreuen sich verschiedener komplexer Verallgemeinerungen von vertrauten geometrischen Eigenschaften. Insbesondere gilt die Cauchy-Schwarz-Ungleichung:*

$$|\langle v, w \rangle| \leq \|v\| \cdot \|w\| \quad \forall v, w \in V \quad (26.2)$$

(iii) *Eine Norm  $\|\cdot\| : V \rightarrow \mathbb{R}$  stammt genau dann von einem inneren Produkt, falls sie die Parallelogrammgleichung*

$$\|v + w\|^2 + \|v - w\|^2 = 2(\|v\|^2 + \|w\|^2) \quad (26.3)$$

*erfüllt.*

*Beweis.* Die Bedingungen an die Form  $\langle \cdot, \cdot \rangle$  sind (vgl. 21.1):  $\langle v, \mu w_1 + w_2 \rangle = \mu \langle v, w_1 \rangle + \langle v, w_2 \rangle$ ,  $\langle v, w \rangle = \langle w, v \rangle$  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

$$\|v + w\|^2 = \|v\|^2 + \underbrace{\langle v, w \rangle + \langle w, v \rangle}_{=2 \operatorname{Re}(\langle v, w \rangle)} + \|w\|^2 \quad (26.4)$$

zunächst die C.-S.-Ungleichung:

· Für  $v \neq 0$  (für  $v = 0$  ist die Aussage trivial) ist die Zerlegung

$$w = \underbrace{\langle v, w \rangle \frac{v}{\|v\|^2}}_{w_{\parallel}} + \underbrace{w - \langle v, w \rangle \frac{v}{\|v\|^2}}_{w_{\perp}} \quad (26.5)$$

§ 26. HILBERT-RÄUME

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

$$\|w\|^2 = \|w_\parallel\|^2 + \underbrace{\|w_\perp\|^2}_{\geq 0} \geq \frac{|\langle v, w \rangle|^2}{\|v\|^2} \quad (26.6)$$

und daraus die Behauptung.

· Damit ist die rechte Seite von (26.4) beschränkt durch

$$\|v + w\|^2 \leq (\|v\| + \|w\|)^2 \quad (26.7)$$

und daraus folgt die Dreiecksungleichung. Ebenso folgt die Gültigkeit der Parallelogrammgleichung für gegebenes inneres Produkt aus (26.4)  $+ \|v - w\|^2$ .

· Für die umgekehrte Richtung sei  $\|\cdot\| : V \rightarrow \mathbb{R}$  eine Norm, welche die Parallelogrammgleichung erfüllt.

Beh.:

$$h(v, w) := \frac{1}{4}(\|v + w\|^2 - \|v - w\|^2) + \frac{i}{4}(\|v - iw\|^2 - \|v + iw\|^2) \quad (26.8)$$

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

$$r(v, w) := \frac{1}{4}(\|v + w\|^2 - \|v - w\|^2) \quad (26.9)$$

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

$$r(v, -iw) = r(iv, w) = r(w, iv) = -r(w, -iv) \quad (26.10)$$

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 (26.3) folgt für  $v, w_1, w_2 \in V$ :

$$\begin{aligned} \|v + w_1 + w_2\|^2 &= -\|v + w_1 - w_2\|^2 + 2\|v + w_1\|^2 + 2\|w_2\|^2 \\ \|v - w_1 - w_2\|^2 &= -\|v + w_1 - w_2\|^2 + 2\|v - w_1\|^2 + 2\|w_1\|^2 \end{aligned} \quad (26.11)$$

und daraus

$$\begin{aligned} 4(r(v, w_1 + w_2) - r(v, w_1) - r(v, w_2)) &= \|v + w_1 + w_2\|^2 - \|v - w_1 - w_2\|^2 \\ &\quad - \|v + w_1\|^2 + \|v - w_1\|^2 - \|v + w_2\|^2 + \|v - w_2\|^2 \\ &= \|v + w_1\|^2 + \|v - w_1\|^2 - 2\|w_1\|^2 - \|v + w_2\|^2 - \|v - w_2\|^2 + 2\|w_2\|^2 \\ &= 2\|v\|^2 - 2\|v\|^2 = 0 \end{aligned} \quad (26.12)$$

$r$  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  $r(v, \alpha w) = r(v, w) \forall \alpha \in \mathbb{Q}$ , 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 26.2.** Ein Hilbertraum ist ein unitärer Raum  $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ , welcher mit der durch die Norm (26.1) induzierten Abstandsfunktion ein vollständiger metrischer Raum ist.

**Beispiel 26.3.**<sup>43</sup> 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. 4.26. 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

$$\langle z, w \rangle = \sum_{i=1}^n \bar{z}^i w^i \tag{26.13}$$

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. 21.4): Für  $v \in V$  mit  $v \neq 0$  ist

$$v^\perp := \text{Ker}(\langle v, \cdot \rangle) \tag{26.14}$$

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 (26.13): Für eine abzählbare Menge (z.B.  $\mathbb{N}$  oder  $\mathbb{Z}$ ) ist der Raum der quadratsummierbaren Folgen

$$\ell^2() := \left\{ (a_k)_{k \in \mathbb{N}} \subset \mathbb{C} \mid \sum_{k \in \mathbb{N}} |a_k|^2 < \infty \right\} \tag{26.15}$$

(Zur Erinnerung: Die Konvergenz auf der rechten Seite ist gleichbedeutend damit, dass die Menge der endlichen Teilsummen

$$\sum_k' |a_k|^2 \tag{26.16}$$

beschränkt ist.) mit dem inneren Produkt

$$\langle (a_k), (b_k) \rangle = \sum_{k \in \mathbb{N}} \bar{a}_k b_k \tag{26.17}$$

---

<sup>43</sup>Wir schreiben, unabhängig vom Vektorraum, im Folgenden immer  $\langle \cdot, \cdot \rangle$  für das innere Produkt und  $\|\cdot\|$  für die Norm.

