



Prof. Dr. Michael Herrmann
Technische Universität Braunschweig
Mathematik – Institut iPDE
michael.herrmann@tu-braunschweig.de

Skript der Vorlesung
Analysis 3
im Wintersemester 2022/23

Version vom 8. Februar 2023

Der Autor ist für Hinweise und Kommentare jederzeit dankbar.

© Michael Herrmann

Dieses Skript ist lizenziert unter **CC BY-SA 4.0**.
<http://creativecommons.org/licenses/by-sa/4.0/deed.de>



Griechisches Alphabet

<i>klein</i>	<i>groß</i>	<i>Name</i>	<i>Laut</i>	<i>klein</i>	<i>groß</i>	<i>Name</i>	<i>Laut</i>
α	A	alpha	a	ν	N	ny	n
β	B	beta	b	ξ	Ξ	xi	x
γ	Γ	gamma	g	o	O	omikron	ö
δ	Δ	delta	d	π	Π	pi	p
ε, ϵ	E	epsilon	ě	ϱ, ρ	P	rho	r
ζ	Z	zeta	z	σ	Σ	sigma	s
η	H	eta	ē	τ	T	tau	t
θ, ϑ	Θ	theta	th	υ	Υ	upsilon	y
ι	I	iota	i	φ, ϕ	Φ	phi	ph, f
κ	K	kappa	k	χ	X	chi	ch
λ	Λ	lambda	l	ψ	Ψ	psi	ps
μ	M	my	m	ω	Ω	omega	ō

Literatur

Es gibt viele sehr gute Lehrbücher zur Maß- und Integrationstheorie, zum Beispiel:

[Bau] HEINZ BAUER: *Maß- und Integrationstheorie*
2. Auflage, de Gruyter 2010

[Els] JÜRGEN ELSTRODT: *Maß- und Integrationstheorie*
8. Auflage, Springer Spektrum 2018

[For] OTTO FORSTER: *Analysis 3*
8. Auflage, Springer Spektrum 2017

Anwendungsorientierte Darstellungen der höherdimensionalen Integration finden sich in der ingenieurmathematischen Literatur, zum Beispiel in:

[AORS] R. ANSORGE, H.J. OBERLE, K. ROTHE, TH. SONAR:
Mathematik für Ingenieure 2. Wiley-VCH 2011, 4. Auflage.

[MV1] K. MEYBERG, P. VACHENAUER:
Höhere Mathematik 1. SpringerVieweg 2001, 6. Auflage.

[MV2] K. MEYBERG, P. VACHENAUER:
Höhere Mathematik 2. SpringerVieweg 2006, 4. Auflage.

Die wesentlichen Definitionen, Theoreme und Beispiele können auch in vielen anderen Büchern gefunden werden:

[ABHKLS] T. ARENS, R. BUSAM, FR. HETTLICH, CH. KARPFINGER, KL.
LICHTENEGGER, H. STACHEL:
Mathematik. Springer Spektrum 2015, 3. Auflage.

Anmerkungen

Bei der Ausarbeitung dieses Skriptes hat der Autor die gelisteten Werke regelmäßig konsultiert und dabei viele Beweisstrategien, Präsentationsideen und Beispiele übernommen. Er hat sich außerdem von anderen Quellen inspirieren lassen und eigene Ausarbeitungen zu früheren Vorlesungen verwendet. Ein besonderer Dank geht an Katia Kleine und Dirk Janßen für das sorgfältige Korrekturlesen.

Inhaltsverzeichnis

1	Maß- und Integrationstheorie	5
1.1	Mengensysteme	5
1.2	Maße auf Ringen und Algebren	12
1.3	Vervollständigung eines Maßes	18
1.4	Das Lebesgue-Maß	29
1.5	Messbare Funktionen	47
1.6	Integrierbare Funktionen	56
1.7	Spezialfälle und Beispiele	67
1.8	Weitere Integralsätze	74
2	Integration in den Anwendungen	83
2.1	Exkurs*: das Riemann-Integral auf Quadern	84
2.2	Satz von Fubini auf Quadern	90
2.3	Exkurs*: Kurven und Kurvenintegrale	93
2.4	Integralsätze auf zweidimensionalen Quadern	97
2.5	Gebietsintegrale auf kompakten Mengen	104
2.6	Integration über spezielle Gebietsklassen	113
2.7	Die Transformationsformel für Integrale	121
2.8	Integralsätze von Gauß und Stokes in 2D	131
3	Vektoranalysis	145
3.1	mehr über Kurvenintegrale	145
3.2	Grundlagen der Potentialtheorie	155
3.3	Frenet-Theorie und Geometrie von Kurven	165
3.4	Grundlagen der Flächentheorie	177
3.5	Flächenintegrale	188
3.6	Integralsätze von Gauß und Stokes in 3D	198

Kapitel 1

Maß- und Integrationstheorie

Vorlesung 01-1 : 26. Oktober

1.1 Mengensysteme

Vorbemerkung Im Folgenden ist X eine beliebige, aber feste Menge¹ und M bezeichnet eine Teilmenge von X , also ein Element der Potenzmenge²

$$\mathcal{P}(X) = \text{Menge aller Teilmengen von } X = \{M : M \subseteq X\}.$$

Eine Teilmenge von $\mathcal{P}(X)$ besteht aus Teilmengen von X und wird Mengensystem in X genannt. Wir bezeichnen in dieser Vorlesung Mengensysteme mit kalligraphischen Großbuchstaben, etwa mit \mathcal{R} .³

Achtung Ein Mengensystem ist eine *Menge von Teilmengen* von X und jedes ihrer Elemente ist eine Teilmenge von X .

Beispiel Für $X = \{1, 2, 3\}$ gilt

$$\mathcal{P}(X) = \{\{\}, \{1\}, \{2\}, \{3\}, \{2, 3\}, \{1, 3\}, \{1, 2\}, \{1, 2, 3\}\},$$

wobei $\{\}$ eine andere Notation für die leere Menge \emptyset ist. Das Mengensystem

$$\mathcal{R} = \{\{\}, \{2, 3\}, \{1, 3\}, \{1, 2\}\}$$

besteht aus allen Teilmengen von X , die eine gerade Anzahl von Elementen besitzen.

¹Später werden wir oft den \mathbb{R}^n oder geometrische Teilmengen des \mathbb{R}^n (Kugeln, Kurven, Flächen, usw.) als Menge X verwenden.

²In der Mathematik wird auch häufig anstelle von $\mathcal{P}(X)$ die alternative Notation 2^X benutzt, da jede Teilmenge $M \subseteq X$ in natürlicher und eindeutiger Weise mit einer Funktion $f_M : X \rightarrow \{0, 1\}$ identifiziert werden kann, wobei $f_M(x) = 1$ dann und nur dann gilt, wenn x ein Element von M ist. Oder anders gesagt: Wir können eine Teilmenge $M \subseteq X$ dadurch charakterisieren, dass wir für jedes x in X eine binäre Entscheidung treffen, nämlich ob x zu M gehört oder nicht.

³Die Formeln $\mathcal{R} = \emptyset$ und $\mathcal{R} = \{\emptyset\}$ beschreiben zwei verschiedene Mengensysteme in X : Das erste enthält überhaupt keine Teilmenge von X , wohingegen das zweite genau ein Element besitzt, nämlich die triviale Teilmenge von X .

über Mengenoperationen Neben den bekannten Standardoperationen der naiven Mengenlehre (Durchschnitt, Vereinigung, Differenz) verwenden wir auch

$$M' := X \setminus M, \quad M_1 \triangle M_2 := (M_1 \setminus M_2) \cup (M_2 \setminus M_1),$$

die den Übergang zur Komplementmenge bzw. die Bildung der symmetrischen Differenz beschreiben.⁴ Siehe dazu auch das nachfolgende Bild. Insbesondere können wir diese Mengenoperationen auf die Elemente eines Mengensystems anwenden und erhalten wieder eine Teilmenge von X und damit ein Element von $\mathcal{P}(X)$.

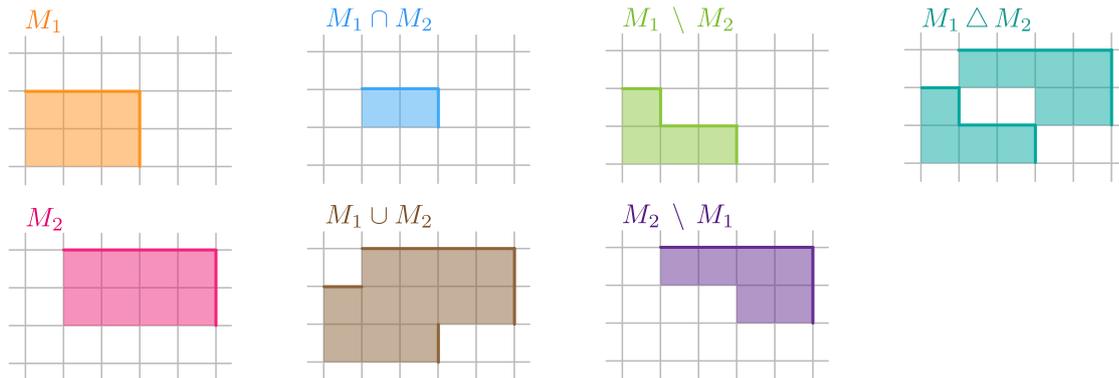


Abbildung Schematische Darstellung einiger Rechenoperationen für Mengen, wobei M_1 und M_2 hier als halboffene Quader gewählt wurden.

Definition Ein nichtleeres Mengensystem $\mathcal{R} \subseteq \mathcal{P}(X)$ wird Ring (in X) genannt, sofern

$$M_1 \cap M_2 \in \mathcal{R}, \quad M_1 \triangle M_2 \in \mathcal{R}$$

für alle $M_1, M_2 \in \mathcal{R}$ erfüllt ist.

Bemerkungen

1. Wegen $M \triangle M = \emptyset$ gilt $\emptyset \in \mathcal{R}$, d.h. die leere Menge ist (als triviale Teilmenge von X) Element jedes Ringes \mathcal{R} .
2. Jeder Ring \mathcal{R} ist per Definition abgeschlossen bzgl. der Bildung des Durchschnitts und der symmetrischen Differenz seiner Elemente.⁵ Für je zwei $M_1, M_2 \in \mathcal{R}$ gilt aber auch

$$M_1 \cup M_2 = (M_1 \triangle M_2) \triangle (M_1 \cap M_2) \in \mathcal{R}$$

sowie

$$M_1 \setminus M_2 = (M_1 \triangle M_2) \cap M_1 \in \mathcal{R},$$

d.h. wir können innerhalb des Ringes auch immer die Vereinigung sowie die (nicht-symmetrische) Differenz zweier Elemente bilden.

⁴In der Literatur schreibt man oftmals auch M^C anstelle von M' .

⁵Wir erinnern noch einmal: Die Elemente von \mathcal{R} sind Teilmengen von X .

3. Jeder Ring \mathcal{R} ist außerdem abgeschlossen bzgl. *endlicher* Vereinigungen und Durchschnitte, denn induktiv können wir

$$\bigcup_{j=1}^n M_j = M_1 \cup \dots \cup M_n \in \mathcal{R}, \quad \bigcap_{j=1}^n M_j = M_1 \cap \dots \cap M_n \in \mathcal{R}$$

für alle $M_1, \dots, M_n \in \mathcal{R}$ zeigen.

Achtung: Ein Ring ist im Allgemeinen aber *nicht* abgeschlossen bzgl. der Bildung *unendlicher* Vereinigungen bzw. Durchschnitte.

4. Wenn wir die Operationen Δ bzw. \cap als abstrakte Addition bzw. Multiplikation ansehen, können wir in jedem Ring \mathcal{R} die Ringaxiome aus der Mathematischen Algebra verifizieren. Siehe dazu die Übungen.

Definition Jeder Ring \mathcal{R} mit $X \in \mathcal{R}$ wird Algebra genannt. Wir sprechen außerdem von einer σ -Algebra, wenn zusätzlich

$$\bigcup_{j=1}^{\infty} M_j = M_1 \cup M_2 \cup \dots \in \mathcal{R}$$

für jede Folge $(M_j)_{j=1}^{\infty}$ aus \mathcal{R} gilt, d.h. wenn die unendliche Vereinigung ihrer Glieder auch zu \mathcal{R} gehört.

Bemerkungen

1. Für Algebren benutzen wir oftmals (aber nicht immer) den Buchstaben \mathcal{A} .
2. Die Formel $M' = M \Delta X$ impliziert, dass eine Teilmenge $M \subseteq X$ genau dann zu einer Algebra \mathcal{A} gehört, wenn ihre Komplementmenge M' Element von \mathcal{A} ist.
3. Eine σ -Algebra ist auch abgeschlossen unter unendlichen Durchschnitten, denn es gilt

$$\bigcap_{j=1}^{\infty} M_j = \left(\bigcup_{j=1}^{\infty} M'_j \right)' \in \mathcal{A}$$

für jede Folge $(M_j)_{j=1}^{\infty} \subset \mathcal{A}$.

4. Es gilt stets

$$\bigcup_{j=1}^{\infty} M_j = \bigcup_{j=1}^{\infty} \tilde{M}_j$$

mit $\tilde{M}_1 := M_1$ sowie

$$\tilde{M}_2 := M_2 \setminus M_1, \quad \tilde{M}_3 := M_3 \setminus (M_1 \cup M_2), \quad \tilde{M}_4 := M_4 \setminus (M_1 \cup M_2 \cup M_3)$$

und so weiter. Die abzählbare Vereinigung beliebiger Mengen kann also auch als abzählbare Vereinigung paarweise disjunkter Mengen geschrieben werden, wobei eine analoge Formel auch für endliche Vereinigungen gilt. Dabei gilt: Wenn alle M_j Elemente des Ringes \mathcal{R} sind, so gehört auch jede Menge \tilde{M}_j zu \mathcal{R} .

5. Ein Paar (X, \mathcal{A}) bestehend aus einer Menge X sowie einer σ -Algebra auf X wird messbarer Raum genannt.

Beispiele

1. Die Potenzmenge von X ist offensichtlich eine σ -Algebra (und damit auch Ring und Algebra). Die Menge aller endlichen Teilmengen von X , also

$$\mathcal{P}_{\text{fin}}(X) = \{M \in \mathcal{P}(X) : M \text{ ist endlich}\}$$

ist zwar immer ein Ring, aber für unendliche Mengen X keine Algebra.

2. In der Menge $X = \mathbb{R}$ wird durch

$$\mathcal{R} = \left\{ \bigcup_{i=1}^m (a_i, b_i] \text{ mit } -\infty < a_1 < b_1 < a_2 < b_2 < \dots < a_m < b_m < +\infty \right\}$$

der rechtsseitige Intervallring definiert (siehe dazu das Bild), den wir im Laufe der Vorlesung immer wieder betrachten werden.⁶ Die Anzahl der Intervalle ist dabei nicht festgelegt, d.h. die Elemente aus \mathcal{R} dürfen aus unterschiedlich vielen Teilintervallen bestehen. Beachte außerdem, dass die Ungleichungen zwischen den a_i und b_i sicherstellen, dass für jedes $M \in \mathcal{R}$ ihre Teilintervalle paarweise disjunkt und außerdem in der natürlichen Reihenfolge nummeriert sind.

Bemerkung: In $X = \mathbb{R}^n$ kann ein analoger Ring definiert werden, sofern die halboffenen Intervalle durch halboffene Quader der Dimension n ersetzt werden. Für $n = 2$ kann jeder halboffene Quader via

$$\{(x, y) \in \mathbb{R}^2 : a < x \leq b, c < y \leq d\} = (a, b] \times (c, d]$$

als das Produkt zweier halboffener Intervalle betrachtet werden und Analoges gilt in höheren Dimensionen.

3. Auf $X = \mathbb{R}$ wird auch durch

$$\mathcal{R} = \{M \subseteq \mathbb{R} : M \text{ ist Vereinigung endlich vieler Intervalle mit endlicher Länge}\}$$

ein Ring definiert, wobei diesmal alle Arten von Intervallen zugelassen sind.

Bemerkung: Die symmetrische Differenz zweier sich überlappender Intervalle, die beide offenen (oder beide abgeschlossenen) sind, besteht aus zwei halboffenen Intervallen. Es kann daher keinen Ring in \mathbb{R} geben, dessen Elemente nur aus offenen (oder nur aus abgeschlossenen) Intervallen zusammengesetzt sind. Mit halboffenen Intervallen ist dies aber möglich, wie das vorherige Beispiel zeigt.

4. Ist \mathcal{R} ein Ring in X und $\check{X} \subseteq X$ eine beliebige Teilmenge von X , so ist

$$\check{\mathcal{R}} := \{M \cap \check{X} : M \in \mathcal{R}\}$$

ein Ring in \check{X} . Analoge Aussagen gelten für Algebren und σ -Algebren.

5. Sei $f : \check{X} \rightarrow X$ eine beliebige Abbildung zwischen zwei Mengen \check{X} und X und sei \mathcal{R} ein Ring in X . Dann wird durch

$$\tilde{\mathcal{R}} := \{f^{-1}(M) : M \in \mathcal{R}\}$$

⁶Das linksseitige Analogon mit halboffenen Intervallen der Bauart $[a_j, b_j)$ hat im Prinzip dieselben Eigenschaften, wird aber weniger häufig verwendet.

ein Ring in \tilde{X} definiert. Ist \mathcal{R} sogar Algebra oder σ -Algebra, so wird auch $\tilde{\mathcal{R}}$ diese Eigenschaft besitzen.

Bemerkung: Die letzten beiden Beispiele zeigen, dass man aus gegebenen Ringen oder Algebren durch gewisse abstrakte Operationen neue Ringe oder Algebren konstruieren kann. Dies wird sehr wichtig werden.

6. Wir werden weiter unten sehen, dass zu jedem Maß auf einem Ring eine spezielle σ -Algebra gehört, auf der die *Lebesguesche Fortsetzung* des Maßes definiert ist.

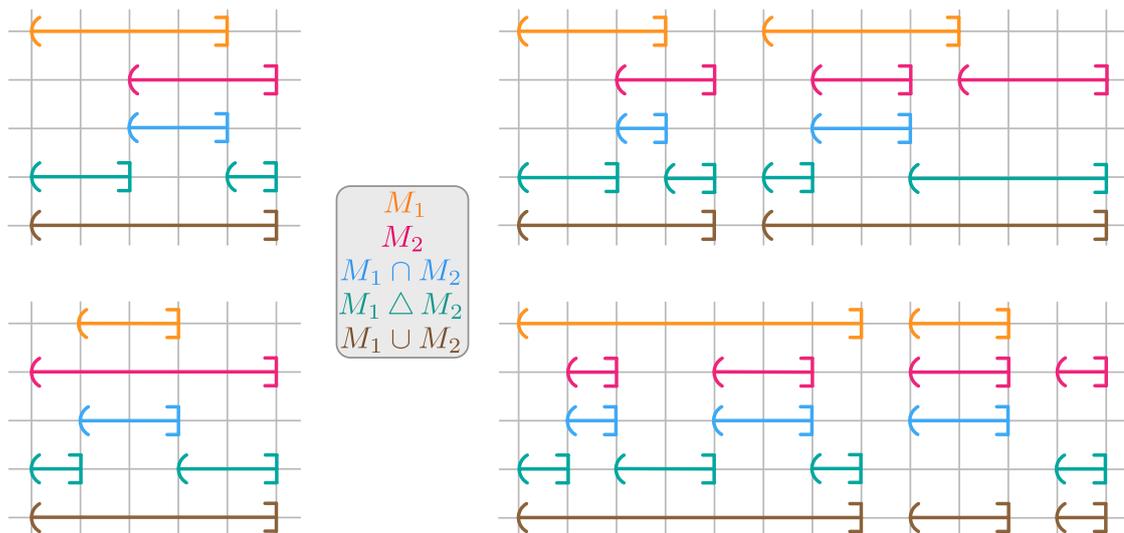


Abbildung Vier Beispiele für die Elemente und die drei symmetrischen Rechenoperationen im rechtsseitigen Intervallring, wobei die verschiedenen Mengen zur besseren Darstellung jeweils übereinander gezeichnet wurden.

Lemma (Schnitt von Mengensystemen) Der Durchschnitt beliebig vieler Ringe (bzw. Algebren bzw. σ -Algebren) in X ist selbst ein Ring (bzw. eine Algebra bzw. eine σ -Algebra) in X .

Beweis Sei I eine beliebige Indexmenge⁷ und sei \mathcal{R}_i für jedes $i \in I$ ein Ring in X . Wir wollen nun zeigen, dass auch

$$\mathcal{R} := \bigcap_{i \in I} \mathcal{R}_i = \{M \subseteq X : M \in \mathcal{R}_i \text{ für jedes } i \in I\}$$

ein Ring in X ist und betrachten dafür zwei beliebige Teilmengen M und N von X , die Elemente von \mathcal{R} sind. Für jedes feste $i \in I$ gilt dann $M, N \in \mathcal{R}_i$ und die Ringeigenschaften implizieren $M \Delta N \in \mathcal{R}_i$ sowie $M \cap N \in \mathcal{R}_i$. Dies gilt aber für jedes $i \in I$ und wir schließen, dass $M \Delta N$ und $M \cap N$ auch Elemente von \mathcal{R} sind. Außerdem enthält \mathcal{R} mindestens die leere Menge und ist daher selbst nichtleer. Der Nachweis der zusätzlichen Algebra- bzw. σ -Algebra-Eigenschaften erfolgt ganz analog. \square

Korollar (Erzeugung von Mengensystemen) Sei \mathcal{E} ein Mengensystem in X . Dann existiert genau ein kleinster Ring in X , der \mathcal{E} enthält.

⁷ I kann hier endlich, abzählbar unendlich oder gar überabzählbar unendlich sein

Beweis Wir bemerken zunächst, dass es mit $\mathcal{P}(X)$ mindestens einen Ring gibt, der \mathcal{E} enthält. Wir betrachten nun den Durchschnitt aller Ringe, die \mathcal{E} enthalten. Dieser ist nach dem Satz selbst ein Ring in X und enthält \mathcal{E} . Er ist nach Konstruktion sogar der kleinste Ring mit dieser Eigenschaft, denn er ist in allen anderen Ringen enthalten. \square

Bemerkungen

1. \mathcal{E} wird der Erzeuger genannt und wir sprechen von dem durch \mathcal{E} erzeugten Ring. Mit analogen Beweisen zeigen wir, dass \mathcal{E} auch eine Algebra sowie eine σ -Algebra erzeugt.
2. Von besonderer Bedeutung ist auch *Borelsche σ -Algebra* im $X = \mathbb{R}^d$ (bzw. in einem allgemeinen metrischen Raum X), die von allen offenen Teilmengen von X erzeugt wird und damit auch alle abgeschlossenen Teilmengen enthält, da deren Komplementmengen ja offen in X sind. Wir werden dies weiter unten genauer studieren und insbesondere die Beziehung zur *Lebesgueschen σ -Algebra* herausarbeiten.

Monotonie und Konvergenz für Folgen von Mengen Eine Folge $(M_j)_{j=1}^{\infty}$ in $\mathcal{P}(X)$ mit

$$M_1 \supseteq M_2 \supseteq M_3 \supseteq M_4 \supseteq \dots \quad \text{bzw.} \quad M_1 \subseteq M_2 \subseteq M_3 \subseteq M_4 \subseteq \dots$$

nennen wir absteigend bzw. aufsteigend, wobei wir

$$\bigcap_{j=1}^{\infty} M_j \quad \text{bzw.} \quad \bigcup_{j=1}^{\infty} M_j$$

in natürlicher Weise als den mengentheoretischen Grenzwert der Folge betrachten. Ein Mengensystem in X heißt monoton, wenn es mit jeder ab- oder aufsteigenden Folge auch ihren mengentheoretischen Grenzwert enthält.

Bemerkungen:

1. Jede σ -Algebra ist ein monotones Mengensystem.
2. Beachte, dass wir bei der mengentheoretischen Konvergenz keine Abstände benutzen (im Gegensatz zu den Konvergenzdefinitionen aus *Analysis 1+2*), sondern ausschließlich mit mengentheoretischen Relationen und Operationen arbeiten.
3. Man kann auch mengentheoretische Grenzwertkonzepte für Folgen einführen, die weder ab- noch aufsteigend sind, aber diese werden wir in dieser Vorlesung nicht verwenden. Siehe dazu etwa [Els, Abschnitt I.2]

Zusatz*

Theorem (äquivalente Charakterisierung der erzeugten σ -Algebra) Sei \mathcal{A} eine Algebra. Dann ist die von ihr erzeugte σ -Algebra gerade das kleinste monotone Mengensystem, das \mathcal{A} enthält.

Beweis* Schritt 0: Wir bezeichnen im Folgenden mit \mathcal{S} bzw. \mathcal{M} die kleinste σ -Algebra bzw. das kleinste monotone Mengensystem, die bzw. das die gegebene Algebra \mathcal{A} enthält, wobei \mathcal{S} gerade die von \mathcal{A} erzeugte σ -Algebra ist.⁸ Da \mathcal{S} monoton ist, gilt $\mathcal{M} \subseteq \mathcal{S}$. Wir werden nun umgekehrt zeigen, dass \mathcal{M} eine σ -Algebra ist, denn dies impliziert $\mathcal{S} \subseteq \mathcal{M}$.

Schritt 1: Für jedes $F \in \mathcal{M}$ betrachten wir das Mengensystem

$$\mathcal{K}(F) := \left\{ Y \subseteq X : Y \setminus F \in \mathcal{M}, F \setminus Y \in \mathcal{M}, Y \cup F \in \mathcal{M} \right\},$$

wobei per Definition die Symmetrie-Eigenschaft

$$E \in \mathcal{K}(F) \quad \Leftrightarrow \quad F \in \mathcal{K}(E)$$

für alle $E, F \in \mathcal{M}$ erfüllt ist.

Schritt 2: Für jedes $F \in \mathcal{M}$ ist $\mathcal{K}(F)$ monoton. In der Tat, für jede aufsteigende Folge $E_1 \subseteq E_2 \subseteq E_3 \subseteq \dots$ in $\mathcal{K}(F)$ und $E := \bigcup_{j=1}^{\infty} E_j$ gilt

$$E \setminus F = \bigcup_{j=1}^{\infty} (E_j \setminus F), \quad F \setminus E = \bigcap_{j=1}^{\infty} (F \setminus E_j), \quad E \cup F = \bigcup_{j=1}^{\infty} (E_j \cup F),$$

wobei $E_j \setminus F$, $F \setminus E_j$ und $E_j \cup F$ für jedes $j \in \mathbb{N}$ wegen $E_j \in \mathcal{K}(F)$ Elemente von \mathcal{M} sind. Die Monotonie von \mathcal{M} garantiert, dass die rechte Seite in jeder der beiden Teilformeln ebenfalls in \mathcal{M} liegt und wir schließen, dass E als unendliche Vereinigung der E_j auch zu $\mathcal{K}(F)$ gehört. Für absteigende Folgen argumentieren wir analog.