§ 26. HILBERT-RÄUME

ein Hilbertraum: Die Wohldefiniertheit, dass nämlich Summen von quadratsummierbaren Folgen wieder quadratsummierbar sind, sowie die Konvergenz von (26.17) für  $(a_k), (b_k) \in \ell^2()$ , 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

$$((a_k^{(n)})_{k \in \mathbb{N}})_{n \in \mathbb{N}} \subset \ell^2() \quad (26.18)$$

eine Cauchy-Folge (von Folgen). Dann folgt aus  $\sum_{k \in \mathbb{N}} |a_k^{(n)} - a_k^{(m)}|^2 < \epsilon$  für  $n, m > N$  gross genug die gleiche Ungleichung für jedes  $k$ , d.h. für jedes  $k \in \mathbb{N}$  ist  $(a_k^{(n)})_{n \in \mathbb{N}}$  eine Cauchy-Folge von komplexen Zahlen. Sei

$$a_k := \lim_{n \rightarrow \infty} a_k^{(n)} \quad (26.19)$$

Für  $n, m > N$  gross genug ist dann jede endliche Teilsumme

$$\sum_k' |a_k^{(n)} - a_k^{(m)}|^2 \leq \sum_{k \in \mathbb{N}} |a_k^{(n)} - a_k^{(m)}|^2 < \epsilon \quad (26.20)$$

und mit  $m \rightarrow \infty$  folgt

$$\sum_k' |a_k^{(n)} - a_k|^2 < \epsilon \quad (26.21)$$

für jede endliche Teilsumme. Aus dem Kriterium 5.13 folgt

$$\sum_{k \in \mathbb{N}} |a_k^{(n)} - a_k|^2 \leq \epsilon \quad (26.22)$$

für  $n \geq N$  und daraus  $(a_k)_{k \in \mathbb{N}} \in \ell^2()$  (genauer:  $(a_k^{(n)} - a_k)_{k \in \mathbb{N}} \in \ell^2()$  und  $\ell^2()$  ist ein Vektorraum), sowie  $(a_k)_{k \in \mathbb{N}} = \lim_{n \rightarrow \infty} (a_k^{(n)})_{k \in \mathbb{N}}$  in der  $\ell^2$ -Norm.

· Ein “Unbeispiel” erhält man durch eine andere Limesbildung in (26.13): Der Raum der endlichen Folgen, d.h.  $\mathcal{F}^{\text{fin}}(\mathbb{C})$  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  $\ell^1$ -Norm, d.h.

$$\sum_{k \in \mathbb{N}} |a_k| \quad (26.23)$$

nicht äquivalent zu (26.17) ist: Die Folgenfolge  $a_k^{(n)} = \frac{1}{k}$  für  $k \leq n$ , 0 sonst, ist in der  $\ell^1$ -Norm keine Cauchy-Folge, in der  $\ell^2$ -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  $= \mathbb{Z}^n$  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

$$\mathcal{C}_c(\mathbb{R}^n, \mathbb{C}) = \{f : \mathbb{R}^n \rightarrow \mathbb{C} \mid f \text{ stetig und } \text{supp}(f) \text{ kompakt}\} \quad (26.24)$$

mit dem inneren Produkt

$$\langle f, g \rangle = \int_{\mathbb{R}^n} \bar{f}(x)g(x) dx \quad (26.25)$$

(s. Aufgabe 5.3 für die Positivität). Dieser unitäre Raum ist ebenfalls nicht vollständig, wobei zusätzlich zum Verhalten bei  $\infty$  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

$$\begin{aligned} b_1 &:= \frac{v_1}{\|v_1\|} \\ b_2 &:= \frac{b_2 - \langle b_1, v_2 \rangle b_1}{\|b_2 - \langle b_1, v_2 \rangle b_1\|} \\ &\vdots \end{aligned} \quad (26.26)$$

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

$$e_i = (0, \dots, 0, \underset{\substack{\uparrow \\ i\text{-te Stelle}}}{1}, 0, \dots, 0, \dots) \quad (26.27)$$

orthonormal, d.h.  $\langle e_i, e_j \rangle = \delta_{ij}$ , aber keine Basis von  $\ell^2(\mathbb{N})$  (der nicht von abzählbarer Dimension ist<sup>44</sup>), und  $E$  kann auch nicht zu einer ONB ergänzt werden, denn

$$\begin{aligned} E^\perp &= \{(a_k)_{k \in \mathbb{N}} \mid \langle e_i, (a_k)_{k \in \mathbb{N}} \rangle = 0 \ \forall i\} \\ &= \{(a_k)_{k \in \mathbb{N}} \mid a_i = 0 \ \forall i\} \\ &= \{0\} \end{aligned} \quad (26.28)$$

(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  $\ell^2$ -Norm) durch endliche Linearkombinationen der  $e_i$  approximiert werden: Die Folgen

$$a_k^{(n)} = \begin{cases} a_k & k \leq n \\ 0 & k > n \end{cases} \quad (26.29)$$

erfüllen  $(a_k^{(n)})_k \in \mathcal{F}^{\text{fin}}(\mathbb{N}, \mathbb{C}) \subset \ell^2(\mathbb{N})$  und

$$\lim_{n \rightarrow \infty} \|(a_k^{(n)} - a_k)_{k \in \mathbb{N}}\| = 0 \quad (26.30)$$

**Definition 26.4.** 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.

<sup>44</sup> $E$  ist aber eine Basis von  $\mathcal{F}^{\text{fin}}(\mathbb{N}, \mathbb{C})$ !