Schritt 3: Wir zeigen das Zwischenergebnis

$$\mathcal{M} \subseteq \mathcal{K}(F)$$

für alle $F \in \mathcal{M}$ und beginnen mit dem Spezialfall $F \in \mathcal{A}$. Für jedes $E \in \mathcal{A}$ gehören sowohl $E \setminus F$, $F \setminus E$ und $E \cup F$ zu $\mathcal{A} \subseteq \mathcal{M}$ und wir schließen, dass E auch in $\mathcal{K}(F)$ liegt. Da E beliebig war, ist ganz \mathcal{A} in dem monotonen Mengensystem $\mathcal{K}(F)$ enthalten und das Zwischenergebnis ergibt sich unmittelbar aus der Minimalität von \mathcal{M} . Im allgemeinen Fall gilt nur $F \in \mathcal{M}$ und wir müssen leicht anders argumentieren. Für ein beliebig gewähltes $E \in \mathcal{A}$ folgt $F \in \mathcal{K}(E)$ aus dem Spezialfall und die Symmetrie-Eigenschaft liefert $E \in \mathcal{K}(F)$. Insbesondere gilt wieder $\mathcal{A} \subseteq \mathcal{K}(F)$ und damit auch das Zwischenergebnis.

Schritt 4: Für beliebige $E, F \in \mathcal{M}$ folgt $E \in \mathcal{K}(F)$ aus dem Zwischenergebnis, aber per Definition meint dies gerade $E \setminus F \in \mathcal{M}$, $F \setminus E \in \mathcal{M}$ sowie $E \cup F \in \mathcal{M}$. Insbesondere ist \mathcal{M} abgeschlossen unter der Bildung von Differenzen und Vereinigungen, wobei zusätzlich $X \in \mathcal{A} \subseteq \mathcal{M}$ gilt. Hieraus folgt (siehe dazu die Übungen), ist \mathcal{M} eine Algebra ist. Sei nun $(E_j)_{j=1}^{\infty}$ eine beliebige Folge in \mathcal{M} . Dann gilt

$$\bigcup_{j=1}^{\infty} E_j = \bigcup_{j=1}^{\infty} \tilde{E}_j \quad \text{mit} \quad \tilde{E}_1 := E_1, \quad \tilde{E}_2 := E_1 \cup E_2, \quad \tilde{E}_3 := E_1 \cup E_2 \cup E_3, \quad \dots,$$

wobei die Folge $(\tilde{E}_j)_{j=1}^{\infty}$ aufsteigend ist. Die Monotonie von \mathcal{M} impliziert, dass \mathcal{M} sogar eine σ -Algebra ist. \square

Bemerkung Der Beweis ist zwar kompliziert, aber auch sehr elegant und illustriert wichtige Beweistechniken in der Maß- und Integrationstheorie. Wir werden ihn weiter unten benutzen, um die Eindeutigkeit gewisser Maßfortsetzungen zu zeigen.

⁸Wir müssen streng genommen noch zeigen, dass der Durchschnitt monotoner Mengensysteme wieder monoton ist, aber dies gelingt ganz analog zum obigen Lemma.

1.2 Maße auf Ringen und Algebren

Definition Eine Abbildung $\mu : \mathcal{R} \rightarrow [0, \infty]$ auf einem Ring \mathcal{R} nennen wir Maß, falls $\mu(\emptyset) = 0$ gilt und sofern

$$\mu\left(\bigcup_{j=1}^{\infty} M_j\right) = \sum_{j=1}^{\infty} \mu(M_j)$$

für jede Folge $(M_j)_{j=1}^{\infty}$ in \mathcal{R} erfüllt ist, deren Glieder paarweise disjunkt sind⁹ und für die die unendliche Vereinigung auf der linken Seite der Formel zu \mathcal{R} gehört.

Bemerkungen

1. Die Mengen aus \mathcal{R} werden μ -messbar genannt und den Wert $\mu(M)$ bezeichnen wir als Maß der Menge M bzgl. μ . Im Fall von $\mu(M) = 0$ sagen wir, M sei eine μ -Nullmenge.
2. Die Reihenformel wird σ -Additivität von μ genannt. Diese umfasst via

$$\mu\left(\bigcup_{j=1}^n M_j\right) = \sum_{j=1}^n \mu(M_j)$$

auch die endliche Additivität für paarweise disjunkte Mengen, denn wir können einfach $M_{n+1} = M_{n+2} = \dots = \emptyset$ setzen. Insbesondere gilt immer die Implikation

$$M_1 \cap M_2 = \emptyset \quad \implies \quad \mu(M_1 \cup M_2) = \mu(M_1) + \mu(M_2)$$

für alle $M_1, M_2 \in \mathcal{R}$.

Achtung: Die Summen- und Reihenformeln dürfen wirklich nur für paarweise disjunkte Mengen verwendet werden. Im Allgemeinen gilt zum Beispiel

$$\mu(M_1) + \mu(M_2) = \mu(M_1 \cup M_2) + \mu(M_1 \cap M_2)$$

sowie

$$\mu\left(\bigcup_{j=1}^{\infty} M_j\right) \leq \sum_{j=1}^{\infty} \mu(M_j),$$

wobei der jeweilige Beweis eine Übungsaufgabe ist und wir in der zweiten Formel wieder voraussetzen müssen, dass die unendliche Vereinigung der M_j zu \mathcal{R} gehört.

3. Jedes Maß ist monoton, denn

$$M_1 \subseteq M_2 \quad \implies \quad \mu(M_1) \leq \mu(M_2)$$

ergibt sich aus der Summenformel $\mu(M_2) = \mu(M_1) + \mu(M_2 \setminus M_1)$ sowie der Ungleichung $\mu(M_2 \setminus M_1) \geq 0$.

⁹Dies meint wie üblich, dass $M_i \cap M_j = \emptyset$ für alle $i \neq j$ gilt.

4. \mathcal{R} ist in dieser Definition ein Ring und damit im Allgemeinen keine σ -Algebra. Die unendliche Vereinigung $\bigcup_{j=1}^{\infty} M_j$ muss also nicht unbedingt zu \mathcal{R} gehören. Beachte aber, dass endliche Vereinigungen immer zu \mathcal{R} gehören und dass dann die entsprechende Additivitätsformel im paarweise disjunkten Fall erfüllt ist.¹⁰
5. Es ist ausdrücklich zugelassen, dass μ einzelnen Mengen $M \in \mathcal{R}$ den Wert $+\infty$ zuordnet. In vielen Anwendungen wird aber μ entweder allen Mengen aus \mathcal{R} einen endlichen Wert oder nur der Menge X den Wert $+\infty$ zuordnen.
6. Ist \mathcal{R} eine Algebra und gilt $\mu(X) < \infty$, so nennen wir μ finites (oder endliches) Maß. Wir sprechen hingegen von einem σ -finiten Maß, falls $X = \bigcup_{j=1}^{\infty} M_j$ mit $\mu(M_j) < \infty$ für alle j gilt, d.h. wenn X als abzählbare Vereinigung von Mengen mit endlichem Maß dargestellt werden kann.¹¹
7. Maße sind hier immer nichtnegativ, d.h. es gilt $\mu(M) \geq 0$ für alle $M \in \mathcal{R}$. Später werden wir *signierte Maße* kennenlernen, bei denen auch negative Werte zugelassen sind. Diese können aber immer eindeutig in einen Positiv- und einen Negativteil zerlegt werden, wobei beide Teile ein Maß in unserem Sinne sind.
8. Ein 3-Tupel (X, \mathcal{A}, μ) bestehend aus einer Menge X , einer σ -Algebra auf X sowie einem Maß μ auf \mathcal{A} wird oftmals Maßraum genannt. Beachte, dass hier μ auf einer σ -Algebra und nicht nur auf einem Ring definiert sein muss.

Ausblick Die zentrale Idee ist, dass die Funktion μ den „Flächeninhalt“ oder das „Volumen“ für jede Menge $M \in \mathcal{R}$ liefert bzw. vorschreibt. Wir werden im Fortgang der Vorlesung sehen, wie ausgehend von μ ein Maß- sowie ein Integrationsbegriff aufgebaut werden können. Insbesondere werden wir auch anderen Teilmengen von X (aber nicht allen) ein Maß zuordnen und für sehr viele Funktionen auf X (aber nicht für alle) in sinnvoller Weise ein entsprechendes Integral definieren. Analog geht man in der modernen Stochastik vor. Dort beschreibt X die möglichen Werte einer Zufallsvariable und μ die Wahrscheinlichkeit gewisser Ereignisse. Ausgehend von diesen können dann die Wahrscheinlichkeiten anderer Ergebnisse berechnet und der Erwartungswert sowie die Varianz als spezielle Integrale eingeführt werden.

Beispiele

1. Sei X eine beliebige Menge und $\mathcal{R} = \mathcal{P}(X)$. Dann wird durch

$$\mu(M) := \text{Anzahl der Elemente von } M$$

das sogenannte Zählmaß auf X eingeführt. Dieses ist genau dann *finit* bzw. σ -finit, wenn X nur endlich viele bzw. abzählbar viele Elemente enthält.

2. Eine abgewandelte Variante ist

$$\mu(M) := \text{Anzahl der Elemente von } M \cap Y,$$

¹⁰In der Literatur fordert man manchmal in der Definition eines Maßes, dass \mathcal{R} sogar eine σ -Algebra ist und spricht andernfalls von einem *Prämaß*. Manche Autoren führen außerdem das verwandte Konzept eines *Inhalts* ein, wobei die Funktion μ dann nicht σ -additiv, sondern nur additiv sein muss.

¹¹In diesem Falle können wir X via $X = \bigcup_{j=1}^{\infty} \tilde{M}_j$ auch als disjunkte Vereinigung schreiben, wobei jede der Mengen $\tilde{M}_j = M_j \setminus \bigcup_{k=1}^j M_k$ als Teilmenge von M_j auch ein endliches Maß besitzt.

wobei Y eine beliebige, aber vorgegebene Teilmenge von X ist, zum Beispiel eine endliche Menge $Y = \{y_1, \dots, y_m\}$ oder eine abzählbare Menge $Y = \{y_1, y_2, \dots\}$. Dieses Maß zählt für jedes $M \subseteq X$, wie viele Punkte von Y innerhalb M liegen.

3. Im rechtsseitigen Intervallring wird durch

$$\mu\left(\bigcup_{i=1}^m (a_i, b_i]\right) = \sum_{i=1}^m (b_i - a_i)$$

in sehr natürlicher Weise ein σ -finites Maß definiert, das dem eindimensionalen *Lebesgue-Maß* und *Lebesgue-Integral* zugrunde liegt (siehe dazu weiter unten). Wir werden dies weiter unten sehr viel genauer studieren, wollen aber festhalten, dass wir in der Formel wieder vorausgesetzt haben, dass die Teilintervalle der Menge paarweise disjunkt sind, d.h. dass $-\infty < a_1 < b_1 < \dots < a_m < b_m < +\infty$ gilt.

Bemerkung: Die Nichtnegativität und die Additivität von μ folgen unmittelbar, aber der Nachweis der σ -Additivität ist subtiler. Nach einigen Vereinfachungen und einer eventuellen Umsortierung von Indizes können wir das Problem jedoch auf den folgenden Spezialfall zurückführen: Sei $(\xi_j)_{j=0}^\infty$ eine monoton fallende Folge, die für $j \rightarrow \infty$ gegen $\xi \in \mathbb{R}$ konvergiert. Sei außerdem $M_j = (\xi_j, \xi_{j-1}]$ und $M = (\xi, \xi_0]$. Dann gilt $M = \bigcup_{j=1}^\infty M_j$ und das Teleskopsummenprinzip liefert

$$\sum_{j=1}^n \mu(M_j) = \sum_{j=1}^n (\xi_{j-1} - \xi_j) = \xi_0 - \xi_n.$$

Im Grenzübergang $n \rightarrow \infty$ erhalten wir schließlich

$$\sum_{j=1}^\infty \mu(M_j) = \lim_{n \rightarrow \infty} \sum_{j=1}^n \mu(M_j) = \xi_0 - \xi = \mu(M)$$

und damit das gewünschte Ergebnis.¹²

4. Ist $W : \mathbb{R} \rightarrow \mathbb{R}$ eine monoton wachsende und rechtsseitig stetige Funktion,¹³ so liefert

$$\mu\left(\bigcup_{i=1}^m (a_i, b_i]\right) = \sum_{i=1}^m W(b_i) - W(a_i)$$

das *Lebesgue-Stieltjes-Maß* zu W auf dem rechtsseitigen Intervallring, wobei die Lebesguesche Variante gerade dem Spezialfall $W(s) = s$ entspricht. Besonders wichtig ist auch der Fall

$$0 \leq W(s) \leq 1 \text{ für alle } s \in \mathbb{R} \quad \text{mit} \quad 0 = \lim_{s \searrow -\infty} W(s) \text{ und } 1 = \lim_{s \nearrow +\infty} W(s),$$

¹²Beachte, dass wir hier das halboffene Intervall M durch eine absteigende Folge von halboffenen Intervallen erzeugt haben. Ist $(\xi_j)_{j=0}^\infty$ eine monoton wachsende Folge mit Grenzwert ξ , so sind die Intervalle $M_j = (\xi_{j-1}, \xi_j]$ aufsteigend, aber $\bigcup_{j=1}^\infty M_j$ ist diesmal das offene Intervall (ξ_0, ξ) und damit nicht in \mathcal{R} enthalten. Insbesondere gibt es in diesem Fall kein Problem mit der σ -Additivität.

¹³Oftmals wird zusätzlich gefordert, dass der linksseitigen Grenzwert $\lim_{\varepsilon \searrow 0} W(s - \varepsilon)$ für jedes $s \in \mathbb{R}$ existiert, wobei dieser wegen der Monotonie niemals größer als $W(s) = \lim_{\varepsilon \searrow 0} W(s + \varepsilon)$ sein kann. Man nennt dann W eine càdlàg-Funktion (*continue à droite, limite à gauche*).

in dem W als Verteilungsfunktion einer reellen Zufallsvariable angesehen werden kann. Ist W sogar stetig differenzierbar mit Ableitung $w : \mathbb{R} \rightarrow \mathbb{R}$, so kann die obige Formel nach dem Fundamentalsatz der Analysis auch als

$$\mu\left(\bigcup_{i=1}^m (a_i, b_i]\right) = \sum_{i=1}^m \int_{a_i}^{b_i} w(x) dx$$

geschrieben werden, wobei im stochastischen Kontext w gerade die *Dichte* der Wahrscheinlichkeitsverteilung W ist und der Bedingung $\int_{-\infty}^{+\infty} w(x) dx = 1$ genügt.

Bemerkung: Die rechtsseitige Stetigkeit von W ist wichtig für die σ -Additivität. Mit den Notationen aus dem vorherigen Beispiel gilt

$$\sum_{j=1}^n \mu(M_j) = \sum_{j=1}^n W(\xi_j) - W(\xi_{j-1}) = W(\xi_0) - W(\xi_n)$$

und im Limes $n \rightarrow \infty$ ergibt sich auf der rechten Seite nur dann der richtige Wert $W(\xi_0) - W(\xi) = \mu(M)$, wenn W rechtsseitig stetig ist.

Definition Ein Maß μ auf einem Ring \mathcal{R} heißt vollständig, wenn jede Teilmenge einer μ -Nullmenge selbst μ -messbar ist.

Bemerkungen

1. Bei einem vollständigen Maß gilt also die Implikation

$$N \subseteq M, \quad M \in \mathcal{R}, \quad \mu(M) = 0 \quad \implies \quad N \in \mathcal{R},$$

wobei die Monotonie und nicht Negativität von μ garantiert, dass N wegen

$$0 \leq \mu(N) \leq \mu(M) = 0$$

auch eine μ -Nullmenge ist.

2. Wir werden sehen, dass vollständige Maße deutlich bessere Eigenschaften als unvollständige besitzen.
3. Viele Maße sind nicht vollständig, weil der zugrunde liegende Ring \mathcal{R} „zu klein“ ist. Wir werden im nächsten Abschnitt diskutieren, wie dieses Problem behoben werden kann.

Lemma (Grenzwertformel für auf- oder absteigende Folgen) Sei μ ein Maß auf der σ -Algebra \mathcal{A} . Dann gelten die Implikationen

$$M_1 \subseteq M_2 \subseteq M_3 \subseteq M_4 \subseteq \dots \quad \implies \quad \mu\left(\bigcup_{j=1}^{\infty} M_j\right) = \lim_{j \rightarrow \infty} \mu(M_j)$$

und

$$M_1 \supseteq M_2 \supseteq M_3 \supseteq M_4 \supseteq \dots \quad \implies \quad \mu\left(\bigcap_{j=1}^{\infty} M_j\right) = \lim_{j \rightarrow \infty} \mu(M_j)$$

sofern $(M_j)_{j=1}^{\infty}$ eine Folge in \mathcal{A} ist, für die $\mu(M_j) < \infty$ für alle $j \in \mathbb{N}$ gilt.

Beweis *Teil 1*: Wir benutzen das *Zwiebelschalenprinzip* (siehe auch die Abbildung) und betrachten die Mengen

$$K_1 := M_1, \quad K_2 := M_2 \setminus M_1, \quad \dots, \quad K_n := M_n \setminus M_{n-1}, \quad \dots$$

die alle zu \mathcal{A} gehören. Sie sind paarweise disjunkt und erfüllen

$$\bigcup_{j=1}^n K_j = M_n = \bigcup_{j=1}^n M_j, \quad \sum_{j=1}^n \mu(K_j) = \mu(M_n)$$

für jedes $n \in \mathbb{N}$ nach Voraussetzung und Konstruktion sowie aufgrund der Additivität von μ . Außerdem gilt

$$\bigcup_{j=1}^{\infty} K_j = M := \bigcup_{j=1}^{\infty} M_j$$

und mit der σ -Additivität von μ erhalten wir mit

$$\mu(M) = \sum_{j=1}^{\infty} \mu(K_j) = \lim_{n \rightarrow \infty} \sum_{j=1}^n \mu(K_j) = \lim_{n \rightarrow \infty} \mu(M_n)$$

die gewünschte Grenzwertaussage.

Teil 2: Diesmal verwenden wir die Mengen

$$K_1 := M_1 \setminus M_2, \quad K_2 := M_2 \setminus M_3, \quad \dots, \quad K_n := M_n \setminus M_{n+1}, \quad \dots,$$

die wieder paarweise disjunkt sind und sämtlich in \mathcal{A} liegen. Insbesondere gilt

$$M_j = K_j \cup M_{j+1}, \quad K_j \cap M_{j+1} = \emptyset, \quad \mu(M_j) = \mu(K_j) + \mu(M_{j+1})$$

für alle $j \in \mathbb{N}$ und mit dem Teleskopsummenprinzip ergibt sich

$$\sum_{j=1}^n \mu(K_j) = \sum_{j=1}^n (\mu(M_j) - \mu(M_{j+1})) = \mu(M_1) - \mu(M_{n+1})$$

für alle $n \in \mathbb{N}$. Unsere Definitionen sowie die σ -Additivität von μ garantieren weiterhin

$$\bigcup_{j=1}^{\infty} K_j = M_1 \setminus M, \quad \sum_{j=1}^{\infty} \mu(K_j) = \mu(M_1 \setminus M) \quad \text{mit} \quad M := \bigcap_{j=1}^{\infty} M_j,$$

wobei außerdem

$$\mu(M_1) = \mu(M) + \mu(M_1 \setminus M)$$

erfüllt ist. Nach Einsetzen der Teilresultate erhalten wir die Konvergenzaussage

$$\mu(M_1) - \mu(M) = \sum_{j=1}^{\infty} \mu(K_j) = \lim_{n \rightarrow \infty} \sum_{j=1}^n \mu(K_j) = \lim_{n \rightarrow \infty} (\mu(M_1) - \mu(M_{n+1})),$$

und diese ist nach den Rechenregeln für konvergente Zahlenfolgen äquivalent zur Behauptung. \square

Bemerkungen

1. Der Beweis zeigt für eine auf- bzw. absteigende Folge, dass $\mu(M_j)$ für $j \rightarrow \infty$ von unten bzw. oben gegen den jeweiligen Grenzwert konvergiert. Alternativ kann dies auch mit der Monotonie von Maßen begründet werden.
2. Ein analoges Ergebnis gilt auch in Ringen, wobei dann explizit gefordert werden muss, dass die unendliche Vereinigung bzw. der unendliche Schnitt der Mengen M_j auch im Ring liegt.
3. Bei aufsteigenden Folgen kann auf die Forderung, dass alle M_j endliches Maß besitzen, verzichtet werden, wobei dann die Limesformel im Sinne von $\infty = \infty$ erfüllt sein kann. Bei absteigenden Folgen braucht man nur zu fordern, dass $\mu(M_j)$ ab einem gewissen Index j_* endlich ist.

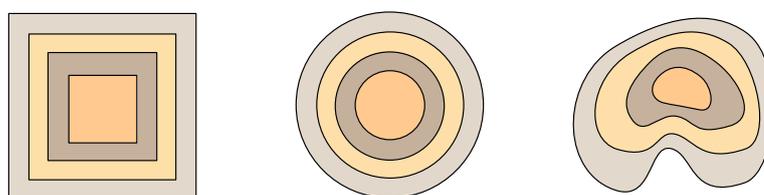


Abbildung Das *Zwiebelschalenprinzip* besagt, dass eine endliche oder unendliche Folge von entweder auf- oder absteigenden Mengen (zum Beispiel Quadrate, Kreisscheiben oder was anderes) in sehr natürlicher Weise als disjunkte Vereinigung anderer Mengen (eckige, runde, oder andere ringförmige Schalen) dargestellt werden kann.

1.3 Vervollständigung eines Maßes

Vorbemerkung Wir zeigen, dass jedes unvollständige Maß μ auf einem Ring \mathcal{R} zu einem vollständigen Maß fortgesetzt werden kann, wobei dieses dann auf einer σ -Algebra definiert ist, die größer als \mathcal{R} ist. Diese Konstruktion ist sehr abstrakt, aber auch sehr elegant und spiegelt sehr tiefe Erkenntnisse wider. Sie ist auch für den Fortgang dieser Vorlesung wichtig, da sie uns das Lebesgue-Maß und später dann das Lebesgue-Integral bereitstellen wird, auf dem viele unserer Resultate beruhen.

Definition Zu jedem Maß μ auf einem Ring \mathcal{R} in X wird durch

$$\mu^*(Y) = \inf \left\{ \sum_{j=1}^{\infty} \mu(M_j) : Y \subseteq \bigcup_{j=1}^{\infty} M_j \text{ mit } M_j \in \mathcal{R} \text{ für alle } j \in \mathbb{N} \right\}$$

eine Funktion $\mu^* : \mathcal{P}(X) \rightarrow [0, \infty]$ definiert, die wir das äußere Maß von μ nennen.

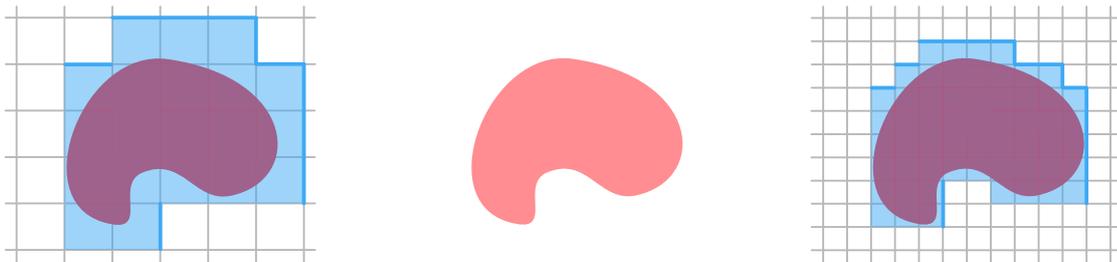


Abbildung Eine Menge $Y \subset X = \mathbb{R}^2$ (rot) sowie zwei endliche Überdeckungen (jeweils blau) mit Mengen M_j , wobei diese hier paarweise disjunkt und als halboffene Quadrate gewählt wurden (siehe dazu das zweidimensionale Lebesgue-Maß weiter unten). Das äußere Maß von Y , also der Wert $\mu^*(Y)$, quantifiziert das Infimum bzgl. aller Überdeckungen von Y , wobei jede aus abzählbar unendlich vielen Mengen des Ringes \mathcal{R} bestehen kann. Beachte auch, dass die quadratischen Raster im Hintergrund nicht wichtig sind und nur der besseren Darstellung wegen gezeichnet wurden.

Diskussion

1. Anschaulich kann die Formel wie folgt verstanden werden. Um $\mu^*(Y)$ für eine beliebige Teilmenge $Y \subseteq X$ zu bestimmen, betrachten wir zunächst alle Möglichkeiten, Y durch die Vereinigung von höchstens abzählbar vielen Mengen M_j aus dem Ring zu überdecken, wobei auch endliche Überdeckungen zugelassen sind (dies entspricht dem Fall $M_j = \emptyset$ für $j \geq j_*$) und die M_j nicht unbedingt paarweise disjunkt sein müssen. Anschließend suchen wir eine Überdeckung von Y , sodass die Summe aller $\mu(M_j)$ möglichst klein wird, wobei es im Allgemeinen kein Minimum, sondern nur ein Infimum geben wird und bei einer optimalen Überdeckung die Mengen M_j in aller Regel paarweise disjunkt sein werden.
2. Die Funktion μ^* wird im Allgemeinen kein Maß im Sinne unserer Definition von oben sein, denn auf der gesamten Potenzmenge von X ist sie nicht additiv, sondern nur subadditiv. Insbesondere haben wir den Definitionsbereich mit $\mathcal{P}(X)$ „zu groß“ gewählt. Die Einschränkung von μ^* auf ein geeignet gewähltes kleineres Mengensystem wird aber die gesuchte Vervollständigung von μ sein.

3. *Ausblick*: In vielen Ringen kann auch ein *inneres Maß* μ_* konstruiert werden, bei dem $Y \subseteq X$ mittels geeigneter Mengen $M_j \in \mathcal{R}$ von innen ausgeschöpft wird. Dieses Konzept brauchen wir aber im Moment noch nicht.

Bemerkungen

1. Für jede Menge $M \in \mathcal{R}$ gilt $\mu^*(M) = \mu(M)$.

Beweis: Für jede abzählbare Überdeckung von M ergibt sich

$$\mu(M) \leq \mu\left(\bigcup_{j=1}^{\infty} M_j\right) \leq \sum_{j=1}^{\infty} \mu(M_j)$$

aus der Monotonie sowie der Subadditivität von μ und die Infimumsbildung über alle Überdeckungen liefert $\mu(M) \leq \mu^*(M)$. Andererseits können wir M trivialerweise durch $M_1 = M$ und $M_2 = M_3 = \dots = \emptyset$ überdecken und erhalten $\mu^*(M) \leq \mu(M)$. \square

2. Das äußere Maß μ^* ist monoton, denn die Implikation

$$\tilde{Y} \subseteq Y \quad \implies \quad \mu^*(\tilde{Y}) \leq \mu^*(Y)$$

gilt für alle $Y, \tilde{Y} \in \mathcal{P}(X)$.

Beweis: Jede Überdeckung von Y überdeckt trivialerweise auch \tilde{Y} (aber nicht umgekehrt). Insbesondere bilden wir bei der Bestimmung von $\mu^*(\tilde{Y})$ das Infimum über eine Menge, die mindestens genauso groß ist wie die Menge, die wir bei der Berechnung von $\mu^*(Y)$ betrachten. Daher kann $\mu^*(\tilde{Y})$ niemals größer als $\mu^*(Y)$ sein. \square

3. Für jedes $Y \in \mathcal{P}(X)$ und jedes $\varepsilon > 0$ existiert eine Folge $(M_j)_{j=1}^{\infty}$ von Mengen aus \mathcal{R} mit

$$Y \subseteq \bigcup_{j=1}^{\infty} M_j, \quad \mu^*(Y) \leq \sum_{j=1}^{\infty} \mu(M_j) \leq \mu^*(Y) + \varepsilon,$$

d.h. $\mu^*(Y)$ kann beliebig genau durch Maße von Mengen aus dem Ring \mathcal{R} angenähert werden. Diese nützliche Approximation von oben ergibt sich aus den Eigenschaften des Infimums (siehe *Analysis 1*).

Lemma (σ -Subadditivität von μ^*) Es gilt

$$\mu^*\left(\bigcup_{k=1}^{\infty} Y_k\right) \leq \sum_{k=1}^{\infty} \mu^*(Y_k)$$

für jede Folge $(Y_k)_{k=0}^{\infty}$ in $\mathcal{P}(X)$.

Beweis Um die gewünschte Abschätzung für gegebene Mengen Y_k zu zeigen, setzen wir $Y := \bigcup_{k=1}^{\infty} Y_k$ und wählen $\varepsilon > 0$ beliebig. Außerdem können wir annehmen, dass $\mu^*(Y_k) < \infty$ für alle $k \in \mathbb{N}$ gilt, da andernfalls die Behauptung trivialerweise richtig

ist. Für jedes k können wir Y_k von oben approximieren, d.h. wir wählen eine Folge $(M_{k,j})_{j=1}^{\infty}$ in \mathcal{R} mit

$$Y_k \subseteq \bigcup_{j=1}^{\infty} M_{k,j}, \quad \sum_{j=1}^{\infty} \mu(M_{k,j}) \leq \mu^*(Y_k) + \varepsilon 2^{-k}.$$

Wegen

$$Y \subseteq \bigcup_{k=1}^{\infty} \bigcup_{j=1}^{\infty} M_{k,j}$$

erhalten wir

$$\begin{aligned} \mu^*(Y) &\leq \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} \mu(M_{k,j}) = \sum_{k=1}^{\infty} \left(\sum_{j=1}^{\infty} \mu(M_{k,j}) \right) \leq \sum_{k=1}^{\infty} (\mu^*(Y_k) + \varepsilon 2^{-k}) \\ &\leq \sum_{k=1}^{\infty} \mu^*(Y_k) + \varepsilon \sum_{k=1}^{\infty} 2^{-k} = \sum_{k=1}^{\infty} \mu^*(Y_k) + \varepsilon, \end{aligned}$$

wobei wir die Definition von μ^* , die Eigenschaften von Reihen nichtnegativer reeller Zahlen sowie die Formel $\sum_{k=1}^{\infty} 2^{-k} = 1$ benutzt haben. Da ε beliebig war, folgt die Behauptung. \square

Bemerkungen

1. Es gilt auch die (endliche) Subadditivität

$$\mu^*\left(\bigcup_{k=1}^l Y_k\right) \leq \sum_{k=1}^l \mu^*(Y_k),$$

denn wir können das Lemma mit der speziellen Wahl $Y_{l+1} = Y_{l+2} = \dots = \emptyset$ auswerten.

2. Mit $l = 2$ sowie $Y_1 = Y \cap E$ und $Y_2 = Y \cap E'$ erhalten wir den Spezialfall

$$\mu^*(Y) \leq \mu^*(Y \cap E) + \mu^*(Y \cap E'),$$

der für alle Teilmengen Y und E von X gilt.

Definition Eine Menge $E \subseteq X$ heißt μ^* -messbar, sofern

$$\mu^*(Y) = \mu^*(Y \cap E) + \mu^*(Y \cap E')$$

für jedes $Y \subseteq X$ gilt. Die Menge aller μ^* -messbaren Mengen bezeichnen wir mit \mathcal{A}^* .

Interpretation

1. Diese Definition wird sich als ausgesprochen mächtig erweisen, denn sie erlaubt es uns ein gegebenes Maß auf einem Ring \mathcal{R} in natürlicher Weise auszudehnen und auch Mengen, die gar nicht zu \mathcal{R} gehören, in sinnvoller Weise ein Maß zuzuweisen. Siehe dazu das nachfolgende Theorem.

- Auf einer sehr informellen Ebene können wir μ^* -Messbarkeit als „gut“ bzw. ihr Fehlen als „schlecht“ interpretieren und damit den Inhalt der Definition wie folgt verstehen: Wenn wir eine beliebige — und vielleicht sehr schlechte — Menge Y mittels einer guten Menge E in zwei disjunkte Teile zerlegen, so sind die Einzelteile $Y \cap E$ und $Y \cap E'$ vielleicht nicht gut, aber eben *nicht* schlechter als Y . Insbesondere gilt immer noch die Summenformel bzgl. μ^* .
- Für jedes $E \in \mathcal{A}^*$ und alle $Y \subseteq X$ gilt wegen der Subadditivität von μ^* auch

$$\mu^*(E) \leq \mu^*(E \cap Y) + \mu^*(E \cap Y'),$$

aber diesmal wird es sich im Allgemeinen um eine Ungleichung handeln. Denn wenn wir eine gute Menge E mittels einer schlechten Menge Y disjunkt zerlegen, so werden die entstehenden Teile schlechter als E sein.

- Es wird sich zeigen, dass erstaunlich viele Mengen gut sind und wir werden uns ziemlich anstrengen müssen, um die Existenz schlechter Mengen nachzuweisen. An einer Tatsache werden wir aber nicht vorbeikommen: Für viele wichtige Maße μ gibt es Mengen, die nicht μ^* -messbar sind.

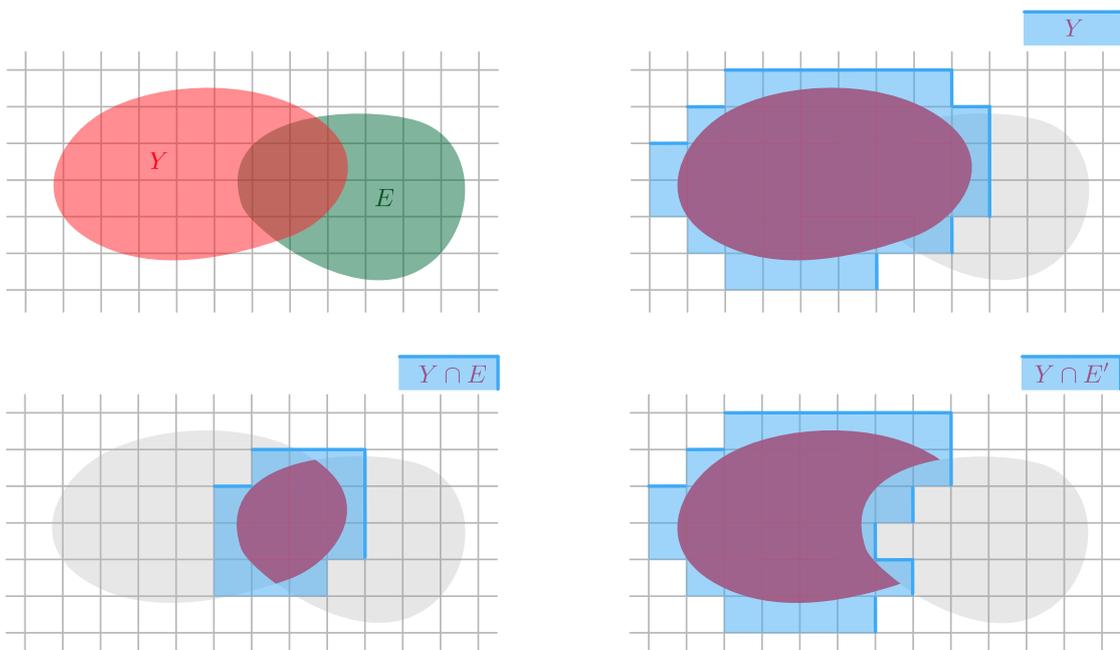


Abbildung Mittels einer Menge $E \subseteq X$ (grün) kann jede andere Menge $Y \subseteq X$ (rot) in zwei disjunkte Teile zerlegt werden. Ist E μ^* -messbar, so wird die Summe von $\mu^*(Y \cap E)$ und $\mu^*(Y \cap E')$ gerade $\mu^*(Y)$ ergeben, wobei die drei äußeren Maße jeweils separat und durch Infimumsbildung bzgl. geeigneter Überdeckungen zu bestimmen sind. Ist E nicht μ^* -messbar, so können die Teile $Y \cap E$ und $Y \cap E'$ zusammen mehr äußeres Maß haben als Y , aber dieser Fall kann graphisch *nicht* dargestellt werden.

Theorem (Eigenschaften von μ^* und Lebesguescher Abschluss von μ)
Es gelten die folgenden Aussagen:

- Es gilt $\mathcal{R} \subseteq \mathcal{A}^*$ und $\mu^*(M) = \mu(M)$ für alle $M \in \mathcal{R}$.
- \mathcal{A}^* ist eine σ -Algebra in X .
- Die Einschränkung von μ^* auf \mathcal{A}^* ist ein vollständiges Maß.

Beweis Teil 1: Seien $M \in \mathcal{R}$ und $Y \subseteq X$ beliebig fixiert. Wir wollen zeigen, dass $\mu^*(Y) = \mu^*(Y \cap M) + \mu^*(Y \cap M')$ gilt und betrachten dazu eine Folge $(M_j)_{j=1}^\infty$ in \mathcal{R} mit $Y \subseteq \bigcup_{j=1}^\infty M_j$. Dann gilt

$$Y \cap M \subseteq \bigcup_{j=1}^\infty M_j \cap M, \quad Y \cap M' \subseteq \bigcup_{j=1}^\infty M_j \cap M',$$

wobei die Mengen $M_j \cap M$ und $M_j \cap M'$ alle zum Ring \mathcal{R} gehören.¹⁴ Die Definition von μ^* garantiert daher

$$\mu^*(Y \cap M) \leq \sum_{j=1}^\infty \mu(M_j \cap M), \quad \mu^*(Y \cap M') \leq \sum_{j=1}^\infty \mu(M_j \cap M')$$

und nach Summation auch

$$\mu^*(Y \cap M) + \mu^*(Y \cap M') \leq \sum_{j=1}^\infty \left(\mu(M_j \cap M) + \mu(M_j \cap M') \right) = \sum_{j=1}^\infty \mu(M_j),$$

wobei wir die Rechenregeln für Reihen nichtnegativer Zahlen sowie die Additivität von μ benutzt haben, die die Summenformel $\mu(M_j) = \mu(M_j \cap M) + \mu(M_j \cap M')$ für jedes $j \in \mathbb{N}$ sicherstellt. Durch Infimumsbildung bzgl. aller zugelassenen Folgen $(M_j)_{j=1}^\infty$ erhalten wir

$$\mu^*(M \cap Y) + \mu^*(M \cap Y') \leq \mu^*(Y).$$

Weil die umgekehrte Ungleichung ein Spezialfall der Subadditivität von μ^* ist, haben wir insgesamt gezeigt, dass M wirklich μ^* -messbar ist und damit zu \mathcal{A}^* gehört. Wir hatten auch weiter oben schon $\mu^*(M) = \mu(M)$ hergeleitet.

Teil 2a: Aus $\mu^*(\emptyset) = \mu(\emptyset) = 0$ folgt

$$\emptyset, X \in \mathcal{A}^*$$

und die Definition von μ^* impliziert außerdem die Äquivalenz

$$E \in \mathcal{A}^* \Leftrightarrow E' \in \mathcal{A}^*.$$

Seien nun E, F beliebig in \mathcal{A}^* fixiert. Wir wollen zeigen, dass auch $E \cup F$ zu \mathcal{A}^* gehört und bemerken zunächst, dass $Y \cap (E \cup F)' = Y \cap E' \cap F'$ gilt und dass sich

$$\mu^*(Y \cap (E \cup F)) \leq \mu^*(Y \cap E) + \mu^*(Y \cap E' \cap F)$$

aus der Subadditivität von μ^* sowie der Formel $Y \cap (E \cup F) = (Y \cap E) \cup (Y \cap E' \cap F)$ ergibt. Dies liefert die Abschätzungskette

$$\begin{aligned} \mu^*(Y) &\leq \mu^*(Y \cap (E \cup F)) + \mu^*(Y \cap (E \cup F)') \\ &\leq \mu^*(Y \cap E) + \mu^*(Y \cap E' \cap F) + \mu^*(Y \cap E' \cap F') \\ &\leq \mu^*(Y \cap E) + \mu^*(Y \cap E') \\ &\leq \mu^*(Y), \end{aligned}$$

¹⁴Wir benutzen hier die Definition von Ringen und die Formel $M_j \cap M' = M_j \setminus M = (M_j \triangle M) \cap M_j$.

wobei wir im ersten Schritt die Subadditivität von μ^* , im zweiten die Hilfsformeln von oben sowie im dritten bzw. vierten Schritt die μ^* -Messbarkeit von F bzw. E ausgewertet haben. Insbesondere sind alle Ungleichungen in der Kette eigentlich Gleichungen und wegen des ersten Gleichheitszeichens schließen wir, dass $E \cup F$ zu \mathcal{A}^* gehört. Wir haben also bewiesen, dass das Mengensystem \mathcal{A}^* abgeschlossen ist unter der Bildung von endlichen Vereinigungen sowie von Komplementmengen. Hieraus folgt, dass \mathcal{A}^* eine Algebra ist.¹⁵

Vorlesung 02-2 : 04. November

Teil 2b: Wir betrachten eine Folge $(E_j)_{j=1}^\infty$ in \mathcal{A}^* und wollen zeigen, dass auch

$$E := \bigcup_{j=1}^{\infty} E_j$$

ein Element von \mathcal{A}^* ist, wobei wir o.B.d.A annehmen dürfen, dass die Folgenglieder paarweise disjunkt sind, d.h. dass $E_i \cap E_j = \emptyset$ für $i \neq j$ gilt.¹⁶ Dazu wählen wir $Y \subseteq X$ beliebig und bemerken zunächst, dass die Formel $E_1 \cap E_2 = \emptyset$ bzw. $E_2 \subset E_1'$ das Hilfsresultat

$$(Y \cap (E_1 \cup E_2)) \cap E_1 = ((Y \cap E_1) \cup (Y \cap E_2)) \cap E_1 = Y \cap E_1$$

bzw.

$$(Y \cap (E_1 \cup E_2)) \cap E_1' = ((Y \cap E_1) \cup (Y \cap E_2)) \cap E_1' = Y \cap E_2$$

impliziert. Anschließend leiten wir damit die Formel

$$\begin{aligned} \mu^*(Y \cap (E_1 \cup E_2)) &= \mu^*((Y \cap (E_1 \cup E_2)) \cap E_1) + \mu^*((Y \cap (E_1 \cup E_2)) \cap E_1') \\ &= \mu^*(Y \cap E_1) + \mu^*(Y \cap E_2) \end{aligned}$$

her, wobei das erste Gleichheitszeichen eine direkte Konsequenz der μ^* -Messbarkeit von E_1 ist. Durch vollständige Induktion ergibt sich

$$\mu^*(Y \cap (E_1 \cup \dots \cup E_n)) = \mu^*(Y \cap E_1) + \dots + \mu^*(Y \cap E_n)$$

für jedes $n \in \mathbb{N}$, wobei auch $E_1 \cup \dots \cup E_n$ ein Element von \mathcal{A}^* ist (siehe Teil 2a). Daher gilt auch

$$\begin{aligned} \mu^*(Y) &= \mu^*(Y \cap (E_1 \cup \dots \cup E_n)) + \mu^*(Y \cap (E_1 \cup \dots \cup E_n)') \\ &\geq \sum_{j=1}^n \mu^*(Y \cap E_j) + \mu^*(Y \cap E'), \end{aligned}$$

¹⁵Wir hatten dies schon in den Übungen diskutiert. Die Formeln

$$E \cap F = (E' \cup F')' \quad \text{und} \quad E \Delta F = (E \cup F')' \cup (E' \cup F)',$$

stellen nämlich sicher, dass wir innerhalb von \mathcal{A}^* auch immer Durchschnitte und symmetrische Differenzen bilden können.

¹⁶Jede abzählbare Vereinigung von Elementen eines Ringes kann nämlich auch als disjunkte Vereinigung geschrieben werden. Siehe dazu die Bemerkung oben nach der Definition von Algebren.

wobei wir für den zweiten Term die Formel $(E_1 \cup \dots \cup E_n)' \supseteq E'$ sowie die Monotonie von μ^* verwendet haben, und der Grenzübergang $n \rightarrow \infty$ liefert

$$\mu^*(Y) \geq \sum_{j=1}^{\infty} \mu^*(Y \cap E_j) + \mu^*(Y \cap E').$$

Andererseits folgt

$$\mu^*(Y) \leq \mu^*(Y \cap E) + \mu^*(Y \cap E'), \quad \mu^*(Y \cap E) \leq \sum_{j=1}^{\infty} \mu^*(Y \cap E_j)$$

aus der Subadditivität von μ^* . Nach Vergleich der oberen und unteren Schranken für $\mu^*(Y)$ ergibt sich schließlich

$$\mu^*(Y) = \mu^*(Y \cap E) + \mu^*(Y \cap E'), \quad \mu^*(Y \cap E) = \sum_{j=1}^{\infty} \mu^*(Y \cap E_j).$$

Da Y beliebig war, folgt $E \in \mathcal{A}^*$ aus dem ersten Teil dieser Formel und wir haben insgesamt gezeigt, dass \mathcal{A}^* wirklich eine σ -Algebra ist.

Teil 3: Es gilt $\mu^*(\emptyset) = 0$ und per Definition auch $\mu^*(E) \in [0, \infty]$ für alle $E \in \mathcal{A}^*$. Wir haben auch schon die σ -Additivität von μ^* gezeigt, denn wenn wir den zweiten Teil der letzten Formel für den Spezialfall $Y = E$ auswerten, erhalten wir

$$\mu^*\left(\bigcup_{j=1}^{\infty} E_j\right) = \sum_{j=1}^{\infty} \mu^*(E_j).$$

Insbesondere ist μ^* ein Maß auf \mathcal{A}^* . Wir betrachten nun eine beliebige μ^* -Nullmenge $N \in \mathcal{A}^*$ sowie eine beliebige Teilmenge E von N . Für alle $Y \subseteq X$ gilt

$$\mu^*(Y) \leq \mu^*(Y \cap E) + \mu^*(Y \cap E') \leq \mu^*(N) + \mu^*(Y) \leq \mu^*(Y),$$

wobei die erste Ungleichung wieder aus der Subadditivität von μ^* folgt und die dritte wegen $\mu^*(N) = 0$ gilt. Die zweite ergibt sich hingegen aufgrund von $Y \cap E \subseteq E \subseteq N$ und $Y \cap E' \subseteq Y$ aus der Monotonie von μ^* . Wir erhalten für jedes E und alle Y die Formel $\mu^*(Y) = \mu^*(Y \cap E) + \mu^*(Y \cap E')$, die schließlich $E \in \mathcal{A}^*$ und damit die Vollständigkeit von μ^* impliziert. \square

Bemerkungen

1. Die Einschränkung von μ^* auf \mathcal{A}^* — die oftmals als $\mu^*|_{\mathcal{A}^*}$ geschrieben wird — ist die Lebesguesche Fortsetzung von μ .
2. Das Theorem wird oftmals wie folgt verwendet: Wenn ein Maß μ auf einem Ring \mathcal{R} gegeben ist, so können wir dieses durch μ^* auf \mathcal{A}^* ersetzen und anschließend mit einem vollständigen Maß auf einer σ -Algebra weiterarbeiten. Beachte auch, dass sich dabei die Finitheit oder σ -Finitheit von μ auf μ^* überträgt.
3. Die Aussagen des Theorems können wie folgt verschärft werden: \mathcal{A}^* enthält sogar alle μ^* -Nullmengen aus $\mathcal{P}(X)$, denn jede Teilmenge $N \subseteq X$ mit $\mu^*(N) = 0$ ist μ^* -messbar.

Beweis: Für jedes $Y \subseteq X$ folgt

$$\mu(Y) \leq \mu^*(Y \cap N) + \mu(Y \cap N') \leq \mu^*(N) + \mu^*(Y) \leq 0 + \mu^*(Y)$$

aus der Additivität sowie der Monotonie von μ^* und wir schließen, dass alle Ungleichungen eigentlich Gleichungen sind. Daher gilt $N \in \mathcal{A}^*$. \square

4. Die σ -Algebra \mathcal{A}^* ist im Allgemeinen deutlich größer als die vom Ring \mathcal{R} erzeugte σ -Algebra $\mathcal{A}^\#$. Siehe dazu auch das Theorem weiter unten zur Charakterisierung von \mathcal{A}^* .
5. Es gilt die Formel

$$(\mu^*|_{\mathcal{A}^*})^* = \mu^*,$$

d.h. μ und seine Lebesguesche Fortsetzung besitzen dasselbe äußere Maß. Insbesondere kann die Lebesguesche Fortsetzung von μ nicht weiter fortgesetzt werden (jedenfalls nicht mit dem von uns benutzten Verfahren).

Bemerkung: Der Beweis der Formel gelingt in wenigen Zeilen und braucht nur die Definition und die σ -Subadditivität von äußeren Maßen sowie die Formel $\mu^*|_{\mathcal{R}} = \mu$, die sich aus dem ersten Teil des Theorems ergibt.

**-Hausaufgabe (10 Superbonuspunkte): Führen Sie diesen Beweis!

6. Für jedes σ -finite Maß μ kann es neben der Lebesgueschen Fortsetzung im Prinzip keine weitere geben. Wir wollen diesen Aspekt hier nicht vertiefen, sondern verweisen auf die Literatur — zum Beispiel [Els, Abschnitt II.5] — sowie den Eindeigkeitssatz weiter unten, der zumindest ein Teilproblem abdeckt.

Theorem (Approximationsatz in \mathcal{A}^*) Für jedes $E \in \mathcal{A}^*$ mit $\mu^*(E) < \infty$ und alle $\varepsilon > 0$ existiert ein $F \in \mathcal{R}$, sodass $\mu^*(E \triangle F) \leq \varepsilon$.

Beweis Wahl der approximierenden Menge: Für gegebenes E und ε wählen wir eine Approximation von oben, d.h. eine Folge $(M_j)_{j=1}^\infty \subseteq \mathcal{R}$ mit

$$E \subseteq \bigcup_{j=1}^{\infty} M_j, \quad \sum_{j=1}^{\infty} \mu^*(M_j) \leq \mu^*(E) + \frac{1}{2} \varepsilon.$$

wobei $\mu^*(M_j) = \mu(M_j)$ für alle $j \in \mathbb{N}$ durch das vorherige Theorem sichergestellt ist. Nach den Grenzwertformeln (ausgewertet für das Maß μ^* und die σ -Algebra \mathcal{A}^*) konvergiert $\mu^*(\bigcup_{j=1}^n M_j)$ für $n \rightarrow \infty$ von unten gegen $\mu^*(\bigcup_{j=1}^{\infty} M_j)$ und daher können wir $n \in \mathbb{N}$ so fixieren, dass

$$\mu^*\left(\bigcup_{j=1}^{\infty} M_j\right) \leq \mu^*\left(\bigcup_{j=1}^n M_j\right) + \frac{1}{2} \varepsilon$$

gilt. Wir setzen

$$F := \bigcup_{j=1}^n M_j \in \mathcal{R}, \quad A := \bigcup_{j=1}^{\infty} M_j \in \mathcal{A}^*$$

und zeigen, dass F die gewünschte Ungleichung erfüllt, wobei wir neben den obigen Abschätzungen auch die disjunkte Zerlegungsformel

$$E \Delta F = (E \setminus F) \cup (F \setminus E) = (E \cap F') \cup (F \cap E'),$$

verwenden.

Abschätzung des Maßes der Differenzmenge: Die μ^* -Messbarkeit von $F = A \cap F$ impliziert

$$\mu^*(A) = \mu^*(A \cap F) + \mu^*(A \cap F') = \mu^*(F) + \mu^*(A \cap F')$$

und zusammen mit $E \subseteq A$, der Monotonie von μ^* sowie der Wahl von n ergibt sich hieraus

$$\mu^*(E \cap F') \leq \mu^*(A \cap F') = \mu^*(A) - \mu^*(F) \leq \frac{1}{2} \varepsilon.$$

Andererseits folgt

$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E') = \mu^*(E) + \mu^*(A \cap E')$$

aus der μ^* -Messbarkeit von $E = A \cap E$ und wir erhalten

$$\mu^*(F \cap E') \leq \mu^*(A \cap E') = \mu^*(A) - \mu^*(E) \leq \sum_{j=1}^{\infty} \mu^*(M_j) - \mu^*(E) \leq \frac{1}{2} \varepsilon,$$

wobei wir außerdem $F \subseteq A$, die Monotonie und σ -Additivität von μ^* sowie die Wahl der M_j benutzt haben. Insgesamt ergibt sich

$$\mu^*(E \Delta F) = \mu^*(E \cap F') + \mu^*(F \cap E') \leq \frac{1}{2} \varepsilon + \frac{1}{2} \varepsilon = \varepsilon,$$

denn das vorherige Theorem garantiert die Additivität von μ^* auf \mathcal{A}^* . □

Bemerkungen

1. Die σ -Algebra \mathcal{A}^* enthält in aller Regel deutlich mehr Elemente als der Ring \mathcal{R} , aber der Approximationssatz stellt sicher, dass wir jede der beim Lebesgueschen Abschluss hinzukommenden Mengen (sofern sie endliches Maß besitzen) beliebig genau durch Elemente des Ringes \mathcal{R} approximieren können.
2. Beachte auch, dass die Bedingung $E \in \mathcal{A}^*$ (also die μ^* -Messbarkeit von E) in den Beweis eingeflossen ist und sicherstellt, dass das Maß der Menge $F \setminus E = F \cap E'$ nach oben abgeschätzt werden kann. Insbesondere gibt es keine analoge Aussage für nicht μ^* -messbare Mengen $Y \subseteq X$.