**Fakten 26.5.** · 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

$$\text{span}_{\mathbb{Q}(i)}\{b_k \mid k = 1, 2, \dots\} = \left\{ \sum_{k=1}^{\infty} a_k b_k \mid a_k \in \mathbb{Q} + i\mathbb{Q}, \text{ fast alle } 0 \right\} \quad (26.31)$$

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

$$\begin{aligned} i : \ell^2(\mathbb{N}) &\rightarrow \mathcal{H} \\ (a_k)_{k \in \mathbb{N}} &\mapsto \sum_{k=1}^{\infty} a_k b_k \end{aligned} \quad (26.32)$$

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

$$\{|k\rangle \mid k \in \mathbb{N}\} \quad \text{bzw.} \quad \{|\text{Realtex}\rangle\} \quad (26.33)$$

### 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

$$(\|f\|_2)^2 = \int_{\mathbb{R}^n} |f(x)|^2 dx < \infty \quad (26.34)$$

(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. (25.26) 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:  $n = 1, \alpha, \beta \in \mathbb{R}$ .

$$f(x) = \begin{cases} |x|^\alpha & 0 < |x| < 1 \\ |x|^\beta & |x| \geq 1 \\ \text{beliebig} & \text{für } x = 0 \end{cases} \quad (26.35)$$

Dann gilt:

$$\int_{\mathbb{R}} |f(x)|^2 dx < \infty \iff \beta < -\frac{1}{2} \text{ und } \alpha > -\frac{1}{2} \quad (26.36)$$

· Die meisten anderen Fälle kann man mit Majoranten auf (26.36) zurückführen. In der diskreten Version sind all Funktionen lokal beschränkt, ausserdem existieren keine Nullmengen.

**Definition 26.6.** · 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

$$f_K^B := \max(\min(f, B \cdot \chi_K), -B \cdot \chi_K) \in \mathcal{L}^1(\mathbb{R}^n) \quad (26.37)$$

( $f_K^B$  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. (26.37) 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 26.7.** · 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 26.8.** 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 25.4 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

§ 26. HILBERT-RÄUME

**Lemma 26.9.** 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 (25.4) □

**Definition 26.10.**

$$\mathcal{L}^2(\mathbb{R}^n) = \mathcal{L}^2(\mathbb{R}^n, \mathbb{C}) := \{f \in \mathcal{M}(\mathbb{R}^n, \mathbb{C}) \mid |f|^2 \in \mathcal{L}^1(\mathbb{R}^n)\}$$

**Lemma 26.11.** (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

$$(f, g) \mapsto \langle f, g \rangle := \int_{\mathbb{R}^n} \bar{f}(x)g(x)dx \in \mathbb{C} \tag{26.38}$$

ist eine (konjugiert) symmetrische Sesquilinearform.

*Beweis.* Die Aussagen folgen mit

$$|f + g|^2 \leq (|f|^2 + |g|^2)^2 \leq (2 \max(|f|, |g|))^2 \leq 4 \max(|f|^2, |g|^2) \tag{26.39}$$

bzw.  $|\bar{f} \cdot g| \leq \max(|f|^2, |g|^2)$

ähnlich wie bei Reihen aus dem Majorantenkriterium, hier 26.9. (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:

$$\langle f, g \rangle = 0 \quad \forall f \in \mathcal{L}^2(\mathbb{R}^n) \Leftrightarrow g \text{ ist eine Nullfunktion} \tag{26.40}$$

und der Raum der Nullfunktionen ist ein Unterraum von  $\mathcal{L}^2(\mathbb{R}^n)$ .

**Definition 26.12.**

$$L^2(\mathbb{R}^n) := \mathcal{L}^2(\mathbb{R}^n) / \text{Nullfunktionen}$$

im Sinne der Quotientenbildung (19.15). 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 26.13.**  $(L^2(\mathbb{R}^n), \langle \cdot, \cdot \rangle)$  ist ein Hilbertraum

*Beweis.* Hier folgen die algebraischen Eigenschaften und die Positivität direkt aus der Wohldefiniertheit der von (26.38) 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

$$\|f_k - f_{k-1}\| \leq 2^{-k} \quad k \geq 2 \quad (26.41)$$

Wir beweisen dann wie bei  $\ell^2()$  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

$$g_1 := |f_1|, \quad g_2 := |f_1| + |f_2 - f_1|, \dots, \quad g_k = g_{k-1} + |f_k - f_{k-1}|, \dots \quad (26.42)$$

Dann ist  $(g_k) \subset \mathcal{L}^2(\mathbb{R}^n)$  punktweise monoton wachsend mit

$$\|g_k\| \leq \|f_1\| + \sum_{i=2}^k \|f_i - f_{i-1}\| \quad (26.43)$$

wegen der Dreiecksungleichung für die Norm und (26.41). Daher ist  $\left(\int_{\mathbb{R}^n} g_k^2\right)$  beschränkt und mit B. Levi 25.3 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 26.8 und

$$|f_k| \leq |f_1| + \sum_{i=2}^k |f_i - f_{i-1}| = g_k \leq g \quad (\text{fast überall}) \quad (26.44)$$

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 26.9, dass  $f - f_k \in \mathcal{L}^2(\mathbb{R}^n)$  und mit 25.4, 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. ???. 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 24.6 unter dem Integral durch Funktionen in  $\mathcal{C}_c$  approximieren lassen.)