Zusatz*

Theorem (Darstellungssatz für \mathcal{A}^*) Sei μ ein σ -finites Maß. Dann existiert für jedes $E \in \mathcal{A}^*$ ein $F \in \mathcal{A}^\#$ sowie ein $N \in \mathcal{A}^*$ mit

$$E = F \cup N, \quad F \cap N = \emptyset, \quad \mu^*(F) = \mu^*(E), \quad \mu^*(N) = 0,$$

wobei $\mathcal{A}^\# \subseteq \mathcal{A}^*$ die von \mathcal{R} erzeugte σ -Algebra bezeichnet.

Beweis* Spezialfall: Wir betrachten zunächst ein endliches Maß mit $X \in \mathcal{R}$ und $\mu(X) < \infty$. Mit fixiertem $E \in \mathcal{A}^*$ gilt $E' \in \mathcal{A}^*$ sowie $\mu^*(E') \leq \mu^*(X) = \mu(X) < \infty$ und wir können für jedes $n \in \mathbb{N}$ eine Folge $(M_{n,j})_{j=1}^{\infty} \subset \mathcal{R}$ mit

$$E' \subseteq M_n := \bigcup_{j=1}^{\infty} M_{n,j}, \quad \sum_{j=1}^{\infty} \mu^*(M_{n,j}) \leq \mu^*(E') + 1/n$$

wählen, wobei sich

$$M_n \in \mathcal{A}^{\#}, \quad \mu^*(M_n) \leq \mu^*(E') + 1/n$$

unmittelbar aus den Eigenschaften von $\mathcal{A}^{\#}$ und μ^* ergibt. Wir setzen

$$G := \bigcap_{n=1}^{\infty} M_n \in \mathcal{A}^{\#}, \quad N := G \setminus E' \in \mathcal{A}^*$$

und erhalten

$$G = E' \cup N, \quad E' \cap N = \emptyset, \quad \mu^*(G) = \mu^*(E'), \quad \mu^*(N) = 0,$$

denn nach Konstruktion gilt $\mu^*(G) = \mu(E') + \mu^*(N)$ sowie

$$E' \subseteq G \subseteq M_n, \quad \mu^*(E') \leq \mu^*(G) \leq \mu^*(M_n) \leq \mu^*(E') + 1/n$$

für alle $n \in \mathbb{N}$. Die Behauptung ergibt sich schließlich mit

$$F := G'$$

nach elementaren Mengenoperationen sowie einfachen Rechnungen mit Zahlen.

Allgemeiner Fall: Nach Voraussetzung gilt

$$X = \bigcup_{k=1}^{\infty} X_k \quad \text{mit} \quad X_k \in \mathcal{R} \quad \text{und} \quad \mu^*(X_k) < \infty \quad \text{für alle} \quad k \in \mathbb{N}.$$

Wir können alle Argumente von oben für $X_k \setminus E$ anstelle von $E' = X \setminus E$ durchführen. Dies liefert für jedes $k \in \mathbb{N}$ zwei Mengen $G_k \in \mathcal{A}^{\#}$ und $N_k \in \mathcal{A}^*$ mit

$$G_k = (X_k \setminus E) \cup N_k, \quad (X_k \setminus E) \cap N_k = \emptyset, \quad \mu^*(G_k) = \mu^*(X_k \setminus E), \quad \mu^*(N_k) = 0.$$

Diese Formeln gelten analog auch mit

$$\tilde{G}_k := G_k \cap X_k \in \mathcal{A}^{\#}, \quad \tilde{N}_k := N_k \cap X_k \in \mathcal{R}^*$$

anstelle von G_k, N_k und implizieren

$$E \cap X_k = \tilde{F}_k \cup \tilde{N}_k, \quad \tilde{F}_k \cap \tilde{N}_k = \emptyset, \quad \mu^*(\tilde{F}_k) = \mu^*(E \cap X_k), \quad \mu^*(\tilde{N}_k) = 0$$

mit $\tilde{F}_k := X_k \setminus \tilde{G}_k$. Die Behauptung ergibt sich schließlich mit

$$F := \bigcup_{k=1}^{\infty} \tilde{F}_k \in \mathcal{A}^{\#} \quad N := \bigcup_{k=1}^{\infty} \tilde{N}_k \in \mathcal{A}^*$$

aus der σ -Additivität von μ^* auf \mathcal{A}^* .

Bemerkungen

1. Es gilt auch die folgende Umkehrung: Für jedes $F \in \mathcal{A}^\#$ und jedes $N \subseteq X$ mit $\mu^*(N) = 0$ liegt $E = F \cup N$ in \mathcal{A}^* und es gilt $\mu^*(E) = \mu^*(F)$.

Beweis: Die erste Behauptung folgt unmittelbar aus $F \in \mathcal{A}^\# \subseteq \mathcal{A}^*$, $N \in \mathcal{A}^*$ (siehe die Bemerkung weiter oben) und der Tatsache, dass \mathcal{A}^* ein Ring ist. Die zweite ergibt sich via

$$\mu^*(E) \leq \mu^*(F) + \mu^*(N) \leq \mu^*(F) + 0 = \mu^*(F) \leq \mu^*(E)$$

aus der Additivität sowie der Monotonie von μ^* . □

2. Das Theorem und die vorherige Bemerkung zeigen klar den Unterschied zwischen den beiden σ -Algebren \mathcal{A}^* und $\mathcal{A}^\#$. Insbesondere liefert μ^* durch Einschränkung zwar immer auch ein Maß auf $\mathcal{A}^\#$, aber dieses wird im Allgemeinen nicht vollständig sein und nicht alle μ^* -Nullmengen enthalten. Das Mengensystem \mathcal{A}^* ist gerade die kleinste σ -Algebra, die sowohl die Elemente von $\mathcal{A}^\#$ als auch alle μ^* -Nullmengen umfasst.

Theorem (ein Eindeutigkeitsatz für Fortsetzungen) Sei μ ein σ -finites Maß auf dem Ring \mathcal{R} und sei $\mathcal{A}^\#$ die von \mathcal{R} erzeugte σ -Algebra. Sind μ_1, μ_2 zwei Maße auf $\mathcal{A}^\#$ mit

$$\mu_1(M) = \mu_2(M) \quad \text{für alle } M \in \mathcal{R},$$

so gilt $\mu_1 = \mu_2$.

*Beweis** Spezialfall: Wir nehmen $\mu(X) < \infty$ an und betrachten das Mengensystem

$$\mathcal{E} := \{E \in \mathcal{A}^\# : \mu_1(E) = \mu_2(E)\},$$

dass nach Voraussetzung den Ring \mathcal{R} enthält und selbst in $\mathcal{A}^\#$ enthalten ist. Die Konvergenzsätze für auf- bzw. absteigende Folgen von Mengen mit endlichem Maß zeigen, dass \mathcal{E} auch monotonen Mengensystem ist und die äquivalente Charakterisierung von $\mathcal{A}^\#$ liefert $\mathcal{A}^\# \subseteq \mathcal{E}$. Hieraus folgt $\mathcal{A}^\# = \mathcal{E}$ sowie die Behauptung.

Allgemeiner Fall: Wir wählen eine geeignete Ausschöpfung von X , d.h. eine Folge $(X_k)_{k=1}^\infty$ von Mengen aus \mathcal{R} wie im Beweis des Darstellungssatzes. Für alle $k \in \mathbb{N}$ und $m \in \{1, 2\}$ definieren wir via

$$\mu_k(M) := \mu(M \cap X_k), \quad \mu_{k,m}(E) = \mu_m(M \cap X_k),$$

neue Maße, die jeweils finit sind. Der Spezialfall garantiert

$$\mu_{k,1}(E) = \mu_{k,2}(E)$$

für jedes $E \in \mathcal{A}^\#$ und alle $k \in \mathbb{N}$. Die Behauptung ergibt sich im Limes $k \rightarrow \infty$. □

1.4 Das Lebesgue-Maß

Vorbemerkung In diesem Abschnitt führen wir das d -dimensionale Lebesgue-Maß ein, das es uns erlaubt, sehr vielen, (aber nicht allen) Teilmengen des \mathbb{R}^d in sinnvoller Weise ein d -dimensionales Volumen zuzuweisen, wobei wir für $d = 2$ alternativ vom *Flächeninhalt* reden. Punkte im \mathbb{R}^d bezeichnen wir dabei meist mit

$$x = (x_1, \dots, x_d) = (x_l)_{l=1}^d,$$

wobei jede Komponente x_l eine reelle Zahl ist.

Quader im \mathbb{R}^d Eine Menge der Bauart

$$(a, b] = \{x \in \mathbb{R}^d : a_l < x_l \leq b_l \text{ für alle } l = 1, \dots, d\}$$

bezeichnen wir als d -dimensionalen und halboffenen Quader, wobei

$$a = (a_1, \dots, a_d), \quad b = (b_1, \dots, b_d)$$

zwei Punkte in \mathbb{R}^d sind. Analog werden offene und abgeschlossene Quader eingeführt.¹⁷

Bemerkungen

1. Bei halboffenen und offenen Quadern setzen wir immer stillschweigend voraus, dass

$$-\infty < a_l < b_l < +\infty$$

gilt. Bei abgeschlossenen Quadern ist auch immer $a_l = b_l$ zugelassen, wobei dann entartete Quader entstehen.

Beispiel: Die Mengen

$$Q = \{(x_1, x_2) : 0 \leq x_1 \leq 1, x_2 = 1\}, \quad Q = \{(x_1, x_2) : x_1 = 1, x_2 = 1\}$$

sind zweidimensionale abgeschlossene Quader mit Höhe 0, wobei im zweiten Fall auch die Breite verschwindet und Q nur den Punkt $(1, 1)$ enthält.

2. Die Zahlen $b_l - a_l$ sind die Kantenlängen von Q und in natürlicher Weise können die 2^d Eckpunkte eingeführt werden. Im Fall von $b_1 - a_1 = \dots = b_d - a_d$ nennen wir Q auch Würfel.
3. Für $d = 1$ sind Quader Intervalle mit den Randpunkten $a, b \in \mathbb{R}$. Für $d = 2$ ist $a = (a_1, a_2)$ bzw. $b = (b_1, b_2)$ der linke untere bzw. der rechte obere Eckpunkt des Quaders, wobei im halboffenen Fall der Punkt a nicht zum Quader gehört.

¹⁷Es gibt natürlich auch halboffene Quader der Bauart $(a, b]$, aber diese werden wir nur gelegentlich betrachten.

4. Es gilt

$$(a, b] = (a_1, b_1] \times \dots \times (a_d, b_d],$$

d.h. jeder d -dimensionale Quader kann als Kreuzprodukt von d eindimensionalen Quadern interpretiert werden.

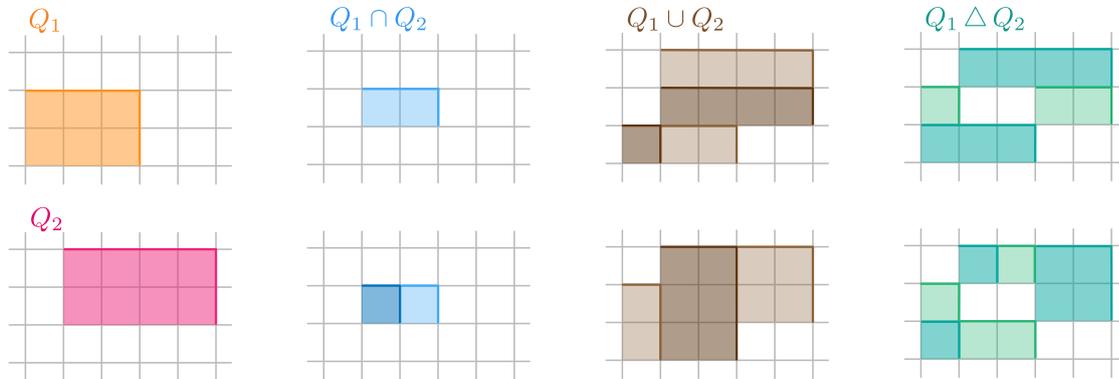


Abbildung Der Durchschnitt, die Vereinigung und die symmetrische Differenz zweier halboffener Quader Q_1 und Q_2 können jeweils als endliche Vereinigung paarweise disjunkter, halboffener Quader geschrieben werden, wobei diese Zerlegung nicht eindeutig ist.

Quaderring Das Mengensystem aller endlichen Vereinigungen halboffener Quader nennen wir \mathcal{Q} . Es handelt sich um einen Ring, der insbesondere die Durchschnitte, Vereinigungen sowie symmetrischen Differenzen halboffener Quader enthält und in dem jedes Element $M \in \mathcal{Q}$ als

$$M = \bigcup_{i=1}^m (a_i, b_i]$$

geschrieben werden kann.¹⁸ Dabei ist die Menge der Teilquader nicht festgelegt und jede Ecke eines Teilquaders ist ein Punkt im \mathbb{R}^d , d.h. es gilt

$$a_i = (a_{i,1}, \dots, a_{i,d}), \quad b_i = (b_{i,1}, \dots, b_{i,d}).$$

Wir können (und werden) dabei immer annehmen, dass die Teilquader von M paarweise disjunkt sind. Siehe dazu das Bild sowie die Übungen.

Lebesgue-Maß auf Quaderring Für jedes $M \in \mathcal{Q}$ wird durch

$$\lambda(M) = \sum_{i=1}^m \lambda((a_i, b_i]) = \sum_{i=1}^m \prod_{l=1}^d (b_{i,l} - a_{i,l})$$

in natürlicher Weise ein Maß auf dem Ring \mathcal{Q} definiert. Insbesondere ist $\lambda(Q)$ für jeden halboffenen Quader Q gerade das Produkt seiner *Kantenlängen*.¹⁹

Definition Das (d -dimensionale) Lebesgue-Maß ist die Lebesguesche Fortsetzung von λ .

¹⁸ \mathcal{Q} ist gerade der Ring, der von allen halboffenen Quadern im \mathbb{R}^d erzeugt wird.

¹⁹Es wird sich zeigen, dass dies auch für alle abgeschlossenen und offenen Quader gilt.

Bemerkungen

1. Wir bezeichnen (wie üblich) die Fortsetzung von λ ebenfalls mit λ (und nicht mit λ^* wie im letzten Abschnitt) und schreiben \mathcal{L} (statt \mathcal{A}^*) für die entsprechende σ -Algebra. Die Elemente von \mathcal{L} werden Lebesgue-messbar bzw. Lebesgue-Mengen genannt.²⁰
2. \mathcal{L} enthält alle Teilmengen $M \subseteq \mathbb{R}^d$, denen man in sinnvoller Weise ein Volumen zuordnen kann, wobei alles auf der sehr einfachen Volumenformel für Quader sowie den abstrakten Konstruktionen aus dem letzten Kapitel beruht. Letztere stellen dabei sicher, dass \mathcal{L} „sehr groß“ ist. Siehe dazu auch die Diskussion zu den alternativen Zugängen zum Volumenproblem.
3. Die Berechnung von $\lambda(M)$ ist für viele M nicht einfach und erfordert nach unserem gegenwärtigen Wissensstand meist nichttriviale Approximationsschritte und Grenzübergänge mit Mengen und den Summen ihrer Maße. Wir werden aber später sehen, dass $\lambda(M)$ in praktisch relevanten Fällen oftmals durch iterierte Integrale berechnet werden kann.

Verschiebungsinvarianz des Lebesgue-Maßes Für jedes $E \in \mathcal{L}$ und alle $\xi \in \mathbb{R}^d$ gilt

$$E + \xi \in \mathcal{L} \quad \text{mit} \quad \lambda(E + \xi) = \lambda(E),$$

wobei $E + \xi = \{x + \xi : x \in E\} \subseteq \mathbb{R}^d$ die Verschiebung der Menge E um den Vektor ξ beschreibt. Für einen Quader $E = Q$ ist die Aussage klar (Q und $Q + \xi$ haben die gleichen Kantenlängen) und für allgemeinere Mengen ergibt sie sich unmittelbar aus den Definitionen zum äußeren Maß und der entsprechenden Messbarkeit aus dem letzten Abschnitt. Insbesondere gilt $\lambda^*(Y) = \lambda^*(Y + \xi)$ für alle $Y \in \mathbb{R}^d$.

Verallgemeinerung: Es wird sich zeigen, dass das Lebesgue-Maß auch invariant unter Spiegelungen und Rotationen ist. Insbesondere werden wir später für eine *lineare* Abbildung $T : \mathbb{R}^d \rightarrow \mathbb{R}^d$ die Formel

$$\lambda(T(E)) = |\det T| \lambda(E)$$

herleiten, aus der sich dann (in mehreren Schritten) die *Transformationsregel für das Lebesgue-Integral* ergibt.

Bemerkung: Man kann zeigen, dass jedes andere verschiebungsinvariante Maß auf \mathbb{R}^d ein skalares Vielfaches vom Lebesgue-Maß sein muss, wobei der universelle Faktor gerade das Maß des Einheitswürfels ist. Wir wollen diese Aussage hier nicht beweisen.

andere Quader Sind $a, b \in \mathbb{R}^d$ die Eckpunkte eines halboffenen Quaders (d.h. $a_l < b_l$ für alle $l = 1, \dots, d$), so gilt

$$\lambda([a, b]) = \lambda((a, b)) = \lambda([a, b)) = \lambda((a, b]) = \prod_{l=1}^d (b_l - a_l),$$

²⁰Viele Autoren statten λ und \mathcal{L} mit dem (meist oberen) Index d aus (Raumdimension), aber wir werden dies nicht tun.

d.h. alle anderen Quader mit denselben Eckpunkten sind auch Lebesgue-messbar und besitzen dasselbe Volumen (nämlich das Produkt der Kantenlängen). Der Beweis dieser sehr intuitiven Aussage wird in den Übungen besprochen. Insbesondere ist jeder entartete Quader (es gilt $a_l = b_l$ für mindestens ein l) und jede Einpunkt-Menge (ein abgeschlossener Quader der Bauart $[a, a]$) eine Lebesguesche Nullmenge, d.h. das entsprechende d -dimensionale Volumen ist 0.

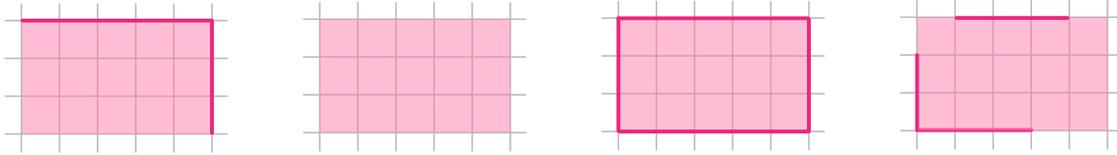


Abbildung Alle Quader sind Lebesgue-messbar und ihr d -dimensionales Volumen ist immer das Produkt ihrer Kantenlängen. Insbesondere ist es nicht wichtig, welche Randpunkte dazu gehören und welche im Komplement liegen. Beachte aber, dass halboffene Quader (ganz links) bessere Eigenschaften unter Mengenoperationen aufweisen und daher in der Theorie eine herausgehobene Rolle spielen.

Bemerkungen

1. Jede abzählbare Teilmenge des \mathbb{R}^d ist eine Lebesguesche Nullmenge.

Beweis: Für jede Menge $N = \{a_1, a_2, \dots\} \subseteq \mathbb{R}^d$ ergibt sich $N \in \mathcal{L}$ sowie

$$\lambda(N) = \sum_{j=1}^{\infty} \lambda([a_j, a_j]) = \sum_{j=1}^{\infty} 0 = 0$$

aus der σ -Additivität von λ . □

2. Ein wichtiges *Standardbeispiel* ist die Dirichlet-Menge

$$D = \{(x_1, \dots, x_d) : 0 \leq x_l \leq 1 \text{ und } x_l \in \mathbb{Q} \text{ für alle } l = 1, \dots, d\},$$

die den abgeschlossenen Einheitswürfel in \mathbb{Q}^d darstellt. Sie ist abzählbar und damit eine Lebesguesche Nullmenge, aber sie ist *nicht* Jordan-messbar (siehe dazu weiter unten). Ihre charakteristische Funktion wird daher Lebesgue-, aber nicht Riemann-integrierbar sein.

3. Es gibt aber auch überabzählbare Nullmengen, zum Beispiel die weiter unten diskutierte *Cantor-Menge*.

Lebesgue-Stieltjes Maße Wenn $w : \mathbb{R}^d \rightarrow \mathbb{R}$ eine nichtnegative und hinreichend gute Funktion ist, so wird auch durch

$$\lambda_w(Q) := \int_Q w(x) dx$$

ein Maß auf dem Quadering \mathcal{Q} erzeugt, wobei wir die Integration von Funktionen über Teilmengen des \mathbb{R}^d erst weiter unten einführen werden und das Lebesgue-Maß gerade der konstanten Funktion $w \equiv 1$ entspricht. Die Lebesguesche Fortsetzung (im Sinne des vorherigen Abschnitts) wird das *Lebesgue-Stieltjes Maß mit Dichte w* genannt.

Solche Maße spielen zum Beispiel in der Stochastik eine wichtige Rolle, wobei dann zusätzlich $\int_{\mathbb{R}^d} w(x) dx = 1$ gefordert und w als Wahrscheinlichkeitsdichte einer vektoriellen Zufallsvariablen Ξ interpretiert wird. Insbesondere quantifiziert die reelle Zahl $\lambda_w(E)$ die Wahrscheinlichkeit, dass Ξ Werte in der Menge E annimmt, und ist daher *nicht* verschiebungsinvariant.

Borelsche und Lebesgue Mengen

Definition Die von allen offenen Teilmengen des \mathbb{R}^d erzeugte σ -Algebra wird Borelsche σ -Algebra (oder einfach Borel-Algebra) genannt und mit \mathcal{B} bezeichnet.

Lemma (Zerlegung und Messbarkeit offener Mengen) Jede offene Menge $O \subseteq \mathbb{R}^d$ kann als abzählbare Vereinigungen von paarweise disjunkten, halboffenen Quadern dargestellt werden. Insbesondere gilt $O \in \mathcal{L}$.

Beweis *Schritt 1:* Für jedes $n \in \mathbb{N}$ betrachten wir das Mengensystem

$$\mathcal{W}_n := \left\{ (2^{-n} j_1 - 2^{-n}, 2^{-n} j_1] \times \dots \times (2^{-n} j_d - 2^{-n}, 2^{-n} j_d] : j_1, \dots, j_d \in \mathbb{Z} \right\},$$

das aus allen d -dimensionalen Würfeln der Kantenlänge 2^{-n} besteht, deren Eckpunkte alle nur ganzzahlige Vielfache von 2^{-n} als Komponenten besitzen. Die Elemente von \mathcal{W}_n sind dabei paarweise disjunkt und schöpfen insgesamt den ganzen Raum aus²¹, d.h. es gilt

$$\mathbb{R}^d = \bigcup \mathcal{W}_n$$

für jedes $n \in \mathbb{N}$, wobei rechts die Vereinigung aller Elemente von \mathcal{W}_n steht. Wir betrachten außerdem

$$O_n = \bigcup \{W \in \mathcal{W}_n : W \subseteq O\},$$

d.h. die Vereinigung aller Würfel aus \mathcal{W}_n , die ganz in O liegen. Nach Konstruktion gilt

$$O_1 \subseteq O_2 \subseteq \dots \subseteq O$$

und jedes O_n ist als Vereinigung von höchstens abzählbar vielen halboffenen Würfeln in der σ -Algebra \mathcal{L} enthalten.

Schritt 2: Wir wollen zeigen, dass die Vereinigung aller O_n gerade O ist und verweisen auf das nachfolgende Bild für eine Illustration der Beweisidee. Wir fixieren dazu $x \in O$ und wählen $\varepsilon > 0$ mit $B_\varepsilon(x) \subseteq O$ sowie für jedes $n \in \mathbb{N}$ den eindeutig bestimmten Würfel $W_n \in \mathcal{W}_n$ mit $x \in W_n$. Für jedes andere $\tilde{x} \in W_n$ ergibt sich die Abschätzung

$$|x - \tilde{x}|^2 = \sum_{l=1}^d |x_l - \tilde{x}_l|^2 \leq \sum_{l=1}^d 2^{-2n} = 2^{-2n} d$$

und wir schließen, dass für

$$n > \frac{\ln d - 2 \ln \varepsilon}{2 \ln 2}$$

der Würfel W_n ganz in $B_\varepsilon(x)$ und damit auch ganz in O liegt.²² Insbesondere gilt dann $W_n \subseteq O_n \subseteq O$ und daher auch $x \in O_n$. Wir haben damit

$$O = O_1 \cup O_2 \cup O_3 \cup \dots$$

bewiesen und schließen, dass O als abzählbare Vereinigung von Mengen aus \mathcal{L} ebenfalls zur σ -Algebra \mathcal{L} gehört. \square

²¹Die Würfel aus \mathcal{W}_n „parkettieren“ insgesamt den \mathbb{R}^d , wobei die Knotenpunkte auf dem Gitter $2^{-n} \mathbb{Z}^d$ liegen.

²²Beachte, dass unser Argument bei einer nicht-offenen Menge versagt. Bei einem Randpunkt x von U kann es nämlich passieren, dass *alle* entsprechenden Quader W_n jeweils mindestens einen Punkt aus O' enthalten und damit *nicht* vollständig in O liegen.

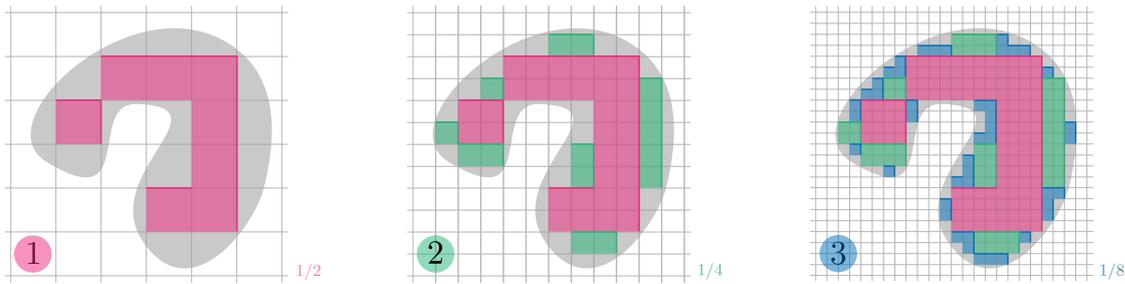


Abbildung Anschauliche Erklärung, dass jede offene Teilmenge O (grau) des \mathbb{R}^d (hier $d = 2$) als Vereinigung von höchstens abzählbar vielen halboffenen Quadern dargestellt werden kann und daher Lebesgue-messbar ist. Im ersten Schritt wählen wir alle Würfel aus dem Mengensystem \mathcal{W}_1 (rot), die ganz in O liegen. Im zweiten bzw. dritten Schritt fügen wir weitere Würfel aus \mathcal{W}_2 (grün) bzw. \mathcal{W}_3 (blau) hinzu und setzen anschließend diese Konstruktion iterativ fort, wobei bei einer beschränkten Menge O im n -ten Schritt immer nur endlich viele Würfel der Kantenlänge 2^{-n} hinzukommen. Insgesamt können wir ganz O von innen ausschöpfen, wobei es wichtig ist, dass O mit jedem Punkt auch eine kleine Kugel um diesen Punkt und damit für große n auch immer einen Würfel aus \mathcal{W}_n vollständig enthält. Der formale Beweis des Lemmas ist übrigens leicht anders aufgeschrieben und benutzt indirekt, dass jeder Würfel in \mathcal{W}_n als disjunkte Vereinigung von genau 2^d Würfeln aus \mathcal{W}_{n+1} betrachtet werden kann. Im Bild meint dies, dass jedes grüne Quadrat aus vier blauen Teilquadraten besteht und dass jedes rote Quadrat die Vereinigung von 4 grünen bzw. 16 blauen Quadraten ist.

Korollar (Borel-Mengen sind Lebesgue-messbar) Es gilt $\mathcal{B} \subseteq \mathcal{L}$.

Beweis Das Lemma impliziert, dass die σ -Algebra \mathcal{L} alle offenen Mengen enthält. Die Behauptung ergibt sich damit unmittelbar aus der Minimalität von \mathcal{B} . \square

Bemerkungen

1. Die Borelsche Algebra enthält auch alle abgeschlossenen Mengen, da diese (per Definition) gerade die Komplemente von offenen Mengen sind. Das Lemma und das Korollar zeigen also, dass *jede offene* sowie *jede abgeschlossene* Teilmenge des \mathbb{R}^n Lebesgue-messbar ist. Das ist auf den ersten Blick sehr überraschend, denn es gibt Borel-Mengen, die sehr wild aussehen (zum Beispiel die fraktalartige Cantor-Menge). Auf den zweiten Blick wird aber klar, dass diese wichtige Eigenschaft des Lebesgue-Maßes die Mächtigkeit der abstrakten Konstruktionen aus dem vorherigen Abschnitt widerspiegelt.
2. Für viele Borel-Mengen (etwa andere Quader, Kugeln, Pyramiden, usw.) kann die Lebesgue-Messbarkeit auch direkt nachgeprüft werden. Siehe dazu auch die Übungen.
3. Es gilt $\mathcal{B} \neq \mathcal{L}$, d.h. es existieren Teilmengen des \mathbb{R}^d , die zwar zur Lebesgue-, aber nicht zur Borel-Algebra gehören. Außerdem gilt $\mathcal{L} \neq \mathcal{P}(\mathbb{R}^d)$, d.h. es gibt Mengen, die nicht Lebesgue-messbar sind und denen man kein Volumen zuordnen kann (jedenfalls nicht in sinnvoller Weise). Es ist allerdings nicht einfach, jeweils ein konkretes Beispiel anzugeben. Siehe dazu weiter unten.

Zusatz*: Man kann sogar zeigen (vgl. etwa die Diskussion und die Literaturverweise in [Els, Abschnitt II.8]), dass die Mächtigkeit von \mathcal{B} bzw. \mathcal{L} die von \mathbb{R} bzw. $\mathcal{P}(\mathbb{R})$ ist. In diesem Sinne gibt es deutlich mehr Lebesguesche als Borelsche Mengen.

4. Die Borel-Algebra wird auch von allen halboffenen Quadern erzeugt.

Beweis: Sei \mathcal{O} das Mengensystem aller offenen Teilmengen des \mathbb{R}^d und \mathcal{A} die von \mathcal{Q} erzeugte σ -Algebra. Im Beweis des Theorems hatten wir gezeigt, dass jedes Element von \mathcal{O} als unendliche Vereinigung von Elementen aus \mathcal{Q} dargestellt werden kann und wir schließen, dass jede σ -Algebra, die \mathcal{Q} enthält, auch \mathcal{O} enthalten wird. Insbesondere gilt dies für \mathcal{A} und die Minimalität von \mathcal{B} garantiert $\mathcal{B} \subseteq \mathcal{A}$. Andererseits kann jeder halboffene Quader als abzählbarer Durchschnitt einer absteigenden Folge von offenen Quadern — oder alternativ als Vereinigung einer aufsteigenden Folge abgeschlossener Quader — dargestellt werden (Übungsaufgabe) und wir schließen, dass auch $\mathcal{A} \subseteq \mathcal{B}$ gilt. \square

5. *Zusatz**: Die letzte Bemerkung und der Darstellungssatz aus dem vorherigen Abschnitt stellen sicher, dass für jede Lebesgue-Menge $E \in \mathcal{L}$ eine Borel-Menge $F \in \mathcal{B}$ sowie eine Lebesguesche Nullmenge $N \in \mathcal{L}$ existiert, sodass

$$E = F \cup N, \quad F \cap N = \emptyset, \quad \lambda(F) = \lambda(E), \quad \lambda(N) = 0$$

erfüllt ist.

Lemma (Approximationssatz für Lebesgue-Mengen) Für jedes $E \subseteq \mathcal{L}$ und jedes $\varepsilon > 0$ existiert eine abgeschlossene Menge A sowie eine offene Menge O , sodass

$$A \subseteq E \subseteq O, \quad \lambda(O \setminus A) \leq \varepsilon$$

gilt.

Beweis offene Menge im ersten Spezialfall: Als Hilfsresultat zeigen wir, dass unter der Zusatzannahme $\lambda(E) < \infty$ immer eine offene Menge O mit

$$E \subseteq O, \quad \lambda(O \setminus E) \leq \frac{1}{2} \varepsilon$$

existiert. Nach Definition des äußeren Maßes existiert eine Folge $(M_j)_{j=1}^{\infty}$ im Ring \mathcal{Q} , sodass

$$E \subseteq \bigcup_{j=1}^{\infty} M_j, \quad \sum_{j=1}^{\infty} \lambda(M_j) \leq \lambda(E) + \frac{1}{4} \varepsilon,$$

wobei jedes M_j via

$$M_j = \bigcup_{i=1}^{m_j} Q_{j,i}, \quad \lambda(M_j) = \sum_{i=1}^{m_j} \lambda(Q_{j,i})$$

als endliche und paarweise disjunkte Vereinigung von halboffenen Quadern dargestellt werden kann. Wir können wir außerdem immer einen offenen Quader $O_{j,i}$ mit

$$Q_{j,i} \subseteq O_{j,i}, \quad \lambda(O_{j,i}) \leq \lambda(Q_{j,i}) + \frac{\varepsilon}{2^{j+2} m_j}$$

wählen. Dann ist

$$O := \bigcup_{j=1}^{\infty} \bigcup_{i=1}^{m_j} O_{j,i}$$

eine offene Menge²³ mit $E \subseteq O$ und ihr Maß genügt der Abschätzung

$$\lambda(O) \leq \sum_{j=1}^{\infty} \sum_{i=1}^{m_j} \left(\lambda(Q_{j,i}) + \frac{\varepsilon}{2^{j+2} m_j} \right) \leq \sum_{j=1}^{\infty} \left(\lambda(M_j) + \frac{\varepsilon}{2^{j+2}} \right) \leq \lambda(E) + \frac{1}{4} \varepsilon + \frac{1}{4} \varepsilon,$$

wobei wir $\sum_{j=1}^{\infty} 2^{-(j+2)} = \frac{1}{4}$ verwendet haben. Die gewünschte Ungleichung ergibt sich nun via $\lambda(O \setminus E) = \lambda(O) - \lambda(E)$.

offene Menge im zweiten Spezialfall: Wir beweisen, dass das Hilfsresultat auch im Fall $\lambda(E) = \infty$ gilt. Für jedes $n \in \mathbb{N}$ setzen wir dazu

$$E_n := E \cap (W_{n+1} \setminus W_n), \quad W_n := (-n, +n] \times \dots \times (-n, +n]$$

und bemerken, dass $\lambda(E_n) \leq \lambda(W_{n+1} \setminus W_n) \leq \lambda(W_{n+1}) - \lambda(W_n) < \infty$ gilt. Damit können wir das schon bewiesene Hilfsresultat mit E_n und $\varepsilon/2^n$ anstelle von E und ε auswerten und erhalten für jedes $n \in \mathbb{N}$ eine offene Menge O_n , sodass

$$E_n \subseteq O_n, \quad \lambda(O_n \setminus E_n) \leq \frac{\varepsilon}{2^{n+1}}.$$

Mit $O := \bigcup_{n=1}^{\infty} O_n$ und wegen $E = \bigcup_{n=1}^{\infty} E_n$ ergibt sich

$$\lambda(O \setminus E) = \lambda\left(\bigcup_{n=1}^{\infty} O_n \setminus E_n\right) \leq \sum_{n=1}^{\infty} \frac{\varepsilon}{2^{n+1}} = \frac{1}{2} \varepsilon$$

sowie $E \subseteq O$.

abgeschlossene Menge: Wir wenden das Hilfsresultat auf die Komplementmenge $\tilde{E} := E'$ an und erhalten eine offene Menge \tilde{O} mit $\tilde{E} \subseteq \tilde{O}$ und $\lambda(\tilde{O} \setminus \tilde{E}) \leq \frac{1}{2} \varepsilon$. Mit $A := \tilde{O}'$ folgt

$$A \subseteq E, \quad \lambda(E \setminus A) \leq \frac{1}{2} \varepsilon$$

aus $E \setminus A = \tilde{O} \setminus \tilde{E}$ und wegen

$$\lambda(O \setminus A) = \lambda(O \setminus E) + \lambda(E \setminus A)$$

haben wir insgesamt die Behauptung hergeleitet. □

Vorlesung 03-2 : 11. November

mehr über Lebesgue-Mengen

Korollar (Charakterisierung Lebesguescher Nullmengen) Eine Teilmenge $E \subseteq \mathbb{R}^d$ ist genau dann eine Lebesguesche Nullmenge, wenn für jedes $\varepsilon > 0$ eine Folge $(Q_j)_{j=1}^{\infty}$ von halboffenen Quadern mit

$$E \subseteq \bigcup_{j=1}^{\infty} Q_j, \quad \sum_{j=1}^{\infty} \lambda(Q_j) \leq \varepsilon$$

existiert.

²³Wir hatten in *Analysis 2* gezeigt, dass beliebige Vereinigungen offener Mengen immer offen sind.

Beweis Hinrichtung: Seien $E \in \mathcal{L}$ mit $\lambda(E) = 0$ sowie $\varepsilon > 0$ gegeben und seien A, O wie im vorherigen Lemma. Dann folgt

$$\lambda(O) = \lambda(A) + \lambda(O \setminus A) \leq \lambda(E) + \lambda(O \setminus A) \leq 0 + \varepsilon = \varepsilon,$$

aus der Monotonie und der Additivität von λ . Das Zerlegungslemma angewendet auf O liefert außerdem

$$O = \bigcup_{j=1}^{\infty} Q_j, \quad \lambda(O) = \sum_{j=1}^{\infty} \lambda(Q_j)$$

für eine geeignet gewählte Folge von Quadern und die Behauptung folgt unmittelbar.

Rückrichtung: Sei E eine beliebige Teilmenge von \mathbb{R}^d mit der ε -Eigenschaft. Dann existiert für jedes $n \in \mathbb{N}$ eine Menge $M_n \in \mathcal{Q}$ mit

$$E \subseteq M_n, \quad \lambda(M_n) \leq 1/n,$$

denn wir können die Voraussetzung mit $\varepsilon = 1/n$ auswerten und jedes M_n sogar als Vereinigung von abzählbar vielen und paarweise disjunkten Quadern wählen. Insbesondere ist $M := \bigcup_{n=1}^{\infty} M_n$ wegen $\lambda(M) = \lim_{n \rightarrow \infty} \lambda(M_n) = 0$ eine Nullmenge mit $E \subseteq M$. Die Vollständigkeit und Monotonie von λ garantieren $E \in \mathcal{L}$ sowie $\lambda(E) = 0$. \square

Bemerkungen

1. Im Approximationssatz ist es wichtig, dass die *kleinere* bzw. die *größere* Menge (also A bzw. O) *abgeschlossen* bzw. *offen* ist. Eine *innere* bzw. *äußere* Approximation mit offenen bzw. abgeschlossenen Mengen ist im Allgemeinen nämlich nicht möglich.

Gegenbeispiel: Die Dirichlet-Menge D (siehe oben) besitzt keine inneren Punkte, aber liegt dicht im abgeschlossenen Einheitswürfel. Insbesondere gilt

$$O = \emptyset, \quad A \supseteq D, \quad \lambda(A \setminus O) = \lambda(A) \geq 1$$

für jedes abgeschlossene A und jedes offene O mit $O \subseteq D \subseteq A$.

2. Verschärfung: Das Korollar gilt sinngemäß auch mit offenen oder abgeschlossenen anstelle von halboffenen Quadern (Übungsaufgabe). Insbesondere folgt aus der Variante mit offenen Quadern und dem Satz von Heine-Borel aus *Analysis 1*, dass jede *kompakte* Nullmenge auch schon durch *endlich viele* Quader überdeckt werden kann.

Cantor-Mengen Für einen gegebenen Parameter $\delta \in (0, 1)$ definieren wir — siehe auch die Abbildung — durch die folgende Rekursionsvorschrift eine Folge $(C_{\delta, n})_{n=0}^{\infty}$ von Teilmengen des Einheitsintervalles:

Anfang: Es gilt $C_{\delta, 0} = [0, 1]$.

n -ter Schritt: Aus jedem Teilintervall von $C_{\delta, n-1}$ wird das mittlere Stück mit relativer Länge δ entfernt, wodurch zwei abgeschlossene Teilintervalle von $C_{\delta, n}$ entstehen.

Durch Induktion können wir leicht zeigen, dass jede Menge $C_{\delta,n}$ aus 2^n paarweise disjunkten Intervallen der Länge

$$l_{\delta,n} = \left(\frac{1-\delta}{2}\right)^n$$

besteht, wobei diese für $n \rightarrow \infty$ gegen 0 konvergiert. Die eigentliche Cantor- δ -Menge ist durch

$$C_\delta := \bigcap_{n=0}^{\infty} C_{\delta,n}$$

gegeben und besitzt sehr bemerkenswerte Eigenschaften.

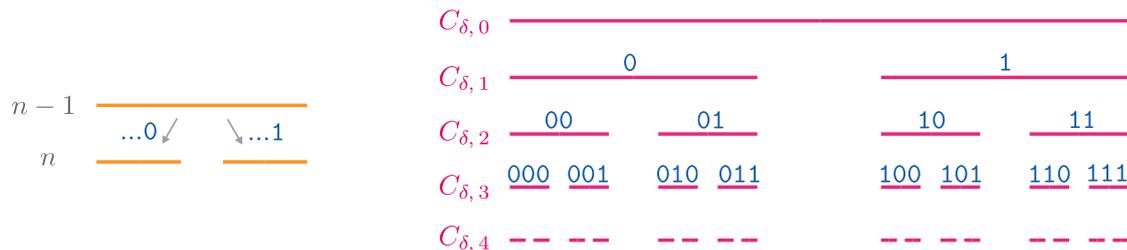


Abbildung Links: Der Rekursionvorschrift bei der Konstruktion der Cantor- δ -Menge C_δ (dargestellt für $\delta = 1/5$). Die farbigen Intervalle sind dabei alle abgeschlossen. Rechts: Die ersten fünf Mengen $C_{n,\delta}$. Die blauen Zahlenkolonnen illustrieren, dass jeder Punkt von $C_\delta = \bigcap_{n=0}^{\infty} C_{\delta,n}$ in eindeutiger Weise einer unendlichen Folge von Nullen und Einsen entspricht, wobei die n -te Ziffer gerade angibt, ob wir uns im n -ten Schritt in der Intervall-Hierarchie nach links oder rechts bewegen müssen, um diesen Punkt zu erreichen. Das Intervall $C_{\delta,0}$ trägt dabei keine Nummer und die Zahlenkolonnen für $C_{\delta,4}$ enthalten jeweils 4 Ziffern, die aber aus Platzgründen nicht angegeben werden konnten.

Lemma (Eigenschaften der Cantor- δ -Menge) Für jeden zugelassenen Wert von δ gelten die folgenden Aussagen:

1. C_δ ist kompakt.
2. C_δ ist Lebeguesche Nullmenge.
3. C_δ enthält keine inneren Punkte.
4. C_δ ist überabzählbar.

Beweis Teil 1: Jedes $C_{\delta,n}$ ist als endliche Vereinigung abgeschlossener Intervalle selbst abgeschlossen und deshalb ist auch der unendliche Durchschnitt C_δ abgeschlossen (siehe *Analysis 2*). Da C_δ außerdem als Teilmenge von $[0, 1]$ beschränkt ist, folgt die Kompaktheit aus dem Satz von Heine-Borel (wieder *Analysis 2*).

Teil 2: Der erste Teil impliziert, dass sowohl C_δ als auch alle Mengen $C_{\delta,n}$ Lebesguemessbar sind und mit Induktion zeigen wir

$$\lambda(C_{\delta,n}) = 2^n l_n = (1-\delta)^n.$$

Die Grenzwertsätze für absteigende Folgen liefern daher $\lambda(C_\delta) = \lim_{n \rightarrow \infty} \lambda(C_{\delta,n}) = 0$.

Teil 3: Wir nehmen als Antithese an, es gäbe ein abgeschlossenes Intervall $[\alpha, \beta]$ in C_δ und fixieren $n \in \mathbb{N}$ hinreichend groß mit $l_n < \beta - \alpha$. Nach Annahme gilt $\alpha, \beta \in C_{\delta,n}$, aber nach der Wahl von n können beide Punkte nur in verschiedenen Teilintervallen von $C_{\delta,n}$ liegen. Dann existiert aber mindestens ein Punkt γ zwischen α und β , der

nicht zu $C_{\delta,n}$ und damit auch nicht zu C_δ gehört. Das ist ein Widerspruch und wir haben gezeigt, dass die Menge C_δ kein abgeschlossenes Intervall enthält. Dann kann sie aber keine eindimensionale Kugel um einen Punkt enthalten, denn eine solche Kugel ist ein offenes Intervall und enthält immer auch ein abgeschlossenes Intervall.

Teil 4: Durch Einführung einer geschickten Indizierung aller Teilintervalle (siehe nochmal das Bild) sowie unter Ausnutzung des Intervallschachtelungsprinzips können wir zeigen, dass jedem Punkt $x \in \Omega_\delta$ in eindeutiger und umkehrbarer Weise eine Folge aus Nullen und Einsen zugeordnet werden kann. Insbesondere besitzen Ω_δ und der Folgenraum

$$\Xi := \left\{ (\xi_j)_{j=1}^\infty : \xi_j \in \{0, 1\} \text{ für alle } j \in \mathbb{N} \right\}$$

dieselbe Mächtigkeit, aber letzterer ist bekanntlich überabzählbar.²⁴ \square

Bemerkung Die Cantor-Mengen zeigen sind Beispiele für eindimensionale Lebesgue-Nullmengen, die „sehr viele“ Punkte enthalten. Analoge Konstruktionen können auch in höheren Dimensionen durchgeführt werden.

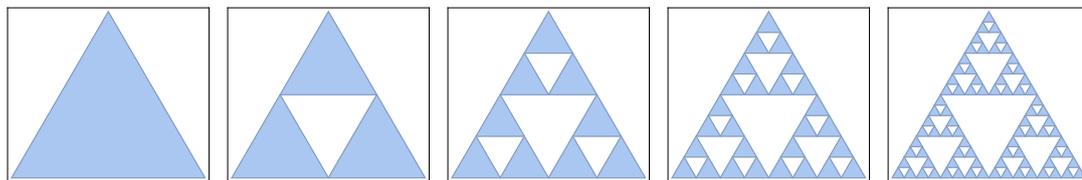


Abbildung Das *Sierpiński-Dreieck* wird ebenfalls rekursiv definiert und ist eine zweidimensionale Lebesguesche Nullmenge.

Theorem (Existenz nicht-messbarer Mengen) \mathcal{L} ist eine echte Teilmenge der Potenzmenge $\mathcal{P}(\mathbb{R}^d)$.

Beweis* *Auswahl einer speziellen Menge:* Wir nennen zwei Punkte $x, \tilde{x} \in \mathbb{R}^d$ äquivalent und schreiben $x \sim \tilde{x}$, falls alle Komponenten von $x - \tilde{x}$ rational sind, d.h. falls $x - \tilde{x} \in \mathbb{Q}^d$ gilt. Dies definiert eine Äquivalenzrelation auf \mathbb{R}^d , zu der es unendlich viele, paarweise disjunkte Äquivalenzklassen gibt, wobei jede auch Repräsentanten im halboffenen und d -dimensionalen Einheitswürfel

$$W := (0, 1] \times \dots \times (0, 1] = (0, 1]^d$$

enthält.²⁵ Nach dem Auswahlaxiom können wir nun aus jeder Äquivalenzklasse einen Repräsentanten aus W wählen und erhalten so eine Teilmenge $Y \subseteq W$, wobei $y_1 \approx y_2$

²⁴Wir können die Überabzählbarkeit von Ξ auch leicht mit einem Cantorschen Diagonalargument nachweisen. Beachte auch, dass Ξ in natürlicher Weise mit der Potenzmenge $\mathcal{P}(\mathbb{N})$ identifiziert werden kann.

²⁵Die Äquivalenzklassen — siehe zum Beispiel die Vorlesung *Lineare Algebra* — sind die Mengen der Bauart

$$[x] = \{ \tilde{x} : \tilde{x} \sim x \}.$$

In unserem Fall enthält jede dieser Mengen abzählbar unendlich viele Elemente und daher muss es insgesamt überabzählbar unendlich viele von ihnen geben. Für jedes $x \in \mathbb{R}^d$ enthält $[x]$ mindestens einen Repräsentanten $\tilde{x} \in [x] \cap W$. Zum Beispiel $\tilde{x} = x - \xi$ mit $\xi = (\xi_1, \dots, \xi_d) \in \mathbb{Z}^d$, sofern ξ_l für jedes $l = 1 \dots d$ die größte ganze Zahl ist, für die $\xi_l < x_l$ gilt. Alternativ und mithilfe der Gaußschen Klammer des Aufrundens kann dies als $\xi_l = \lceil x_l - 1 \rceil$ bzw. $\tilde{x}_l = x_l - \lceil x_l - 1 \rceil$ formuliert werden.

für je zwei verschiedene Punkte $y_1, y_2 \in Y$ gilt. Wir wollen durch einen indirekten Beweis zeigen, dass Y nicht Lebesgue-messbar ist, und nehmen daher $Y \in \mathcal{L}$ als Antithese an.

Vorbereitungen für das Widerspruchsargument: Die Menge $\tilde{W} \cap \mathbb{Q}^d$ mit

$$\tilde{W} := (-1, +1] \times \dots \times (-1, +1] = (-1, +1]^d$$

ist abzählbar unendlich und wir fixieren irgendeine abzählende Folge $(\xi_n)_{n=1}^\infty$. Wir betrachten außerdem die Mengen

$$Y_n := \xi_n + Y = \{\xi + y : y \in Y\}$$

und bemerken, dass diese die folgenden Eigenschaften besitzen:

(1) Für alle $n_1 \neq n_2$ gilt

$$Y_{n_1} \cap Y_{n_2} = \emptyset,$$

denn andernfalls gäbe es ein $x \in \mathbb{R}^d$ sowie zwei verschiedene Punkte $y_1, y_2 \in Y$ mit $x = \xi_{n_1} + y_1 = \xi_{n_2} + y_2$ und es würde $y_1 - y_2 \in \mathbb{Q}^d$ und damit $y_1 \sim y_2$ gelten, was aber der Definition von Y widerspricht.

(2) Unsere Definition impliziert

$$\bigcup_{n=1}^{\infty} Y_n \subseteq \check{W}$$

mit $\check{W} := (-1, 2]^d$, denn es gilt offensichtlich $Y_n \subseteq \check{W}$ für alle $n \in \mathbb{N}$.

(3) Es gilt

$$W \subseteq \bigcup_{n=1}^{\infty} Y_n.$$

In der Tat, für jedes $x \in W$ existiert ein $y \in Y$ mit $x \sim y$ bzw. $x - y \in \mathbb{Q}^d$. Wegen $x - y \in \check{W}$ können wir außerdem $n \in \mathbb{N}$ wählen, sodass $x - y = \xi_n$ und damit $x \in Y_n$ gilt.

(4) Für jedes $n \in \mathbb{N}$ folgt

$$Y_n \in \mathcal{L}, \quad \lambda(Y_n) = \lambda(Y)$$

aus unserer Antithese sowie der Verschiebungsinvarianz von λ .