(2) Treppenfunktionen mit rationalen Werten auf rationalen Quadern sind abzählbar

§ 26. HILBERT-RÄUME

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

$$\int_{\mathbb{R}^n} |f - \varphi|^2 < \epsilon \tag{26.45}$$

(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 Vektorraum<sup>45</sup>

$$\mathcal{S}(\mathbb{R}^n, \mathbb{C}) = \{f \in \mathcal{C}^\infty(\mathbb{R}^n, \mathbb{C}) \mid \|x^\alpha D^\beta f\|_\infty < \infty \forall \alpha, \beta \in (\mathbb{Z}_{\geq 0})^n\} \tag{26.47}$$

heisst “Schwartz-Raum des  $\mathbb{R}^n$ ” und ist Ausgangspunkt vieler beliebter Untersuchungen.

· Ein vermutlich aus den Physik-Vorlesungen oder -Übungen schon vertrautes Beispiel eines vollständigen Orthonormalsystems sind die Funktionen

$$\chi_k(t) := e^{2\pi i k t} \quad \text{für } k \in \mathbb{Z} \tag{26.48}$$

im Hilbertraum der periodischen Funktionen

$$\begin{aligned} L^2_{\text{per.}}(\mathbb{R}) &:= \{f \in \mathcal{M}(\mathbb{R}, \mathbb{C}) \mid f(t+1) = f(t) \text{ f.ü.}, \int_0^1 |f(t)|^2 dt < \infty\} / \text{Nullfunktionen} \\ &\cong \{f : [0, 1) \rightarrow \mathbb{C} \mid f \text{ messbar und quadratintegrierbar}\} / \text{Nullfunktionen} \end{aligned} \tag{26.49}$$

mit dem inneren Produkt

$$\langle f, g \rangle = \int_0^1 \bar{f}(t)g(t)dt \tag{26.50}$$

Es gilt

$$\langle \chi_k, \chi_l \rangle = \int_0^1 e^{2\pi i(l-k)t} dt = \begin{cases} \frac{1}{2\pi i(l-k)} e^{2\pi i(l-k)t} \Big|_0^1 = 0 & k \neq l \\ \int_0^1 dt = 1 & k = l \end{cases} \tag{26.51}$$

und man zeigt, dass für alle  $f \in L^2_{\text{per.}}(\mathbb{R})$  die Fourier-Reihe

$$t \mapsto \sum_{k=-\infty}^{\infty} a_k e^{2\pi i k t} \tag{26.52}$$

<sup>45</sup>Hier sind  $\alpha, \beta$  sogenannte Multi-Indizes, und

$$x^\alpha D^\beta f(x) := (x^1)^{\alpha_1} \dots (x^n)^{\alpha_n} \cdot (\partial_{x^1})^{\beta_1} \dots (\partial_{x^n})^{\beta_n} f(x) \tag{26.46}$$

in der  $L^2_{\text{per.}}$ -Norm gegen  $f$  konvergiert, mit den Fourier-Koeffizienten

$$a_k = \langle \chi_k, f \rangle = \int_0^1 f(t) e^{-2\pi i k t} dt \in \mathbb{C} \quad (26.53)$$

Der durch (26.52) gestiftete Isomorphismus

$$L^2_{\text{per.}}(\mathbb{R}) \cong \ell^2(\mathbb{Z}) \quad (26.54)$$

impliziert insbesondere die *Parsevalsche Gleichung*

$$\sum_{k=-\infty}^{\infty} |a_k|^2 = \int_0^1 |f(t)|^2 dt \quad (26.55)$$

Für die historisch und praktisch interessante Frage der punktweisen Konvergenz von (26.52), 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.

## § 27 Lineare Operatoren

Hilberträume<sup>46</sup> bilden also das mathematische Substrat der Quantenmechanik. Da aber je zwei unendlich-dimensionale Hilberträume isomorph zueinander sind (vgl. 26.5), 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 27.1.** 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. 8.5.
- Gemäss 26.1 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

---

<sup>46</sup>Es interessieren wie gesagt nur separable Hilberträume und wir lassen dieses Adjektiv jetzt weg.

## § 27. 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

$$U_0^{\mathbb{N}} := \{(u_k)_{k \in \mathbb{N}} \subset U \mid (u_k)_{k \in \mathbb{N}} \text{ ist Cauchy-Folge}\} \quad (27.1)$$

aller Cauchy-Folgen in  $U$  die Äquivalenzrelation

$$(u_k) \sim (u'_k) \iff \text{Die Reissverschlussfolge } (u_1, u'_1, u_2, u'_2, \dots) \text{ ist eine Cauchy-Folge.} \quad (27.2)$$

und definiert

$$\overline{U} = U_0^{\mathbb{N}} / \sim \quad (27.3)$$

mit der (wohl-definierten) Norm

$$\|[(u_k)]\| := \lim_{k \rightarrow \infty} \|u_k\| \quad (27.4)$$

(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 27.2.** *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,*

$$\|A\| := \sup\{\|A(u)\| \mid \|u\| = 1\} < \infty \quad (27.5)$$

*Beweis. " $\Rightarrow$ ": Ist  $A$  beschränkt, so gilt (leichte Übung...)*

$$\|A(u_1) - A(u_2)\| = \|A(u_1 - u_2)\| \leq \|A\| \cdot \|u_1 - u_2\| \quad (27.6)$$