Herleitung des Widerspruchs: Im Fall von $\lambda(Y) = 0$ können wir via

$$1 = \lambda(E) \leq \sum_{n=1}^{\infty} \lambda(Y) = \sum_{n=1}^{\infty} 0 = 0$$

einen Widerspruch aus (1), (3) und (4) ableiten, wohingegen sich für $\lambda(Y) > 0$ der Widerspruch

$$\infty = \sum_{n=1}^{\infty} \lambda(Y) \leq \lambda(2E) = 2^d$$

aus (1), (2) und (4) ergibt. Wir haben damit gezeigt, dass es keinen möglichen Wert für $\lambda(Y)$ gibt, und schließen, dass Y nicht Lebesgue-messbar sein kann. \square

Diskussion

1. Unser Beweis ist wahrscheinlich der erste in Ihrem Studium, der sich explizit auf das *Auswahlaxiom der Mathematik* bezieht. Es besagt, dass man für ein gegebenes System von Mengen (bei uns die überabzählbar vielen Äquivalenzklassen) aus jeder Menge im System genau ein Element auswählen kann. Besteht das Mengensystem aus nur endlich vielen Mengen, so ergibt sich die Gültigkeit dieser Aussage aus elementarerer Prinzipien, aber für unendliche Mengensysteme muss sie durch Postulat sichergestellt werden. Für eine weiterführende Diskussion verweisen wir auf die Literatur zur mathematischen Logik und zur theoretischen Mengenlehre (die allerdings nicht einfach zu verstehen ist).
2. Es stellt sich heraus, dass ohne das Auswahlaxiom die Existenz von nicht Lebesgue-messbaren Mengen nicht bewiesen werden kann. Insbesondere gibt es keine *konstruktive* oder *algorithmische* Möglichkeit, eine solche anzugeben.
3. Äquivalent zum Auswahlaxiom ist übrigens das *Lemma von Zorn*. Auch dieses werden Sie in ihrem Studium hin und wieder verwenden.
4. Durch eine geringfügige Modifikation der Konstruktionen und Argumente aus dem Beweis des Theorems kann man zeigen, dass jede Lebesgue-Menge mit positivem Maß nicht-messbare Teilmengen besitzt. Beachte aber, dass jede Teilmenge einer Lebesgueschen Nullmenge messbar ist (Vollständigkeit des Lebesgue Maßes).

Zusatz*

Lebesgue-Maß und Jordan-Inhalt

Vorbemerkung In diesem Abschnitt entwickeln wir eine alternative Sicht auf das Lebesgue-Maß und diskutieren den Zusammenhang mit einem anderen Volumenbegriff.

äußeres und inneres Lebesgue-Maß Durch

$$\begin{aligned}\lambda_{\#}(Y) &:= \sup \{ \lambda(K) : K \subseteq \mathbb{R}^d \text{ kompakt mit } K \subseteq Y \}, \\ \lambda^{\#}(Y) &:= \inf \{ \lambda(O) : O \subseteq \mathbb{R}^d \text{ offen mit } Y \subseteq O \}\end{aligned}$$

werden zwei Mengenfunktionen $\lambda_{\#}, \lambda^{\#} : \mathcal{P}(\mathbb{R}^d) \rightarrow [0, \infty]$ definiert, die man inneres und äußeres Lebesgue-Maß von $Y \subset X$ nennt.

Bemerkungen

1. Die Definition von $\lambda^{\#}$ ist leicht anders als die von μ^* im letzten Abschnitt, aber aufgrund des Zerlegungslemmas konzeptionell sehr verwandt. Insbesondere gilt $\lambda^{\#}(Y) = \lambda^*(Y)$ für alle $Y \subseteq \mathbb{R}^d$.
2. Es gilt $\lambda_{\#}(Y) \leq \lambda^{\#}(Y)$ für alle $Y \subseteq \mathbb{R}^d$.

Beweis: Für alle kompakten Mengen K und jede offene Menge O mit $K \subseteq Y \subseteq O$ gilt $\lambda(K) \leq \lambda(O)$. Durch Supremumbildung über K erhalten wir $\lambda_{\#}(Y) \leq \lambda(O)$ und die Behauptung folgt nach Infimumsbildung bzgl. O . \square

3. Beide Mengenfunktionen sind monoton, denn für $Y_1 \subseteq Y_2$ folgt $\lambda_{\#}(Y_1) \leq \lambda_{\#}(Y_2)$ bzw. $\lambda^{\#}(Y_1) \leq \lambda^{\#}(Y_2)$ aus den Eigenschaften des Supremums bzw. Infimums.
4. Die Mengenfunktion $\lambda^{\#}$ ist σ -subadditiv auf $\mathcal{P}(\mathbb{R}^d)$, aber $\lambda_{\#}$ ist σ -superadditiv.
5. Es gilt

$$\lambda_{\#}(Y) = \sup \{ \lambda(A) : A \subseteq \mathbb{R}^d \text{ abgeschlossen mit } A \subseteq Y \},$$

d.h. bei der Berechnung des inneren Maßes können wir auch mit abgeschlossenen anstelle von kompakten Mengen arbeiten.

Lemma (äquivalente Charakterisierung von Lebesgue-Messbarkeit) Eine Teilmenge $E \subseteq \mathbb{R}^d$ gehört genau dann zu \mathcal{L} , wenn $\lambda_{\#}(E)$ und $\lambda^{\#}(E)$ gleich sind, wobei dann sogar $\lambda_{\#}(E) = \lambda(E) = \lambda^{\#}(E)$ gilt.

Beweis* *Hinrichtung*: Mit fixiertem $E \in \mathcal{L}$ betrachten wir die beschränkten sowie Lebesgue-messbaren Mengen

$$E_n := E \cap W_n, \quad W_n := (-n, n] \times \dots \times (-n, n],$$

für die $\lim_{n \rightarrow \infty} \lambda(E_n) = \lambda(E)$ im Sinne der eigentlichen (für $\lambda(E) < \infty$) oder der uneigentlichen (für $\lambda(E) = \infty$) Konvergenz gilt.²⁶ Nach dem Approximationssatz existieren abgeschlossene Mengen A_n sowie offene Mengen O_n mit

$$A_n \subseteq E_n \subseteq O_n, \quad \lambda(O_n \setminus A_n) \leq 1/n,$$

wobei jedes A_n beschränkt und damit kompakt ist.²⁷ Unsere Definitionen implizieren

$$\lambda_{\#}(E) \geq \lambda(A_n) = \lambda(E_n) - \lambda(E_n \setminus A_n) \geq \lambda(E_n) - \lambda(O_n \setminus A_n) = \lambda(E_n) - 1/n$$

und der Limes $n \rightarrow \infty$ garantiert

$$\lambda_{\#}(E) \geq \lambda(E).$$

Andererseits gilt

$$\lambda(O_n) = \lambda(E_n) + \lambda(O_n \setminus E_n) \leq \lambda(E_n) + \lambda(O_n \setminus A_n) \leq \lambda(E_n) + 1/n$$

und nach Grenzübergang erhalten wir

$$\lambda^{\#}(E) \leq \lambda(O) = \lambda(E),$$

wobei $O = \bigcup_{j=1}^{\infty} O_n$ selbst offen ist. Im Kombination mit $\lambda_{\#}(E) \leq \lambda^{\#}(E)$ (siehe die Bemerkungen oben) ergibt sich die Ungleichungskette

$$\lambda_{\#}(E) \leq \lambda^{\#}(E) \leq \lambda(E) \leq \lambda_{\#}(E),$$

und wir schließen, dass alle Terme gleich sind.

²⁶Es gilt offensichtlich $W_n \subseteq W_{n+1}$ sowie $\bigcup_{n=1}^{\infty} W_n = \mathbb{R}^d$.

²⁷Siehe den Satz von Heine-Borel aus *Analysis 2*.

Rückrichtung: Sei $E \subseteq \mathbb{R}^d$ gegeben mit $\lambda_{\#}(E) = \lambda^{\#}(E)$. Für jedes $n \in \mathbb{N}$ existieren aufgrund der Definition des inneren und äußeren Lebesgue-Maßes eine kompakte Menge K_n sowie eine offene Menge O_n mit

$$K_n \subseteq E \subseteq O_n, \quad \lambda(K_n) \leq \lambda_{\#}(E) \leq \lambda(K_n) + 1/n, \quad \lambda(O_n) - 1/n \leq \lambda^{\#}(E) \leq \lambda(O_n).$$

Wir setzen

$$K := \bigcup_{n=1}^{\infty} K_n, \quad O := \bigcap_{n=1}^{\infty} O_n$$

und bemerken, dass beide Mengen zu \mathcal{B} und damit auch zu \mathcal{L} gehören (wobei im Allgemeinen K nicht kompakt und O nicht offen sein wird). Dann gilt

$$K_n \subseteq K \subseteq E \subseteq O \subseteq O_n,$$

für alle $n \in \mathbb{N}$ und in Kombination mit der Voraussetzung erhalten wir

$$0 \leq \lambda(O \setminus K) \leq \lambda(O_n \setminus K_n) \leq (\lambda^{\#}(E) + 1/n) - (\lambda_{\#}(E) - 1/n) = 2/n.$$

Der Grenzübergang $n \rightarrow \infty$ zeigt, dass $O \setminus K$ eine Nullmenge ist, und die Vollständigkeit von λ garantiert, dass $E \setminus K$ als Teilmenge von $O \setminus K$ Lebesgue-messbar ist. Aus $E = K \cup (E \setminus K)$ folgt schließlich $E \in \mathcal{L}$. Außerdem ergibt sich via

$$\lim_{n \rightarrow \infty} \lambda(K_n) = \lambda(K) = \lambda_{\#}(E) = \lambda(E) = \lambda^{\#}(E) = \lambda(O) = \lim_{n \rightarrow \infty} \lambda(O_n)$$

wieder die gewünschte Doppelgleichung. □

Bemerkung Manche Autoren führen das Lebesgue-Maß mittels $\lambda_{\#}$ und $\lambda^{\#}$ ein. Der im vorherigen Abschnitt beschriebene Weg ist aber allgemeiner und kann auch auf andere Maße angewendet werden.

äußeres und inneres Jordan-Maß Für jedes $Y \subseteq \mathbb{R}^d$ setzen wir

$$\begin{aligned} \nu_{\#}(Y) &:= \sup \{ \lambda(R) : R \in \mathcal{Q} \text{ mit } R \subseteq Y \}, \\ \nu^{\#}(Y) &:= \inf \{ \lambda(S) : S \in \mathcal{Q} \text{ mit } Y \subseteq S \}, \end{aligned}$$

wobei $\inf \emptyset = \infty$ vereinbart sei (sodass zum Beispiel $\nu^{\#}(\mathbb{R}^d) = \infty$ gilt).

Definition Eine Menge $E \subseteq \mathbb{R}^d$ heißt Jordan-messbar, sofern

$$\nu_{\#}(E) = \nu^{\#}(E)$$

gilt, wobei dieser gemeinsame Wert dann der Jordan-Inhalt von E genannt und als $\nu(E)$ geschrieben wird.

Bemerkungen

1. Das Mengensystem aller Jordan-messbaren Mengen bezeichnen wir mit \mathcal{J} . Dieses ist ein Ring, aber keine σ -Algebra, und man kann zeigen, dass $\nu : \mathcal{J} \rightarrow [0, \infty)$ zwar additiv, aber nicht σ -additiv ist. Insbesondere ist ν kein Maß, sondern nur ein sogenannter *Inhalt*.

2. Es gilt $\nu_{\#}(Y) \leq \nu^{\#}(Y)$ für alle $Y \subseteq \mathbb{R}^d$.
3. Es gilt $E \in \mathcal{J}$ genau dann, wenn für jedes $\varepsilon > 0$ die Formeln

$$R \subseteq E \subseteq S, \quad \lambda(R \setminus S) \leq \varepsilon$$

für zwei geeignete Mengen R und S aus dem Quaderring \mathcal{Q} erfüllt sind.

4. Die Jordan-Messbarkeit impliziert die Lebesgue-Messbarkeit, wobei dann auch $\nu(E) = \lambda(E)$ gilt (Übungsaufgabe). Die umgekehrte Aussage ist aber falsch. Das Standardbeispiel ist wieder die Dirichlet-Menge D (siehe oben). Für diese berechnen wir

$$\nu_{\#}(D) = 0 < 1 = \nu^{\#}(D),$$

denn es kann nur dann $R \subseteq D \subseteq S$ für $R, S \in \mathcal{Q}$ gelten, wenn R die leere Menge ist und S den ganzen Einheitswürfel enthält.

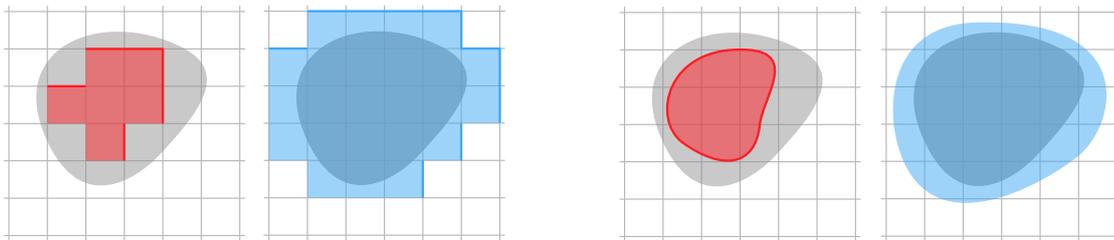


Abbildung Beim Jordan-Inhalt approximieren wir eine gegebene Menge (grau) von innen (rot) und von außen (blau) mit jeweils endlich vielen Quadern (links), wohingegen beim Lebesgue-Maß nach der äquivalenten Charakterisierung innen eine beliebige abgeschlossene und außen eine beliebige offene Menge verwendet werden darf (rechts). Siehe auch die Diskussion im Text.

Diskussion

1. Die Jordan-Messbarkeit entspricht im Wesentlichen dem naiven Flächen- und Volumenbegriff aus der Schule, d.h. dem Kästchenzählen auf Millimeterpapier. Insbesondere kann eine Jordan-messbare Menge E sowohl von innen als auch von außen durch endlich viele Quader angenähert werden, wobei die Betonung sowohl auf *endlich viele* als auch auf *Quader* liegt. Dabei ist es jedoch nicht wichtig, dass diese Quader halboffen sind, und man könnte die Definition alternativ auch mit abgeschlossenen oder offenen Quadern formulieren.
2. Für viele — und eigentlich alle in der Praxis auftretenden — Mengen liefern beide Ansätze dasselbe d -dimensionale Volumen, aber es gibt mehr Lebesgue-messbare als Jordan-messbare Mengen. Insbesondere besitzt aus theoretischer Sicht nur das Lebesgue-Maß wirklich gute Eigenschaften und führt zu einem wesentlich robusteren Integralbegriff.

Theorem (äquivalente Charakterisierung von Jordan-Messbarkeit) Eine Teilmenge $E \subseteq \mathbb{R}^d$ gehört genau dann zu \mathcal{J} , wenn sie beschränkt ist und wenn ihr Rand eine Lebesguesche Nullmenge ist.

Beweis* Hinrichtung: Für fixiertes $E \in \mathcal{J}$ und jedes $n \in \mathbb{N}$ wählen wir zwei Mengen R_n, S_n aus dem Quaderring \mathcal{Q} , sodass

$$R_n \subseteq E \subseteq S_n \quad \nu(E) - 1/n \leq \lambda(R_n) \leq \lambda(S_n) \leq \nu(E) + 1/n.$$

Dann gilt²⁸

$$\text{bnd}(E) = \text{cls}(E) \setminus \text{int}(E) \subseteq B_n := \text{cls}(S_n) - \text{int}(R_n)$$

sowie

$$\lambda(B_n) \leq \lambda(\text{cls}(S_n)) - \lambda(\text{int}(R_n)) = \lambda(S_n) - \lambda(R_n) \leq 2/n,$$

wobei wir benutzt haben, dass

$$\lambda(\text{int}(M)) = \lambda(M) = \lambda(\text{cls}(M))$$

für alle $M \in \mathcal{Q}$ gilt (da diese jeweils endliche Vereinigung paarweise disjunkter Quader sind). Mit $B := \bigcup_{n=1}^{\infty} B_n \in \mathcal{B}$ gilt $\lambda(B) \leq \lambda(B_n) \leq 2/n$ für alle $n \in \mathbb{N}$ und wir schließen, dass B eine Lebesguesche Nullmenge ist. Wegen der Vollständigkeit von λ gilt dies aber auch für den Rand von E . Außerdem ist E als Teilmenge einer endlichen Vereinigung von Quadern beschränkt.

Rückrichtung, Teil 1: Sei $E \subseteq \mathbb{R}^d$ beschränkt mit $\lambda(\text{bnd}(E)) = 0$. Dann sind $\text{int}(E)$ und $\text{cls}(E)$ als Borel-Mengen auch Lebesgue-messbar und in Kombination mit

$$E = \text{int}(E) \cup (E \setminus \text{int}(E)), \quad E \setminus \text{int}(E) \subseteq \text{cls}(E) \setminus \text{int}(E) = \text{bnd}(E)$$

erhalten wir

$$E \in \mathcal{L}, \quad \lambda(\text{int}(E)) = \lambda(E) = \lambda(\text{cls}(E)),$$

da die Vollständigkeit von λ die Lebesgue-Messbarkeit von $E \setminus \text{int}(E)$ sicherstellt. Das Zerlegungslemma liefert eine Folge $(Q_j)_{j=1}^{\infty}$ von paarweise disjunkten Quadern mit

$$\text{int}(E) = \bigcup_{j=1}^{\infty} Q_j, \quad \lambda(\text{int}(E)) = \sum_{j=1}^{\infty} \lambda(Q_j).$$

Für jedes $n \in \mathbb{N}$ setzen wir $R_n := \bigcup_{j=1}^n Q_j$ und bemerken, dass die Abschätzung

$$\nu_{\#}(E) \geq \lambda(R_n) = \sum_{j=1}^n \lambda(Q_j)$$

wegen $R_n \in \mathcal{Q}$ und $R_n \subseteq E$ erfüllt ist. Der Limes $n \rightarrow \infty$ liefert schließlich

$$\nu_{\#}(E) \geq \lambda(\text{int}(E)) = \lambda(E).$$

Rückrichtung, Teil 2: Sei nun $\varepsilon > 0$ beliebig. Dann existiert eine Folge $(M_j)_{j=1}^{\infty}$ in \mathcal{Q} mit

$$\text{cls}(E) \subseteq \bigcup_{j=1}^{\infty} M_j, \quad \sum_{j=1}^{\infty} \lambda(M_j) \leq \lambda(\text{cls}(E)) + \varepsilon = \lambda(E) + \varepsilon,$$

²⁸Wie schon in *Analysis 2* benutzen wir die folgenden Notationen für eine beliebige Menge $U \subseteq \mathbb{R}^d$: $\text{int}(U)$ bzw. $\text{cls}(U)$ ist das Innere bzw. der Abschluss von U und $\text{bnd}(U)$ ist der Rand.

wobei jedes M_j via

$$M_j = \bigcup_{i=1}^{m_j} Q_{j,i}, \quad \lambda(M_j) = \sum_{i=1}^{m_j} \lambda(Q_{j,i})$$

als endliche und paarweise disjunkte Vereinigung von halboffenen Quadern dargestellt werden kann. Außerdem wählen wir für jedes $Q_{j,i}$ einen offenen Quader $O_{j,i}$ sowie einen halboffenen Quader $P_{j,i}$ mit

$$Q_{j,i} \subseteq O_{j,i} \subseteq P_{j,i}, \quad \lambda(P_{j,i} \setminus Q_{j,i}) \leq \frac{\varepsilon}{2^j m_j}$$

und erhalten

$$\text{cls}(E) \subseteq \bigcup_{j=1}^{\infty} \bigcup_{i=1}^{m_j} O_{j,i} \subseteq \bigcup_{j=1}^{\infty} \bigcup_{i=1}^{m_j} P_{j,i}, \quad \sum_{j=1}^{\infty} \sum_{i=1}^{m_j} \lambda(P_{j,i}) \leq \lambda(E) + \varepsilon,$$

wobei wir $\lambda(P_{j,i}) = \lambda(Q_{j,i}) + \lambda(P_{j,i} \setminus Q_{j,i})$ sowie die Nebenrechnung

$$\sum_{j=1}^{\infty} \sum_{i=1}^{m_j} \frac{\varepsilon}{2^j m_j} = \sum_{j=1}^{\infty} \frac{\varepsilon}{2^j} = \varepsilon$$

verwendet haben. Nach Konstruktion überdecken alle offenen Quader $O_{j,i}$ zusammen die kompakte Menge $\text{cls}(E)$ und wir können daher eine endliche Teilüberdeckung wählen. Durch Vereinigung der entsprechenden Quader $P_{j,i}$ erhalten wir schließlich eine Menge $S \in \mathcal{Q}$ mit

$$E \subseteq \text{cls}(E) \subseteq S, \quad \lambda(S) \leq \lambda(E) + \varepsilon,$$

und dies impliziert $\nu^\#(E) \leq \lambda(E) + \varepsilon$. Insgesamt ergibt sich

$$\nu_\#(E) \leq \nu^\#(E) \leq \lambda(E) + \varepsilon \leq \nu_\#(E) + \varepsilon$$

und weil ε beliebig war, folgt $\nu_\#(E) = \nu^\#(E) = \lambda(E)$ bzw. $E \in \mathcal{J}$. □

1.5 Messbare Funktionen

Vorbemerkung Wir betrachten in diesem Abschnitt einen messbaren Raum (X, \mathcal{A}) , der aus einer Menge X und einer σ -Algebra \mathcal{A} in X besteht. Bei der ersten Lektüre können Sie sich vorstellen, dass $X = \mathbb{R}^d$ und $\mathcal{A} = \mathcal{L}$ gilt, aber andere Anwendungen sind auch möglich und wichtig. Im Folgenden bezeichnet \mathcal{B} immer die Borelsche σ -Algebra im Bildraum \mathbb{R} .

Definition Eine Funktion $f : X \rightarrow \mathbb{R}$ heißt messbar (bzgl. \mathcal{A}), falls $f^{-1}(U) \in \mathcal{A}$ für jedes $U \in \mathcal{B}$ gilt, d.h. wenn das Urbild jeder borelschen Teilmenge von \mathbb{R} als Teilmenge von X messbar ist.

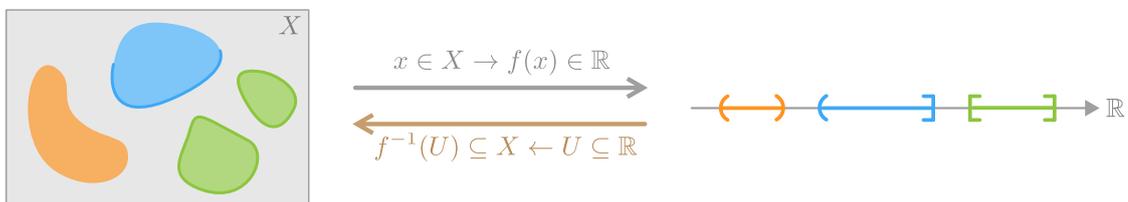


Abbildung Bei einer messbaren Funktion $f : X \rightarrow \mathbb{R}$ gilt: Für jede Borel-Menge aus dem Bildraum (rechts) ist das entsprechende Urbild (links) messbar (siehe den braunen Pfeil). Beachte aber, dass das Bild einer messbaren Menge durchaus eine nicht-borelsche Menge sein kann.

Bemerkungen

1. Für jede messbare Menge $E \in \mathcal{A}$ ist ihre charakteristische Funktion (oder Indikatorfunktion) $\chi_E : X \rightarrow \mathbb{R}$ durch

$$\chi_E(x) := \begin{cases} 1 & \text{falls } x \in E, \\ 0 & \text{sonst} \end{cases}$$

definiert und außerdem messbar (denn für $f^{-1}(U)$ kommen nur \emptyset , E , E' oder X in Frage, je nachdem ob U die 0 und die 1 enthält oder nicht).

2. Die Menge aller messbaren Funktionen auf X kann in natürlicher Weise als reeller Vektorraum betrachtet werden. Siehe dazu die Übungen.

Lemma (äquivalente Charakterisierung von Messbarkeit) Eine Funktion $f : X \rightarrow \mathbb{R}$ ist genau dann messbar, wenn ihre Urbildmenge $f^{-1}((-\infty, c])$ für jedes $c \in \mathbb{R}$ messbar ist.

Beweis Hinrichtung: Die Behauptung folgt unmittelbar aus der Definition, da jedes Intervall $(-\infty, c]$ eine Borel-Menge in \mathbb{R} ist.

Rückrichtung: Sei f eine Funktion mit der Urbild-Eigenschaft. Die Rechenregeln für Urbilder implizieren unmittelbar, dass das Mengensystem

$$\mathcal{M} := \{U \subseteq \mathbb{R} : f^{-1}(U) \in \mathcal{A}\}$$

eine σ -Algebra in \mathcal{A} ist (Übungsaufgabe). Nach Voraussetzung gilt $(-\infty, c] \in \mathcal{M}$ für alle $c \in \mathbb{R}$ und wegen

$$f^{-1}((c_1, c_2]) = f^{-1}((-\infty, c_2]) \setminus f^{-1}((-\infty, c_1])$$

schließen wir, dass \mathcal{M} auch alle halboffenen Intervalle $(c_1, c_2]$ für beliebige c_1, c_2 mit $-\infty < c_1 < c_2 < +\infty$ enthält. Der Zerlegungssatz (ausgewertet in \mathbb{R} , also für $d = 1$) garantiert außerdem, dass die σ -Algebra \mathcal{M} auch alle offenen Teilmengen von \mathbb{R} enthält, und die Minimalität der Borel-Algebra stellt sicher, dass $\mathcal{B} \subseteq \mathcal{M}$ gilt. \square

Bemerkungen

1. Drei weitere äquivalente Charakterisierungen entstehen durch die Forderung, dass alle Urbilder der Bauart

$$f^{-1}((-\infty, c)) \quad \text{oder} \quad f^{-1}((c, +\infty)) \quad \text{oder} \quad f^{-1}([c, +\infty))$$

in der σ -Algebra \mathcal{A} liegen.

2. Per Definition gilt

$$f^{-1}((-\infty, c]) = \{x \in X : f(x) \leq c\}, \quad f^{-1}((-\infty, c)) = \{x \in X : f(x) < c\}$$

und analoge Formeln beschreiben andere Urbildmengen.

3. Mit einem analogen Beweis können wir auch die folgenden zwei Äquivalenzen begründen: Eine Funktion $f : X \rightarrow \mathbb{R}$ ist genau dann messbar, wenn das Urbild jeder offenen (oder jeder abgeschlossenen) Teilmenge von \mathbb{R} unter f zur σ -Algebra \mathcal{A} gehört.
4. Ist $f : X \rightarrow \mathbb{R}$ messbar und ist $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ stetig, so ist auch die Komposition $\varphi \circ f : X \rightarrow \mathbb{R}$ stetig.²⁹
5. Im Fall von $X = \mathbb{R}^d$ unterscheiden wir zwischen Lebesgue-Messbarkeit und Borel-Messbarkeit, wobei sich diese Bezeichnungen auf die Wahl von \mathcal{A} im Urbildraum bezieht.

Klarstellung: Für $X = \mathbb{R}^d$ verwenden wir im *Urbildraum* standardmäßig die d -dimensionale Lebesgue-Algebra und nur ausnahmsweise die d -dimensionale Borel-Algebra (wobei wir dies dann extra betonen werden). Insbesondere meint Messbarkeit bei uns im Zweifel immer Lebesgue-Messbarkeit. Im *Bildraum* \mathbb{R} benutzen wir jedoch ausschließlich die eindimensionale Borel-Algebra, denn bei den Untersuchung zur Messbarkeit einer Funktion f interessieren nur die Urbilder von Borel-Mengen, aber nicht die von Lebesgue-Mengen.

6. Insbesondere gilt für $X = \mathbb{R}^d$: Jede stetige Funktion auf \mathbb{R}^d ist sowohl Lebesgue- als auch Borel-messbar. Da es mehr Lebesgue- als Borel-Mengen gibt, impliziert die zweite Eigenschaft die erste (aber nicht umgekehrt).
7. Sind $f, g : X \rightarrow \mathbb{R}$ zwei messbare Funktionen, so sind die vier Mengen

$$\{x \in X : f(x) < g(x)\}, \quad \{x \in X : f(x) \leq g(x)\}$$

und

$$\{x \in X : f(x) = g(x)\}, \quad \{x \in X : f(x) \neq g(x)\}$$

auch jeweils messbar, wobei der Beweis in den Übungen besprochen wird.

²⁹Dies ergibt sich aus der Tatsache, dass $\varphi^{-1}(O)$ für jede offene Teilmenge $O \subseteq \mathbb{R}$ auch offen in \mathbb{R} ist. Siehe dazu die äquivalente Charakterisierung von Stetigkeit in *Analysis 1+2*.

Lemma (nützliche Resultate) Mit $f : X \rightarrow \mathbb{R}$ sind auch $|f|, f_-, f_+ : X \rightarrow \mathbb{R}$ messbar, wobei diese Funktionen durch

$$|f|(x) := |f(x)|, \quad f_-(x) := -\min\{0, f(x)\}, \quad f_+(x) := +\max\{0, f(x)\}$$

definiert sind.

Beweis Für jedes $c \geq 0$ impliziert die Nebenrechnung

$$|f|(x) \leq c \Leftrightarrow -c \leq f(x) \leq c$$

die Formel

$$|f|^{-1}((-\infty, c]) = f^{-1}([-c, +c]) \in \mathcal{A},$$

wohingegen sich

$$f_-((-\infty, c]) = f^{-1}([-c, 0]) \in \mathcal{A}, \quad f_+((-\infty, c]) = f^{-1}([0, +c]) \in \mathcal{A}$$

aus

$$f_-(x) \leq c \Leftrightarrow -c \leq f(x) \leq 0, \quad f_+(x) \leq c \Leftrightarrow 0 \leq f(x) \leq +c$$

ergibt. Für alle $c < 0$ gilt außerdem

$$|f|^{-1}((-\infty, c]) = f_-^{-1}((-\infty, c]) = f_+^{-1}((-\infty, c]) = \emptyset \in \mathcal{A}$$

und die Behauptung folgt unmittelbar aus dem vorherigen Lemma zur äquivalenten Charakterisierung von Messbarkeit. \square

Bemerkungen

- Wir nennen f_- bzw. f_+ den Negativteil bzw. den Positivteil von f , wobei

$$f(x) = f_+(x) - f_-(x), \quad |f(x)| = f_+(x) + f_-(x), \quad f_+(x) f_-(x) = 0$$

für alle $x \in X$ gilt.

Achtung: Der Negativteil einer nichtnegativen Funktion, d.h. es gilt $f_-(x) \geq 0$ für alle $x \in X$. Es gilt außerdem $f_- = (-f)_+$, d.h. der Negativteil von f ist gerade der Positivteil von $-f$.

- Allgemeiner können wir für zwei messbare Funktionen $f, g : X \rightarrow \mathbb{R}$ zeigen, dass durch

$$\max\{f, g\}(x) := \max\{f(x), g(x)\}, \quad \min\{f, g\}(x) := \min\{f(x), g(x)\}$$

sowie

$$(f + g)(x) := f(x) + g(x), \quad (fg)(x) := f(x)g(x)$$

wier weitere messbare Funktionen definiert werden. Außerdem ist jedes skalare Vielfache einer messbaren Funktion auch messbar. Siehe dazu die Übungen.

Verallgemeinerung Oftmals ist es sinnvoll, auch unendlich große Funktionswerte zuzulassen und sogenannte numerische Funktionen $f : X \rightarrow \overline{\mathbb{R}}$ zu betrachten, die die Menge X in die erweiterten reellen Zahlen

$$\overline{\mathbb{R}} := \{-\infty\} \cup \mathbb{R} \cup \{+\infty\}$$

abbilden.³⁰ Eine solche Funktion wird messbar genannt, falls die beiden Mengen $f^{-1}(\{-\infty\})$ und $f^{-1}(\{+\infty\})$ sowie alle Urbilder $f^{-1}(U)$ von Borel-Mengen $U \subseteq \mathbb{R}$ in \mathcal{A} liegen.

Merkregel: Wir können mit numerischen Funktionen fast genauso gut arbeiten wie mit skalaren Funktionen. An einigen Stellen müssen wir etwas aufpassen und konkrete Formulierungen oder Argumente abändern, aber in aller Regel sind die notwendigen Modifikationen leicht zu verstehen und umzusetzen.

Theorem (Messbarkeit und punktweise Konvergenz) Sei $(f_n)_{n=1}^\infty$ eine Folge numerischen und messbaren Funktionen $f_n : X \rightarrow \overline{\mathbb{R}}$, die via

$$f_n(x) \xrightarrow{n \rightarrow \infty} f(x)$$

für $n \rightarrow \infty$ punktweise gegen $f : X \rightarrow \overline{\mathbb{R}}$ konvergiert (wobei dies im eigentlichen oder uneigentlichen Sinne gelten kann). Dann ist die Grenzfunktion auch messbar.

Beweis Teil 1: Wir nehmen zunächst an, dass f sowie alle Funktionen f_n weder den Wert $-\infty$ noch den Wert $+\infty$ annehmen und fixieren $c \in \mathbb{R}$ beliebig. Die Menge

$$E := \bigcap_{m=1}^{\infty} E_m, \quad E_m := \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} \{x \in X : f_n(x) \leq c + 1/m\}$$

gehört nach Voraussetzung zur σ -Algebra \mathcal{A} , da jede der ganz rechts auftretenden Mengen der Bauart $f_n^{-1}((-\infty, c + 1/m])$ ist und damit zu \mathcal{A} gehört. Wir wollen nun

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

zeigen, da dies wegen der Beliebigkeit von c sofort die Behauptung impliziert. Dazu betrachten wir zunächst ein beliebiges $x \in E$ sowie ein festes $m \in \mathbb{N}$. Wegen $x \in E_m$ existiert ein $k \in \mathbb{N}$ mit $f_n(x) \leq c + 1/m$ für alle $n \geq k$ und die punktweise Konvergenz aus der Voraussetzung garantiert $f(x) \leq c + 1/m$. Da dies für alle $m \in \mathbb{N}$ gilt, folgt schließlich $f(x) \leq c$ wie gewünscht. Gelte nun umgekehrt $f(x) \leq c$ und sei $m \in \mathbb{N}$ beliebig. Aufgrund der punktweisen Konvergenz existiert ein Index $k \in \mathbb{N}$, sodass $f_n(x) \leq c + 1/m$ für alle $n \geq k$ gilt. Wir erhalten daher zunächst $x \in E_m$ und anschließend $x \in E$ wegen der Beliebigkeit von m .

Teile 2: Der allgemeine Fall kann mit analogen Argumenten behandelt werden, sofern einige naheliegende Fallunterscheidungen getroffen werden. Die Details sind eine Übungsaufgabe. \square

Bemerkung Dieses Theorem erscheint zunächst unscheinbar. Es ist aber sehr wichtig und wird in vielen Beweisen eine Rolle spielen. Insbesondere garantiert es, dass die Messbarkeit von Funktionen unter punktweiser Konvergenz erhalten bleibt.

³⁰Diese Bezeichnung ist sicher nicht optimal, hat sich aber so eingebürgert.

Lemma (Messbarkeit spezieller punktwiser Definitionen) Sei $(f_n)_{n=1}^\infty$ eine Folge numerischer Funktionen $f_n : X \rightarrow \overline{\mathbb{R}}$. Dann werden durch

$$x \mapsto \inf \{f_n(x) : n \in \mathbb{N}\}, \quad x \mapsto \sup \{f_n(x) : n \in \mathbb{N}\}$$

sowie

$$x \mapsto \liminf_{n \rightarrow \infty} f_n(x), \quad x \mapsto \limsup_{n \rightarrow \infty} f_n(x)$$

vier weitere numerische Funktionen definiert, die jeweils messbar sind.

Beweis Teil 1: Wir definieren die Folge $(g_n)_{n=1}^\infty$ rekursiv durch

$$g_1(x) = f_1(x), \quad g_{n+1}(x) = \min \{g_n(x), f_{n+1}(x)\}$$

und zeigen induktiv, dass jede Funktion $g_n : X \rightarrow \overline{\mathbb{R}}$ messbar ist (siehe dazu auch die Bemerkungen zu den nützlichen Resultaten) und dass darüber hinaus $g_{n+1}(x) \leq g_n(x)$ für jedes $x \in X$ und alle $n \in \mathbb{N}$ gilt. Insbesondere ist

$$g(x) := \lim_{n \rightarrow \infty} g_n(x)$$

im Sinne der monotonen punktwisen Konvergenz wohldefiniert und das vorherige Theorem liefert die Messbarkeit der Grenzfunktion $g : X \rightarrow \overline{\mathbb{R}}$. Andererseits folgt

$$g(x) = \inf \{f_n(x) : n \in \mathbb{N}\}$$

aus elementaren Überlegungen und wir haben die erste Behauptung bewiesen. Die zweite kann analog hergeleitet werden, sofern wir in der Definition von g_n das Maximum durch das Minimum ersetzen.

Teil 2: Nach dem ersten Teil ist die Funktion

$$h_k(x) := \inf \{f_n(x) : n \in \mathbb{N} \text{ mit } n \geq k\}$$

für jedes $k \in \mathbb{N}$ messbar, wobei außerdem $h_{k+1}(x) \geq h_k(x)$ für jedes $x \in X$ und alle $n \in \mathbb{N}$ gilt. Die punktwise Grenzfunktion $h : X \rightarrow \overline{\mathbb{R}}$ mit

$$h(x) := \lim_{k \rightarrow \infty} h_k(x)$$

ist also wohldefiniert und nach dem Theorem auch messbar. Andererseits gilt

$$h(x) = \liminf_{n \rightarrow \infty} f_n(x)$$

nach der äquivalenten Charakterisierung des kleinsten und größten Häufungspunktes aus *Analysis 1*. Es folgt die dritte Behauptung und die vierte kann analog hergeleitet werden. \square

Korollar Für jede Folge $(f_n)_{n=1}^\infty$ messbarer Funktionen $f_n : X \rightarrow \overline{\mathbb{R}}$ ist die Menge

$$\{x \in X : f_n(x) \text{ konvergiert für } n \rightarrow \infty \text{ gegen einen Grenzwert aus } \overline{\mathbb{R}}\}$$

messbar.

Beweis Das Lemma garantiert, dass durch

$$\underline{f}(x) := \liminf_{n \rightarrow \infty} f_n(x), \quad \overline{f}(x) := \limsup_{n \rightarrow \infty} f_n(x),$$

zwei messbare Funktionen \underline{f} und \overline{f} definiert werden. Insbesondere ist die Menge

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

messbar, aber das ist gerade die Menge in der Behauptung. \square

einfache Funktionen

Definition Eine Funktion $f : X \rightarrow Y$ heißt einfach, wenn es ein $l \in \mathbb{N}$ sowie reelle Zahlen y_1, \dots, y_l und messbare Menge $E_1, \dots, E_l \in \mathcal{A}$ gibt, sodass

$$f(x) = \sum_{j=1}^l y_j \chi_{E_j}(x)$$

für $x \in X$ gilt. Wenn außerdem

$$y_{j_1} \neq y_{j_2} \quad \text{und} \quad E_{j_1} \cap E_{j_2} = \emptyset \quad \text{für alle} \quad j_1 \neq j_2$$

erfüllt ist, so sagen wir, f sei in Normalform gegeben.

Bemerkungen

1. Durch eine geeignete Umdefinition der y_j und E_j sowie der Anzahl l können wir jede einfache Funktion in Normalform bringen.

Beispiel 1 : Für beliebige Mengen $E_1, E_2 \in \mathcal{A}$ gilt

$$y_1 \chi_{E_1}(x) + y_2 \chi_{E_2}(x) = \tilde{y}_1 \chi_{\tilde{E}_1}(x) + \tilde{y}_2 \chi_{\tilde{E}_2}(x) + \tilde{y}_3 \chi_{\tilde{E}_3}(x)$$

für alle $x \in X$, sofern wir

$$\tilde{E}_1 := E_1 \setminus E_2, \quad \tilde{E}_2 := E_2 \setminus E_1, \quad \tilde{E}_3 := E_1 \cap E_2$$

und

$$\tilde{y}_1 := y_1, \quad \tilde{y}_2 := y_2, \quad \tilde{y}_3 := y_1 + y_2$$

setzen (siehe auch die nachfolgende Abbildung).

Beispiel 2 : Für $y_1 = y_2 \neq y_3$ ergibt sich

$$y_1 \chi_{E_1}(x) + y_2 \chi_{E_2}(x) + y_3 \chi_{E_3}(x) = \tilde{y}_1 \chi_{\tilde{E}_1}(x) + \tilde{y}_2 \chi_{\tilde{E}_2}(x)$$

aus

$$\tilde{E}_1 := E_1 \cup E_2, \quad \tilde{E}_2 := E_3, \quad \tilde{y}_1 := y_1 = y_2, \quad \tilde{y}_2 := y_3.$$

2. Eine einfache Funktion besitzt sehr viele verschiedene Darstellungen, aber die ihre Normalform ist eindeutig festgelegt (Übungsaufgabe).

3. Jede einfache Funktion ist messbar, denn für jedes $c \in \mathbb{R}$ ist die Urbildmenge $f^{-1}((-\infty, c])$ gerade die Vereinigung aller E_j mit $y_j \leq c$. In Normalform gilt

$$E_j = f^{-1}(\{y_j\}), \quad f(x) = y_j \Leftrightarrow x \in E_j$$

für alle $j = 1, \dots, l$.

4. Die Menge aller einfachen Funktionen bezeichnen wir mit $\mathcal{E}(X)$, wobei diese in natürlicher Weise ein reeller Vektorraum ist. Außerdem sind das Produkt, die Minimum und das Maximum einfacher Funktionen wieder einfach. Diese Eigenschaften sowie die entsprechenden Normalformen werden in den Übungen genauer diskutiert.



Abbildung Der Übergang zur Normalform für die beiden Beispiele aus der Bemerkung.

Diskussion Einfache Funktionen spielen in der allgemeinen Integrationstheorie ungefähr die Rolle, die Treppenfunktionen beim Riemann-Integral haben (siehe *Analysis 1* für den Fall $d = 1$). Beachte aber, dass jede der Mengen $E_j = f^{-1}(\{y_j\})$ sehr kompliziert sein kann, wohingegen in der Riemannschen Theorie immer vorausgesetzt wird, dass diese eine Vereinigung von endlich vielen Quadern ist. Siehe dazu auch die unten eingeführten Approximationsfolgen für messbare Funktionen.

Beispiel: Die charakteristische Funktion χ_D der Dirichlet-Menge D ist einfach (bzgl. der Lebesgue-Algebra in \mathbb{R}^d), kann aber nicht durch Treppenfunktionen angenähert werden.

Verallgemeinerung Manchmal ist es sinnvoll, auch abzählbar viele Summanden zuzulassen. Wir nennen $f : X \rightarrow \mathbb{R}$ eine verallgemeinerte einfache Funktion, wenn diese als

$$f(x) = \sum_{j=1}^{\infty} y_l \chi_{E_j}(x) \quad \text{oder} \quad f(x) = \sum_{j=-\infty}^{\infty} y_l \chi_{E_j}(x)$$

geschrieben werden kann, wobei wir dann immer voraussetzen, dass das Analogon der Normalform-Bedingung erfüllt ist, d.h. dass die y_j paarweise verschieden und die E_j paarweise disjunkt sind.

Bemerkungen:

1. Die Bedingung der paarweisen Disjunktheit stellt sicher, dass jeder Punkt $x \in X$ zu höchstens einer Menge E_j gehört. Insbesondere sind beide Varianten der unendlichen Summe punktweise immer wohldefiniert, da fast alle Summanden verschwinden.
2. Da \mathbb{N} und \mathbb{Z} gleichmächtig sind, kann jeder der beiden Reihenformeln durch eine geeignete Reindizierung in die andere überführt werden. Je nach Kontext ist mal die eine und mal die andere Darstellung besser geeignet.

gleichmäßige Approximation durch Abrunden Die Formel³¹

$$f_m(x) := \lfloor m f(x) \rfloor / m,$$

liefert für jede messbare Funktion eine approximierende Folge verallgemeinerter einfacher Funktionen $(f_m)_{m=1}^\infty$, wobei die alternative Darstellungsformel

$$f_m(x) = \sum_{j=-\infty}^{+\infty} \frac{j}{m} \chi_{E_{m,j}}(x), \quad E_{m,j} = \left\{ x \in X : \frac{j}{m} \leq f(x) < \frac{j+1}{m} \right\},$$

mit einfachen Argumenten hergeleitet werden kann (siehe dazu die Abbildung). Unsere Konstruktion impliziert

$$0 \leq f(x) - f_m(x) \leq 1/m$$

für alle $x \in X$ und damit

$$\|f_m - f\|_\infty := \sup \{ |f_m(x) - f(x)| : x \in X \} \xrightarrow{m \rightarrow \infty} 0.$$

Insbesondere konvergiert f_m für $m \rightarrow \infty$ *gleichmäßig* und damit auch *punktweise* gegen die messbare Funktion f .

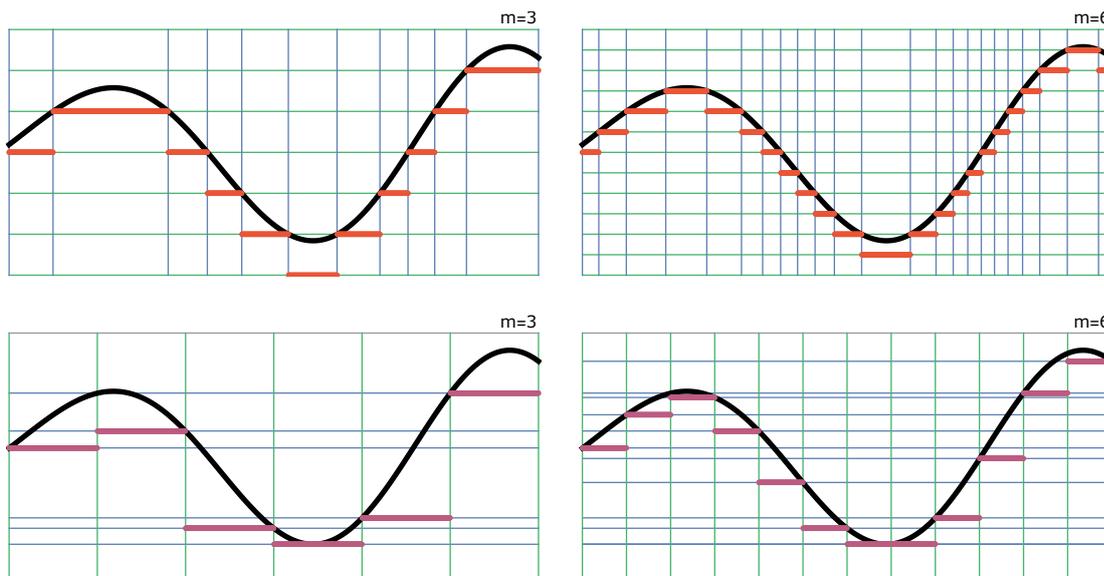


Abbildung *Oben*: Approximation einer beschränkten und messbaren Funktion (schwarz) durch einfache Funktionen (rot) von unten, wobei der Bildbereich zunächst in Intervalle der Dicke $1/m$ zerlegt wird (grüne horizontale Linien) und anschließend die entsprechenden messbaren Mengen im Urbildbereich ermittelt werden (blaue vertikale Linien). *Unten*: Bei der Approximation mit Treppenfunktionen (lila) in der Riemannschen Integrationstheorie ist es genau andersherum: Hier wird eine Zerlegung des Urbildbereichs vorgegeben, die dann eine entsprechende Zerlegung des Bildbereichs festlegt (grüne bzw. blaue Linien nun vertikal bzw. horizontal). Für stetig differenzierbare Funktionen macht das keinen großen Unterschied, aber es gibt Funktionen, die durch einfache Funktionen, jedoch nicht durch Treppenfunktionen approximiert werden können (etwa die charakteristische Funktion χ_D der Dirichlet-Menge).

Bemerkungen:

1. Wenn f zusätzlich beschränkt ist, so ist jede Funktion f_m sogar einfach.

³¹Wir benutzen hier noch einmal die Gauß-Klammer des Abrundens. Es gilt also $\lfloor y \rfloor \in \mathbb{Z}$ sowie $0 \leq y - \lfloor y \rfloor < 1$ für alle $y \in \mathbb{R}$.

2. Die punktweise Konvergenz $f_m(x) \rightarrow f(x)$ ist im Allgemeinen nicht monoton, aber nach Übergang zu der durch $m_n = 2^n$ definierte Teilfolge kann dies sichergestellt werden. Siehe dazu den Beweis des Approximationssatzes.
3. Eine analoge Approximationsfolge kann durch Aufrunden gewonnen werden.
4. Wenn f numerisch und messbar ist, so kann die approximierende Folge ganz analog definiert werden, wobei die Funktionen f_m dann aber numerisch sind.
5. Die Funktionen f_m verdeutlichen auch den wesentlichen Unterschied zwischen der Riemannschen und der Lebesgueschen Integrationstheorie (siehe die Abbildung).

Lemma (monotone Approximation durch einfache Funktionen) Für jede nichtnegative messbare Funktion $f : X \rightarrow \overline{\mathbb{R}}$ existiert eine monoton wachsende Folge $(f_n)_{n=1}^\infty$ von einfachen und nichtnegativen Funktionen $f_n : X \rightarrow \mathbb{R}$, die für $n \rightarrow \infty$ punktweise von unten gegen f konvergiert.³²

Beweis Teil 1: Wir betrachten zunächst eine Funktion $f : X \rightarrow \mathbb{R}$ und setzen

$$g_n(x) := \lfloor 2^n f(x) \rfloor / 2^n.$$

und erhalten eine verallgemeinerte einfache Funktion, die nichtnegativ ist und wegen

$$0 \leq f(x) - g_n(x) \leq \frac{1}{2^n}$$

für $n \rightarrow \infty$ gleichmäßig und damit auch punktweise gegen f konvergiert. Sei nun $x \in X$ fixiert. Dann existiert für jedes $n \in \mathbb{N}$ eine $j_n \in \mathbb{Z}$ mit

$$f(x) \in I_n := \left[\frac{j_n}{2^n}, \frac{j_n + 1}{2^n} \right) \subseteq \mathbb{R},$$

wobei die Definition von I_n sicherstellt, dass I_{n+1} entweder die linke oder die rechte Hälfte von I_n ist. Im ersten bzw. im zweiten Fall ergibt sich

$$g_{n+1}(x) = g_n(x) \quad \text{bzw.} \quad g_{n+1}(x) = g_n(x) + 2^{-(n+1)}$$

und wir schließen insgesamt, dass die Funktionenfolge $(g_n(x))_{n=1}^\infty$ monoton wachsend ist (im nicht-strikten Sinne) und von unten gegen $f(x)$ konvergiert.

Teil 2: Ist f beschränkt, so ist jede Funktion g_n sogar einfach und wir können direkt $f_n := g_n$ setzen. Andernfalls nimmt $f : X \rightarrow \mathbb{R}$ beliebig große Werte an und wir erhalten nach „Abschneiden“ via

$$f_n(x) := \min \{n, g_n(x)\} = n \chi_{E_n}(x) + \chi_{E_n^c}(x) g_n(x), \quad E_n := \{x \in X : f(x) > n\}$$

für jedes $n \in \mathbb{N}$ eine einfache Funktion f_n auf X . Insbesondere ist $(f_n(x))_{n=1}^\infty$ für jedes $x \in X$ nach Konstruktion eine monoton wachsende Zahlenfolge mit Grenzwert $f(x)$.

Teil 3: Wenn f den Wert $+\infty$ auf einer Menge $E \in \mathcal{A}$ annimmt, wenden wir alle Argumente aus den ersten beiden Beweisteilen auf die Funktion

$$\tilde{f} := \chi_{E^c} f, \quad \tilde{f}(x) = \begin{cases} 0 & \text{falls } f(x) = \infty \\ f(x) & \text{sonst} \end{cases}$$

an. Dies liefert eine approximierende Folge $(\tilde{f}_n)_{n=1}^\infty$ einfacher Funktionen und mit

$$f_n := \tilde{f}_n + n \chi_E f, \quad f_n(x) = \begin{cases} n & \text{falls } f(x) = \infty \\ \tilde{f}_n(x) & \text{sonst} \end{cases}$$

ergibt sich die Behauptung. □

³²Eine Funktion $f : X \rightarrow \mathbb{R}$ wird nichtnegativ genannt, wenn $f(x) \geq 0$ für alle $x \in X$ gilt.

Bemerkungen

1. Die Funktionen f_n sind genau dann monoton wachsend, wenn für jedes $x \in X$ die Folge $(f_n(x))_{n=1}^{\infty}$ monoton wachsend ist, d.h. wenn $f_{n+1}(x) \geq f_n(x)$ für alle $n \in \mathbb{N}$ erfüllt ist.
2. Die monotone Approximation ist ein wichtiger technischer Baustein in der Integrationstheorie. Insbesondere werden wir später eine messbare Funktion f in ihren Positivteil f_+ und Negativteil f_- aufspalten und das Lemma auf beide Teile separat anwenden.
3. Ist f sogar beschränkt, so können wir auf den zweiten (und dritten) Beweisschritt verzichten und $f_n = g_n$ setzen. Insbesondere ist dann die Konvergenz im Theorem nicht nur monoton, sondern sogar gleichmäßig.
4. Der Beweis deckt auch nichtnegative numerische Funktionen ab, wobei jede der approximierenden Funktionen weiterhin einfach ist. Insbesondere gilt $f_n(x) = n$ für alle $x \in X$ mit $f(x) = \infty$.