$A$  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 27.3.** *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

$$\overline{A}(v) := \lim_{k \rightarrow \infty} A(u_k) \tag{27.7}$$

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:<sup>47</sup>

**Proposition/Definition 27.4.** *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 27.3 oBdA mit  $\mathcal{D}(A) = V \ \forall A \in \mathcal{B}(V, W)$ .
- Mit der Operatornorm (27.5) 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 ?? 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:

$$m_{f_0} : f \mapsto f_0 \cdot f \tag{27.8}$$

mit Norm

$$\|m_{f_0}\| = \|f_0\|_\infty \tag{27.9}$$

- 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

$$\delta((a_k)_{k \in \mathbb{Z}}) := (ka_k)_{k \in \mathbb{Z}} \tag{27.10}$$

---

<sup>47</sup>Der Sprachgebrauch ist ähnlich zu den meromorphen Funktionen in 18.3, die ebenfalls nicht auf dem gesamten Definitionsbereich definiert sind.

## § 27. LINEARE OPERATOREN

genau dann Sinn, wenn  $\delta((a_k)_{k \in \mathbb{Z}}) \in \ell^2(\mathbb{Z})$ , d.h. wenn

$$(a_k)_{k \in \mathbb{Z}} \in \mathcal{D}(\delta) := \left\{ (a_k)_{k \in \mathbb{Z}} \mid \sum_{k=-\infty}^{\infty} |ka_k|^2 < \infty \right\} \quad (27.11)$$

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  $i : L^2_{\text{per.}}(\mathbb{R}) \rightarrow \ell^2(\mathbb{Z})$ ,  $i(e^{2\pi ikt}) = e_k$  (s. (26.54)),

$$(i^{-1} \circ \delta \circ i)(e^{2\pi ikt}) = ke^{2\pi ikt} = \frac{1}{2\pi i} \frac{d}{dt} e^{2\pi ikt} \quad (27.12)$$

wird  $\delta$  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  $\psi(x)$ ,  $x \in \mathbb{R}$

$$(q\psi)(x) := x\psi(x), \quad (p\psi)(x) := \frac{\hbar}{i} \cdot \frac{d\psi}{dx}(x) \quad (27.13)$$

Diese Operatoren erfüllen (zumindest formal) die *kanonischen Vertauschungsrelationen*

$$[q, p] = q \circ p - p \circ q = i\hbar \cdot \text{id}, \quad (27.14)$$

(Stichwort: Heisenbergsche Unschärferelation), sind aber offenbar nicht beschränkt: Die rechten Seiten in (27.13) sind i.A. nicht quadratintegrierbar, selbst wenn  $\psi$  es ist. Der Schwartzraum (26.47) 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 27.5.** *Es sei  $\mathcal{A}$  eine normierte  $\mathbb{C}$ -Algebra mit Einselement und submultiplikativer Norm 4.53. Dann gilt (27.14) für keine  $p, q \in \mathcal{A}$ . Insbesondere gibt es keine endlich-dimensionalen Darstellungen der KVRn.*

(Es wird hier *nicht* gesagt, dass die Gleichung (27.14) keine Lösung hat, die Operatoren in (27.13) geben ja ein explizites Beispiel. Der Punkt ist nur, dass Operatoren, welche (27.14) erfüllen, nicht beschränkt sein können.)

*Beweis.* Angenommen, (27.14) gälte doch, und zwar oBdA mit  $\hbar = -i$ . Aus der Derivationseigenschaft des Kommutators folgt für  $n = 1, 2, \dots$

$$\begin{aligned} [q, p^n] &= (qp)p^{n-1} - pqp^{n-1} + pqp^{n-1} - pp^{n-1}q \\ &= [q, p]p^{n-1} + p[q, p^{n-1}] = \dots \\ &= \sum_{i=0}^{n-1} p^i [q, p] p^{n-i} \\ &= np^{n-1} \end{aligned} \quad (27.15)$$

Daraus folgt einerseits induktiv, dass  $p^n \neq 0 \forall n$ , andererseits aus den Eigenschaften der Norm

$$\begin{aligned} n \cdot \|p^{n-1}\| &\leq 2\|q\| \cdot \|p^n\| \leq 2\|q\| \cdot \|p\| \cdot \underbrace{\|p^{n-1}\|}_{\neq 0} \\ \Rightarrow n &\leq 2\|q\| \cdot \|p\| \end{aligned} \tag{27.16}$$

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 (21.5).

**Theorem 27.6.** *Es sei  $(\mathcal{H}, \langle \cdot, \cdot \rangle)$  ein Hilbertraum. Dann ist der Raum der stetigen linearen Abbildungen (a.k.a. Funktionale)*

$$\mathcal{B}(\mathcal{H}, \mathbb{C}) = \{\lambda : \mathcal{H} \rightarrow \mathbb{C} \mid \|\lambda\| < \infty\} \tag{27.17}$$

mit der Operatornorm ein (separabler!) Hilbertraum. Die Abbildung

$$\mathcal{H} \ni v \mapsto (v^\flat : w \mapsto \langle v, w \rangle) \in \mathcal{B}(\mathcal{H}, \mathbb{C}) \tag{27.18}$$

ist ein Anti-Isomorphismus von Hilberträumen. Man schreibt (nicht konform mit 19.8)  $\mathcal{H}^* \cong \mathcal{H}$  und nach Dirac (vgl. (26.33))

$$(|\text{blah}\rangle)^\flat = \langle \text{blah} | \tag{27.19}$$

Vor dem Beweis halten wir zunächst die Eigenschaften von orthogonalen Zerlegungen von Hilberträumen fest.

**Proposition/Definition 27.7.** *Für eine Teilmenge  $M \subset \mathcal{H}$  ist*

$$M^\perp := \{v \in \mathcal{H} \mid \langle m, v \rangle = 0 \forall m \in M\} \tag{27.20}$$

ein abgeschlossener Unterraum von  $\mathcal{H}$ . Ist  $M = U$  selbst ein abgeschlossener Unterraum, so ist

$$\mathcal{H} = U \oplus U^\perp \tag{27.21}$$

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),  $\forall m \in M$ :

$$\langle v, m \rangle = \langle \lim_{k \rightarrow \infty} v_k, m \rangle = \lim_{k \rightarrow \infty} \langle v_k, m \rangle = 0 \tag{27.22}$$

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

$$d(v, U) = \inf\{\|v - u\| \mid u \in U\} \tag{27.23}$$

§ 27. LINEARE OPERATOREN

(der “Abstand” von  $v$  zu  $U$ ) und  $(u_k) \subset U$  eine Folge mit

$$d(v, U) = \lim_{k \rightarrow \infty} \|v - u_k\| \quad (27.24)$$

Dann gilt mit der Parallelogrammgleichung (26.3)

$$\begin{aligned} \|u_k - u_l\|^2 &= \|(u_k - v) - (u_l - v)\|^2 \\ &= 2(\|u_k - v\|^2 + \|u_l - v\|^2) - \underbrace{\|u_k + u_l - 2v\|^2}_{=4\|\frac{u_k+u_l}{2}-v\|^2 \geq 4d^2} \\ &\leq 2(\|u_k - v\|^2 + \|u_l - v\|^2) \end{aligned} \quad (27.25)$$

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

$$\|v - u_{\parallel}\| = d(v, U) \quad (27.26)$$

und hängt mit dieser Eigenschaft nicht von der Folge  $(u_k)$  ab (vorausgesetzt, sie erfüllt (27.24), Übung). Wir setzen  $u_{\perp} = v - u_{\parallel}$  und behaupten  $u_{\perp} \in U^{\perp}$ . Dazu betrachten wir für  $u \in U$  und  $t \in \mathbb{R}$

$$\begin{aligned} \|u_{\perp}\|^2 &= d(v, U)^2 \leq \|v - u_{\parallel} + tu\|^2 \quad (\text{Def. von } d(v, U), -u_{\parallel} + tu \in U) \\ &= \|u_{\perp}\|^2 + 2t \operatorname{Re}\langle v - u_{\parallel}, u \rangle + t^2 \|u\|^2 \end{aligned} \quad (27.27)$$

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 26.1)  $\operatorname{Im}\langle u_{\perp}, u \rangle = 0 \forall u \in U$ , 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 27.6.* Es genügt offenbar zu zeigen, dass (27.18) 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

$$|v^{\flat}(w)| = |\langle v, w \rangle| \leq \|v\| \cdot \|w\| \quad (27.28)$$

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  $\lambda$  ist

$$U := \operatorname{Ker}(\lambda) \quad (27.29)$$

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

$$\tilde{w} := w - \frac{\lambda(w)}{\lambda(v_0)} v_0 \in \operatorname{Ker}(\lambda) = U \quad (27.30)$$

(Denn  $\lambda(\tilde{w}) = \lambda(w) - \frac{\lambda(w)}{\lambda(v_0)}\lambda(v_0) = 0$ .) Da  $v_0 \perp U$  gilt also

$$0 = \langle v_0, \tilde{w} \rangle = \langle v_0, w \rangle - \lambda(w) \frac{\langle v_0, v_0 \rangle}{\lambda(v_0)} \quad (27.31)$$

d.h.  $\forall w \in \mathcal{H}$

$$\lambda(w) = \left\langle \frac{\overline{\lambda(v_0)}}{\|v_0\|^2} v_0, w \right\rangle \quad (27.32)$$

Es ist also  $\lambda^\sharp := \frac{\overline{\lambda(v_0)}}{\|v_0\|^2} v_0 \in \mathcal{H}$  ein Urbild von  $\lambda$  unter (27.18),  $\lambda = (\lambda^\sharp)^\flat$ .  $\square$

*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)

$$T(e_k) := e_{k+1} \quad (27.33)$$

auf  $\ell^2(\mathbb{N})$ .  $T$  ist dicht definiert und stetig, die Fortsetzung auf ganz  $\ell^2(\mathbb{N})$  ist

$$T((a_1, a_2, a_3, \dots)) = (0, a_1, a_2, \dots) \quad (27.34)$$

Man sieht leicht, dass  $T$  injektiv, aber nicht surjektiv ist.

· In Fortsetzung der Diskussion um 19.9 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

$$\langle v_1 \oplus v_2, w_1 \oplus w_2 \rangle := \langle v_1, w_1 \rangle + \langle v_2, w_2 \rangle \quad (27.35)$$

(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

$$\langle v_1 \otimes v_2, w_1 \otimes w_2 \rangle := \langle v_1, w_1 \rangle \cdot \langle v_2, w_2 \rangle \quad (27.36)$$

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 27.3. Es gilt dann beispielsweise

$$L^2(\mathbb{R}^n) \overline{\otimes} L^2(\mathbb{R}^m) \cong L^2(\mathbb{R}^{n+m}) \quad (27.37)$$

(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  $\mathcal{H}_1 \overline{\otimes} \mathcal{H}_2$ .)

## § 27. 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 grösseren<sup>48</sup> zuständig ist.

· Natürlich kann man hier auch wieder symmetrische und anti-symmetrische Tensorprodukte und Multilinearformen in Verbindung setzen.

Physikalisch stellt 27.6 eine Korrespondenz zwischen Präparation (den Kets  $|\psi\rangle \in \mathcal{H}$ ) und Registrierung (den Bras  $\lambda = \langle\varphi| \in \mathcal{H}^* \cong \mathcal{H}$ ) 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

$$\mathbb{E}_\psi(A) = \langle\psi, A\psi\rangle \quad \text{mit } \langle\psi, \psi\rangle = 1 \quad (27.38)$$

(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 in 21 untersucht hatten, siehe 21.4 und 21.12 und vgl. auch das Auftauchen der Form (27.38) in 21.13. In den unendlichen Dimensionen eines allgemeinen separablen Hilbertraums ist Platz für einige weitere Neuheiten. Als Erstes müssen wir die in der Fussnote 21 gegebene Definition an die Beobachtung 27.5 anpassen, dass physikalisch relevante Operatoren nicht notwendig beschränkt sind.

**Lemma/Definition 27.8.** *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*

$$\langle v^A, w \rangle = \langle v, Aw \rangle \quad \forall w \in \mathcal{D}(A) \quad (27.39)$$

(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*

$$\begin{aligned} \mathcal{D}(A^\dagger) &:= \{v \in \mathcal{H} \mid \exists v^A \in \mathcal{H} : \langle v^A, w \rangle = \langle v, Aw \rangle \forall w \in \mathcal{H}\} \\ A^\dagger(v) &:= \text{eben jenes } v^A \end{aligned} \quad (27.40)$$

*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  $\|A^\dagger\| = \|A\|$ .*

<sup>48</sup>Wie die Formeln zeigen, ist das Tensorprodukt von separablen Hilberträumen wieder separabel.

*Beweis.* (i) Erfüllen  $v^A$  und  $\tilde{v}^A$  beide (27.39), 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 27.7, dass  $v^A - \tilde{v}^A = 0$ .

(ii) Übungsaufgabe

(iii) Für jedes feste  $v \in \mathcal{H}$  ist die Abbildung

$$\mathcal{H} \ni w \mapsto \langle v, Aw \rangle \in \mathbb{C} \tag{27.41}$$

eine beschränkte Linearform, denn wegen der C.S.U. und  $A \in \mathcal{B}(\mathcal{H})$  ist

$$|\langle v, Aw \rangle| \leq \|v\| \cdot \|A\| \cdot \|w\| \tag{27.42}$$

Wegen 27.6 existiert daher ein  $v^A = A^\dagger(v)$  wie gefordert.

• Aus der für jeden stetigen linearen Operator gültigen Aussage, dass<sup>49</sup>

$$\|A\| = \sup\{|\langle v, Aw \rangle| \mid \|v\| = \|w\| = 1\} \tag{27.45}$$

folgt zunächst

$$\|A^\dagger\| \leq \|A\| \tag{27.46}$$

$A^\dagger$  ist also insbesondere beschränkt. Aus

$$\langle (A^\dagger)^\dagger v, w \rangle = \langle v, A^\dagger w \rangle = \overline{\langle A^\dagger w, v \rangle} = \overline{\langle w, Av \rangle} = \langle Av, w \rangle \tag{27.47}$$

für alle  $v, w \in \mathcal{H}$  folgt dann  $(A^\dagger)^\dagger = A$  und zuletzt durch Wiederholung von (27.46)  $\|A\| = \|(A^\dagger)^\dagger\| \leq \|A^\dagger\|$ .  $\square$

Die Definition in Fussnote ?? bleibt im Wesen unverändert,

**Definition 27.9.** Ein linearer Operator  $U$  heisst unitär, falls  $U^\dagger U = U U^\dagger = \text{id}_{\mathcal{H}}$ .

man muss nur wieder darauf aufpassen, dass

$$\text{injektiv} \Leftrightarrow \text{surjektiv} \tag{27.48}$$

*Beispiel.* Für den in (27.33) definierten Operator  $T$  gilt

$$\langle (b_k), T(a_k) \rangle = 0 + \sum_{k=2}^{\infty} \bar{b}_k a_{k-1} = \sum_{k=1}^{\infty} \bar{b}_{k+1} a_k \tag{27.49}$$

d.h.

$$T^\dagger(b_1, b_2, b_3, \dots) = (b_2, b_3, b_4, \dots) \tag{27.50}$$

---

<sup>49</sup>Einerseits ist

$$\sup\{|\langle v, Aw \rangle| \mid \|v\| = \|w\| = 1\} \leq \sup\{\|v\| \cdot \|Aw\| \mid \|v\| = \|w\| = 1\} = \|A\|, \tag{27.43}$$

andererseits (jedenfalls für  $A \neq 0$ )

$$\begin{aligned} \|A\| &= \sup\{\|Aw\| \mid \|w\| = 1, Aw \neq 0\} \\ &= \sup\left\{\left\langle \frac{Aw}{\|Aw\|}, Aw \right\rangle \mid \|w\| = 1, Aw \neq 0\right\} \\ &\leq \sup\{|\langle v, Aw \rangle| \mid \|v\| = \|w\| = 1\} \end{aligned} \tag{27.44}$$

## § 27. 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. *Projektor* auf die erste Komponente ist

$$P_1((a_k)) = (a_1, 0, \dots) \quad (27.51)$$

$TT^\dagger$  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 (27.48) 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 27.10.** 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 27.11.** Ein linearer Operator  $A$  ist genau dann symmetrisch, wenn die assoziierte quadratische<sup>50</sup> Form, nämlich

$$\mathcal{D}(A) \ni v \mapsto q_A(v) := \langle v, Av \rangle \quad (27.52)$$

reell ist.

*Beweis.* · Ist  $A$  symmetrisch, so ist  $v \in \mathcal{D}(A)$  insbesondere auch in  $\mathcal{D}(A^\dagger)$  und es gilt

$$\overline{\langle v, Av \rangle} = \overline{\langle Av, v \rangle} = \langle v, Av \rangle \quad (27.53)$$

d.h.  $q_A(v) \in \mathbb{R}$ .

· Es sei umgekehrt  $q_A$  reell. Aus der Identität (vgl. (26.8))

$$\langle v, Aw \rangle = \frac{1}{4} (q_A(v+w) - q_A(v-w) + iq_A(v-iw) - iq_A(v+iw)) \quad (27.54)$$

$\forall v, w \in \mathcal{D}(A)$  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 21.12 über die Exponentialfunktion eine Einparametergruppe unitärer Symmetrietransformationen von  $\mathcal{H}$  erzeugt.

**Definition 27.12.** Ein symmetrischer linearer Operator  $A$  heisst selbst-adjungiert, wenn  $\mathcal{D}(A^\dagger) = \mathcal{D}(A)$ .

**Beispiel 27.13.** · Multiplikation mit einer (vernünftigen, beispielsweise stetigen) reellen Funktion  $f$  definiert einen selbst-adjungierten Operator auf  $L^2(\mathbb{R}^n)$ . Definitionsbereich ist

$$\mathcal{D}(m_f) = \{ \psi \in L^2(\mathbb{R}^n) \mid f \cdot \psi \in L^2(\mathbb{R}^n) \} = \mathcal{D}(m_f^\dagger) \quad (27.55)$$

<sup>50</sup>Man 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

$$p\psi(x) := \frac{\hbar}{i} \frac{d\psi}{dx}(x) \tag{27.56}$$

definierte Operator  $p$  ist symmetrisch: Für  $\psi, \varphi \in \mathcal{C}_c^1(\mathbb{R})$  gilt

$$\begin{aligned} \int_{\mathbb{R}} \bar{\varphi}(x) \frac{\hbar}{i} \frac{d\psi}{dx}(x) dx &= - \int_{\mathbb{R}} \frac{\hbar}{i} \frac{d\bar{\varphi}}{dx}(x) \psi(x) dx \quad (\text{partielle Integration}) \\ &= \int_{\mathbb{R}} \frac{\hbar}{i} \frac{d\bar{\varphi}}{dx}(x) \psi(x) dx \end{aligned} \tag{27.57}$$

Zunächst ist zwar  $\mathcal{D}(p^\dagger) \not\supseteq \mathcal{D}(p)$  (beispielsweise geht die Rechnung (27.57) sicher auch durch, wenn  $\varphi$  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 (27.39) zu erfüllen.) Es ist aber nicht-trivial, dass man sich in der Mitte treffen kann.

· Die **formale Rechnung** für  $t \in \mathbb{R}$ ,

$$\left(\exp\left(\frac{i}{\hbar}tp\right)\psi\right)(x) = \sum_{k=0}^{\infty} \frac{1}{k!} \left(t \frac{d}{dx}\right)^k \psi(x) = \psi(x) + t\psi'(x) + \frac{t^2}{2}\psi''(x) + \dots = \psi(x+t) \tag{27.58}$$

(Taylorreihe) weist  $\frac{i}{\hbar}p$  als infinitesimalen Erzeuger der Translationen

$$(T_t\psi)(x) = \psi(x+t) \tag{27.59}$$

aus. (Etwas formaler argumentiert man mit der Differentialgleichung (s. (21.30))

$$\frac{d}{dt}(T_t(\psi)(x)) = \frac{d}{dx}(T_t(\psi))(x) = \frac{i}{\hbar}p(T_t(\psi))(x) \tag{27.60}$$

$T_t$  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:

$$T_2(\chi_{[0,1]}) = 0 \tag{27.61}$$

Die diskrete Version war gerade (27.33).

## § 28. STURM-LIOUVILLE-THEORIE

· Die Untersuchung der Symmetrie von  $\frac{\hbar}{i} \frac{d}{dx}$  auf Funktionen auf  $[0, \infty)$  mittels partieller Integration liefert den Randterm

$$\frac{\hbar}{i} \bar{\varphi}(0)\psi(0) \quad (27.62)$$

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  $\mathcal{D}(p) \neq \mathcal{D}(p^\dagger)$ .<sup>51</sup>

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.

## § 28 Sturm-Liouville-Theorie

[sturmliouville]

Sturm-Liouville-Randwertprobleme sind in der mathematischen Physik allgegenwärtig. Man sollte sie hier als Beispiel der Spektraltheorie ausführlich behandeln. Problem: Vollständigkeit der Eigenfunktionen.

## § 29 Fourier-Transformation

[fouriertrafo]

Anknüpfung an vorherigen § über harmonischen Oszillator (Hermite/Laguerre-Polynome), dann Übergang zu partiellen Differentialgleichungen und höher-dimensionalen Randwertproblemen, Greensche Funktionen. Wohin mit Distributionen?

---

<sup>51</sup>Dies ist natürlich formal nicht ganz korrekt, denn unsere "Funktionen" nehmen ja an Punkten keine echten Werte an...