Vorlesung 04-2 : 18. November

1.6 Integrierbare Funktionen

Vorbemerkung Wir betrachten neben dem messbaren Raum (X, \mathcal{A}) ein Maß μ auf \mathcal{A} und werden schrittweise das entsprechende Integral einführen, indem wir nacheinander

1. nichtnegative einfache,
2. nichtnegative messbare,
3. allgemeine messbare,

Funktionen $f : X \rightarrow \overline{\mathbb{R}}$ diskutieren. Dabei werden wir f in den ersten beiden Fällen immer in sinnvoller Weise ein Integral zuordnen können. Im dritten Fall müssen wir jedoch Bedingungen an f stellen, um gewisse Entartungsfälle auszuschließen.

Bemerkungen:

1. Im Moment ist nicht wichtig, ob μ vollständig und/oder σ -finit ist.
2. Eine wichtige Anwendung in der Analysis ist der Lebesguesche Fall mit $X = \mathbb{R}^d$, $\mathcal{A} = \mathcal{L}$ und $\mu = \lambda$.
3. Vor allem in der stochastischen Analysis werden auch andere Maße und Algebren betrachtet. Zum Beispiel Zählmaße oder Lebesgue-Stieltjes Maße, für die das Maß einer Lebesgue-Menge E durch die Integration einer Dichtefunktion über E gewonnen wird. Siehe dazu auch weiter unten.

maßtheoretische Generalkonvention In der Integrationstheorie vereinbaren wir

$$0 \cdot \infty = 0,$$

wobei wir die Sinnhaftigkeit dieser Formel weiter unten erkennen werden.

Integral nichtnegativer einfacher Funktionen

Definition Für eine einfache und nichtnegative Funktion $f : X \rightarrow \mathbb{R}$ nennen wir

$$I_\mu(f) := \sum_{j=1}^l y_j \mu(E_j)$$

das Integral von f bzgl. μ .

Bemerkungen

1. Im Falle von $(X, \mathcal{A}, \mu) = (\mathbb{R}^d, \mathcal{L}, \lambda)$ sprechen wir von Lebesgue-Integrierbarkeit bzw. vom d -dimensionalen Lebesgue-Integral.
2. Der Wert des Integrals $I_\mu(f)$ hängt nicht von der konkreten Darstellung von f ab, d.h. es ist nicht wichtig, wie viele und welche Zahlen y_j sowie Mengen E_j in der Darstellungsformel für f auftauchen (Übungsaufgabe).

Beispiel: Im eindimensionalen Lebesgue-Fall beschreiben die Formeln

$$f = 2 \chi_{(0,2]} + 3 \chi_{(1,4]} \quad \Longrightarrow \quad I_\lambda(f) = 2 \cdot 2 + 3 \cdot 3 = 13$$

und

$$f = 2 \chi_{(0,1]} + 5 \chi_{(1,2]} + 3 \chi_{(2,4]} \quad \Longrightarrow \quad I_\lambda(f) = 2 \cdot 1 + 5 \cdot 1 + 3 \cdot 2 = 13$$

die Integration derselben einfachen Funktion, wobei die zweite Rechnung ihre Normalform benutzt. Beachte auch, dass

$$I_\lambda(\chi_{(c_1, c_2)}) = I_\lambda(\chi_{[c_1, c_2]}) = I_\lambda(\chi_{[c_1, c_2)}) = I_\mu(\chi_{[c_1, c_2)}) = c_2 - c_1$$

für alle $c_1 \leq c_2$ gilt, denn das Lebesgue-Maß eines Intervalls (bzw. eines eindimensionalen Quaders) hängt nur von seiner Länge ab.

3. Das Integral $I_\mu(f)$ kann den Wert ∞ annehmen, nämlich genau dann wenn $y_j \neq 0$ und $\mu(E_j) = \infty$ für mindestens ein j gilt. Beachte jedoch, dass ein j mit $y_j = 0$ und $\mu(E_j) = \infty$ wegen der Generalkonvention keinen Beitrag zum Integral liefert.
4. Man kann sogar zulassen, dass ein oder mehrere y_j den Wert $+\infty$ annehmen, wobei dann wieder die Generalkonvention zum Tragen kommt: Ein solcher Term liefert einen verschwindenden Beitrag zum Integral für $\mu(E_j) = 0$, andernfalls immer den Beitrag ∞ .

Merkregel: Ein Integral berücksichtigt weder den Wert 0 auf einer Menge vom Maß ∞ noch den Wert ∞ auf einer Menge von Maß 0. Das ist gerade die Motivation für die Generalkonvention.

Bemerkung: Diese Regel gilt auch, wenn ∞ durch eine nichtnegative reelle Zahl ersetzt wird.

5. Das Integral einer verallgemeinerten einfachen Funktion definieren wir durch die entsprechende Reihe, wobei die Nichtnegativität aller Glieder sicherstellt, dass diese immer wohldefiniert ist (mit endlichem oder unendlichem Wert) und die Reihenfolge der Summation beliebig vertauscht werden kann.

Beispiel: Im eindimensionalen Lebesgue-Fall gilt

$$f = \sum_{j=1}^{\infty} j^{-2} \chi_{[j-1, j)} \implies I_{\lambda}(f) = 1 + 1/4 + 1/9 + 1/25 + \dots = \frac{1}{6} \pi^2,$$

sowie

$$f = \sum_{j=1}^{\infty} j^{-1} \chi_{[j-1, j)} \implies I_{\lambda}(f) = 1 + 1/2 + 1/3 + \dots = \infty,$$

6. Die Formel aus der Definition kann, sofern gewisse Nicht-Entartungsbedingungen erfüllt sind, auch für einfache Funktionen mit negativen Werten y_j verwendet werden. Vereinfacht gesagt gilt: Probleme gibt es immer dann, wenn die Terme $-\infty$ und $+\infty$ gleichzeitig als Einzelsummanden auftreten oder durch Addition von anderen Termen erzeugt werden können.

Merkregel: $\infty - \infty$ hat auch in der abstrakten Integrationstheorie keinen Sinn.

7. Die Implikationen

$$I_{\mu}(f) < \infty \implies y_j < \infty \text{ und } \mu(E_j) < \infty \text{ gilt für alle } j = 1, \dots, l$$

und

$$I_{\mu}(f) > 0 \implies \text{es existiert } j = 1, \dots, l \text{ mit } y_j > 0 \text{ und } \mu(E_j) > 0$$

ergeben sich unmittelbar aus der Definition.

8. Durch elementare Abschätzungen zeigen wir

$$0 \leq I_{\mu}(f) \leq \|f\|_{\infty} (\mu(E_1) + \dots + \mu(E_l)), \quad 0 \leq I_{\mu}(f g) \leq \|f\|_{\infty} I_{\mu}(g),$$

$$\text{mit } \|f\|_{\infty} = \sup \{|f(x)| : x \in X\} = \max \{y_1, \dots, y_l\}.$$

Lemma (Monotonie und Linearität des Integrals, Vorversion) Für zwei einfache und nichtnegative Funktionen $f, g : X \rightarrow \mathbb{R}$ gilt³³

$$f \leq g \implies I_{\mu}(f) \leq I_{\mu}(g)$$

sowie

$$I_{\mu}(\alpha f + \beta g) = \alpha I_{\mu}(f) + \beta I_{\mu}(g),$$

wobei α, β zwei reelle Zahlen sind.

Beweis Die Linearität und Monotonie ergeben sich direkt aus den Hausaufgaben zu einfachen Funktionen. \square

Lemma (Satz über die monotone Konvergenz, Version 1) Sei $(f_n)_{n=1}^{\infty}$ eine monoton wachsende Folge einfacher und nichtnegativer Funktionen, die punktweise gegen eine einfache Funktion f konvergiert. Dann gilt

$$I_{\mu}(f) = \lim_{n \rightarrow \infty} I_{\mu}(f_n)$$

im eigentlichen oder uneigentlichen Sinne.

³³Wie üblich meint $f \leq g$ für zwei Funktionen, dass $f(x) \leq g(x)$ für alle $x \in X$ gilt.

Beweis Fall 1: Unter der Annahme $I_\mu(f) = \infty$ existiert (da f einfach ist) ein $F \in \mathcal{A}$ sowie $z \in \mathbb{R}$ mit

$$f \geq z \chi_F, \quad z > 0, \quad \mu(F) = \infty.$$

Die punktweise monotone Konvergenz aus der Voraussetzung impliziert

$$F_1 \subseteq F_2 \subseteq F_3 \subseteq \dots \quad \text{und} \quad F = \bigcup_{n=1}^{\infty} F_n \quad \text{mit} \quad F_n := \left\{ x \in F : f_n(x) \geq \frac{1}{2} z \right\}$$

und damit $\lim_{n \rightarrow \infty} \mu(F_n) = \mu(F) = \infty$. Da andererseits auch $I_\mu(f_n) \geq \frac{1}{2} z \mu(F_n)$ gilt, erhalten wir $\lim_{n \rightarrow \infty} I_\mu(f_n) = \infty$ und damit die Behauptung.

Fall 2: Für $I_\mu(f) < \infty$ betrachten wir die einfachen und nichtnegativen Funktionen $g_n = f - f_n$, sodass für jedes $x \in X$ die Folge $(g_n(x))_{n=1}^{\infty}$ monoton fallend ist und gegen 0 konvergiert. Insbesondere gilt

$$\|g_1\|_\infty \leq \|f\|_\infty < \infty, \quad I_\mu(g_1) \leq I_\mu(f) < \infty$$

und

$$\mu(S) < \infty \quad \text{mit} \quad S := \{x \in X : g_1(x) > 0\}.$$

Wir fixieren nun $\varepsilon > 0$ beliebig und betrachten die Mengen

$$F_n := \{x \in S : 0 \leq g_n(x) < \varepsilon\}, \quad G_n := \{x \in S : g_n(x) \geq \varepsilon\},$$

die zusammen für jedes $n \in \mathbb{N}$ eine disjunkte Zerlegung der Menge S liefern, wobei sich

$$F_1 \subseteq F_2 \subseteq F_3 \subseteq \dots, \quad G_1 \supseteq G_2 \supseteq G_3 \supseteq \dots$$

aus der Monotonie und Konvergenz der g_n ergibt. Wegen

$$\bigcap_{n=1}^{\infty} G_n = \{x \in X : \lim_{n \rightarrow \infty} g_n(x) \geq \varepsilon\} = \emptyset$$

erhalten wir schließlich

$$\lim_{n \rightarrow \infty} \mu(G_n) = 0, \quad \lim_{n \rightarrow \infty} \mu(F_n) = \mu(S).$$

Die Eigenschaften des Integrals einfacher Funktionen implizieren außerdem

$$0 \leq I_\mu(f) - I_\mu(f_n) = I_\mu(g_n) = I_\mu(\chi_{F_n} g_n) + I_\mu(\chi_{G_n} g_n) \leq \varepsilon \mu(F_n) + \|f\|_\infty \mu(G_n)$$

für alle $n \in \mathbb{N}$, wobei die rechte Seite für $n \rightarrow \infty$ gegen die reelle Zahl $\varepsilon \mu(S)$ konvergiert. Da ε beliebig war, folgt die Behauptung. \square

Integral nichtnegativer messbarer Funktionen

Definition Für eine nichtnegative und messbare Funktion $f : X \rightarrow \overline{\mathbb{R}}$ ist ihr Integral bzgl. μ durch

$$I_\mu(f) := \lim_{n \rightarrow \infty} I_\mu(f_n)$$

gegeben, wobei der Wert $+\infty$ zugelassen ist und $(f_n)_{n=1}^{\infty}$ eine monoton wachsende Folge nichtnegativer einfacher Funktionen bezeichnet, die punktweise gegen f konvergiert.

Rechtfertigung Wir hatten im letzten Abschnitt bewiesen, dass für jedes fixierte f immer eine approximierende Folge mit den gewünschten Eigenschaften existiert, wobei die Monotonie (sowohl der Folge als auch des Integralbegriffs) sicherstellt, dass die Zahlenfolge $(I_\mu(f_n))_{n=1}^\infty$ monoton ist und daher von unten gegen einen nichtnegativen Grenzwert $\alpha \in \mathbb{R}$ konvergiert. Wir müssen jedoch noch zeigen, dass dieser nicht von den Details der Approximation abhängt. Wenn $(g_m)_{m=1}^\infty$ eine weitere Approximationsfolge für f ist, setzen wir

$$\beta := \lim_{m \rightarrow \infty} I_\mu(g_m) \in \overline{\mathbb{R}}$$

und betrachten für beliebige $n, m \in \mathbb{N}$ die Funktion

$$h_{n,m} := \min \{ f_n, g_m \},$$

die ebenfalls einfach und nichtnegativ ist. Für festes m ist die Folge $(h_{n,m})_{n=1}^\infty$ monoton wachsend und konvergiert punktweise gegen die einfache Funktion g_m , wobei außerdem

$$h_{n,m} \leq f_n, \quad I_\mu(h_{n,m}) \leq I_\mu(f_n) \leq \alpha$$

für alle n gilt. Das Lemma zur monotonen Konvergenz garantiert daher

$$I_\mu(g_m) \leq \alpha$$

im Limes $n \rightarrow \infty$ und nach einer nochmaligen Anwendung des Lemmas erhalten wir

$$\beta \leq \alpha$$

nach Grenzübergang $m \rightarrow \infty$. Aus Symmetriegründen (bzw. nach einer Argumentation mit vertauschten Rollen) muss aber auch

$$\alpha \leq \beta$$

und damit $\alpha = \beta$ gelten. Insgesamt haben wir bewiesen, dass zwei verschiedene Approximationen von f denselben Wert für $I_\mu(f)$ liefern. \square

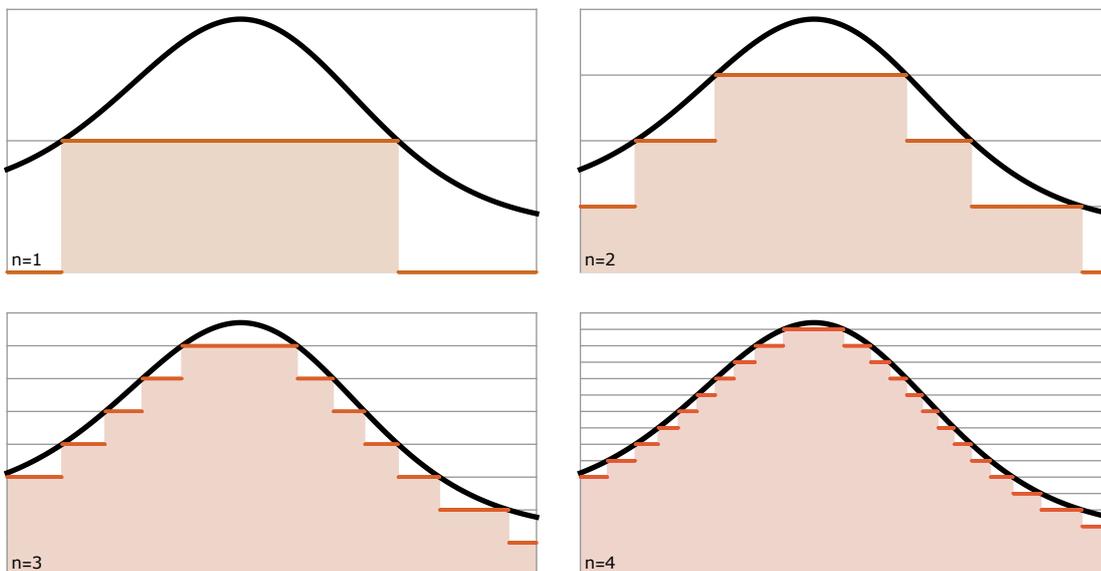


Abbildung In der Definition des Integrals einer nichtnegativen messbaren Funktion (schwarz) wird diese als punktweiser Limes einer monoton wachsenden Folge einfacher Funktionen (rot) betrachtet.

Bemerkung Das Integral nichtnegativer und messbarer Funktionen ist ebenfalls monoton und linear, wobei die Monotonie auch mit der Linearität begründet werden kann (Übungsaufgabe).

Lemma (Satz über die monotone Konvergenz, Version 2) Für jede monotone wachsende Folge $(f_n)_{n=1}^{\infty}$ nichtnegativer messbarer Funktionen $f_n : X \rightarrow \overline{\mathbb{R}}$ gilt

$$I_{\mu}(f) = \lim_{n \rightarrow \infty} I_{\mu}(f_n),$$

wobei $f : X \rightarrow \overline{\mathbb{R}}$ mit $f(x) = \lim_{n \rightarrow \infty} f_n(x)$ den punktweisen Limes bezeichnet.³⁴

Beweis Wir wählen für jedes f_n eine monoton wachsende Folge $(g_{n,m})_{m=1}^{\infty}$ einfacher und nichtnegativer Funktionen, sodass $g_{n,m}$ für $m \rightarrow \infty$ gegen f_n konvergiert. Durch

$$h_k := \max\{g_{n,m} : n \leq k, m \leq k\},$$

wird eine weitere monoton wachsende Folge $(h_k)_{k=1}^{\infty}$ einfacher Funktionen definiert, die punktweise gegen eine Grenzfunktion konvergiert. Da nach Konstruktion

$$g_{n,m}(x) \leq h_k(x) \leq f(x)$$

für alle $x \in X$ und alle $n, m \leq k$ gilt, schließen wir (indem wir erst k , dann m und am Ende auch n gegen ∞ schicken), dass dieser Grenzwert f sein muss und die Definition des Integrals garantiert

$$I_{\mu}(f) = \lim_{k \rightarrow \infty} I_{\mu}(h_k).$$

Andererseits gilt nach Konstruktion und aufgrund der Eigenschaften des Integrals einfacher Funktionen die Formel

$$h_k \leq f_k \leq f, \quad I_{\mu}(h_k) \leq I_{\mu}(f_k) \leq I_{\mu}(f)$$

für alle $k \in \mathbb{N}$. Die Behauptung ergibt sich schließlich im Grenzübergang $k \rightarrow \infty$. \square

Korollar Die logische Äquivalenz

$$I_{\mu}(f) = 0 \quad \iff \quad \mu(\{x \in X : f(x) > 0\}) = 0$$

gilt für jede nichtnegative und messbare Funktion f .

Beweis Sei $S := \{x \in X : f(x) > 0\}$. Wir nehmen zunächst an, dass $I_{\mu}(f) = 0$ erfüllt ist und bemerken für jedes $n \in \mathbb{N}$, dass das Integral der messbaren Funktion

$$g_n := \min\{nf, \chi_S\},$$

verschwindet, denn wegen $0 \leq g_n \leq nf$ ergibt sich $0 \leq I_{\mu}(g_n) \leq n I_{\mu}(f) \leq 0$. Die Folge $(g_n)_{n=1}^{\infty}$ ist andererseits monoton wachsend und konvergiert punktweise gegen χ_S , sodass sich

$$\mu(S) = I_{\mu}(\chi_S) = \lim_{n \rightarrow \infty} I_{\mu}(g_n) = 0$$

³⁴Für jedes $x \in X$ folgt die Existenz und die Nichtnegativität von $f(x)$ direkt aus dem monotonen Wachstum der Zahlenfolge $(f_n(x))_{n=1}^{\infty}$. Die Messbarkeit von f hatten wir bereits im Abschnitt über messbare Funktionen bewiesen.

aus dem Satz über die monotone Konvergenz nichtnegativer Funktionen ergibt. Gilt umgekehrt $\mu(S) = 0$, so definieren wir

$$f_n := \min\{f, n \chi_S\}.$$

Dann gilt $I_\mu(f_n) \leq n I_\mu(\chi(S)) = 0$ für jedes $n \in \mathbb{N}$ und der Satz über die monotone Konvergenz liefert diesmal $I_\mu(f) = \lim_{n \rightarrow \infty} I_\mu(f_n) = 0$. \square

Bemerkung Es gilt sogar die schärfere Aussage

$$I_\mu(f) > 0 \quad \iff \quad \text{es existiert } \varepsilon > 0 \text{ mit } \mu(\{x \in X : f(x) > \varepsilon\}) > 0,$$

wobei die Rückrichtung sofort aus der Monotonie des Integrals folgt und die Hinrichtung sich aus dem Korollar sowie den Hausaufgaben ergibt.

integrierbare Funktionen

Definition Eine messbare Funktion $f : X \rightarrow \overline{\mathbb{R}}$ heißt integrierbar, wenn ihr Positiv- und ihr Negativteil jeweils ein endliches Integral besitzen. Dabei wird

$$I_\mu(f) := I_\mu(f_+) - I_\mu(f_-)$$

das Integral von f bzgl. μ genannt.

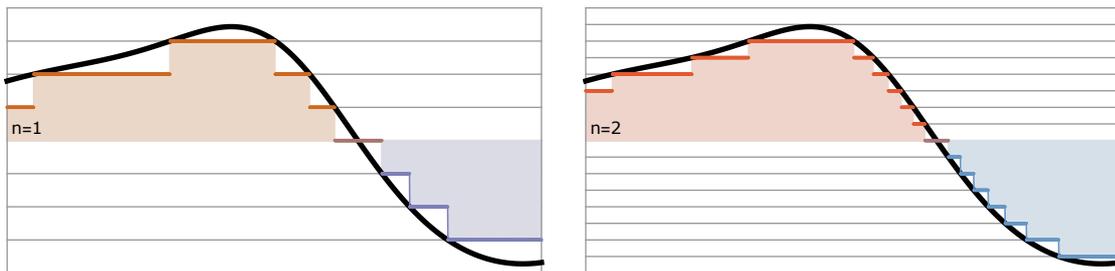


Abbildung Eine messbare Funktion f ist genau dann integrierbar, wenn sowohl ihr Positivteil f_+ als auch ihr Negativteil f_- jeweils ein endliches Integral besitzen. Jedes Teilintegral kann unabhängig vom anderen durch eine monoton wachsende Approximationsfolge einfacher Funktionen ermittelt werden und liefert einen positiven bzw. negativen Beitrag zum Gesamtintegral. Im Bild sind zwei rote bzw. blaue Approximationen von f_+ bzw. f_- dargestellt, wobei die für f_- wie üblich an der horizontalen Achse gespiegelt wurden und daher scheinbar negativ Werte annehmen.

Bemerkungen

1. Jede integrierbare Funktion ist messbar, aber die umgekehrte Aussage ist falsch, da $I_\mu(f_-) = \infty$ oder $I_\mu(f_+) = \infty$ gelten kann.
2. Eine einfache Funktion $f = \sum_{j=1}^l y_j \chi_{E_j}$ ist genau dann integrierbar, wenn $\mu(E_j) < \infty$ für alle $j = 1, \dots, l$ mit $y_j \neq 0$ gilt, wobei sich dann die Formel $I_\mu(f) = \sum_{j=1}^l y_j \mu(E_j)$ ergibt.
3. **Achtung:** Eine nichtnegative Funktion f besitzt zwar immer ein Integral, aber sie wird nur dann integrierbar genannt, wenn $I_\mu(f) < \infty$ gilt.³⁵ Für eine beliebige messbare Funktion f schreiben wir analog auch

$$I_\mu(f) = +\infty \quad \text{falls} \quad I_\mu(f_-) < \infty \quad \text{und} \quad I_\mu(f_+) = \infty$$

³⁵Das ist etwas verwirrend, hat sich aber so eingebürgert.

sowie

$$I_\mu(f) = -\infty \quad \text{falls} \quad I_\mu(f_-) = \infty \quad \text{und} \quad I_\mu(f_+) < \infty,$$

obwohl f in beiden Fällen nicht integrierbar ist. Im Fall von $\infty - \infty$, also für $I_\mu(f_-) = \infty = I_\mu(f_+)$ kann aber $I_\mu(f)$ nicht in sinnvoller Weise definiert werden.

4. Später benutzen wir oftmals die alternativen Notationen

$$\int_X f(x) \, d\mu(x) \quad \text{oder} \quad \int_X f \, d\mu$$

anstelle von $I_\mu(f)$. Ist $E \in \mathcal{A}$ eine messbare Menge, so schreiben wir auch

$$\int_E f(x) \, d\mu(x) = \int_E f \, d\mu = I_\mu(\chi_E f)$$

und meinen damit das Integral der Funktion $\chi_E f$, die auf E mit f übereinstimmt, aber auf dem Komplement E' verschwindet.

5. Im Lebesgueschen Fall $X = \mathbb{R}^d$, $\mathcal{A} = \mathcal{L}$ und $\mu = \lambda$ schreiben wir oftmals auch

$$\int_E f(x) \, dx \quad \text{statt} \quad \int_E f(x) \, d\lambda(x),$$

d.h. wir geben das zugrundeliegende Maß nicht explizit an.

Achtung: Für $d > 1$ ist $x = (x_1, \dots, x_d)$ ein Vektor und E eine Lebesgue-messbare Teilmenge des \mathbb{R}^d . Wir werden später sehen, dass ein d -dimensionales Integral unter gewissen Voraussetzungen als „Verschachtelung“ niedrig-dimensionaler Integrale interpretiert werden kann. Für $d = 2$ und $E = [a_1, b_1] \times [a_2, b_2]$ gilt zum Beispiel

$$\int_E f(x) \, dx = \int_{a_2}^{b_2} \left(\int_{a_1}^{b_1} f(x_1, x_2) \, dx_1 \right) dx_2 = \int_{a_1}^{b_1} \left(\int_{a_2}^{b_2} f(x_1, x_2) \, dx_2 \right) dx_1,$$

d.h. das zweidimensionale Lebesgue-Integral über einem Quader kann mittels eindimensionaler Integrale berechnet werden, wobei die Variable des äußeren Integrals als Parameter im inneren Integral auftaucht.

6. Das Integral einer integrierbaren Funktion ist eine reelle Zahl und nicht etwa ein Element von X oder gar eine Funktion. Insbesondere verallgemeinert unser Integralbegriff das *bestimmte Integral* aus *Analysis 1*, aber nicht das unbestimmte Integral (das sowieso nur im Eindimensionalen existiert).

einige Integralsätze

Lemma (elementare Eigenschaften des Integrals)

1. Eine messbare Funktion f ist genau dann integrierbar, wenn $|f|$ integrierbar ist. Außerdem gilt

$$|I_\mu(f)| \leq I_\mu(|f|).$$

2. Gilt $f = f_2 - f_1$ für zwei nichtnegative und integrierbare Funktionen f_1 und f_2 , so ist auch f integrierbar mit

$$I_\mu(f) = I_\mu(f_2) - I_\mu(f_1).$$

Hierbei sei vorausgesetzt, dass $f_1(x) < \infty$ oder $f_2(x) < \infty$ für jedes $x \in X$ gilt, dass also in keinem Punkt $\infty - \infty$ auftaucht.

3. Seien f, g integrierbar und α, β zwei reelle Zahlen. Dann ist auch $\alpha f + \beta g$ integrierbar mit

$$I_\mu(\alpha f + \beta g) = \alpha I_\mu(f) + \beta I_\mu(g).$$

4. Sind f, g integrierbar, so gilt

$$f \leq g \implies I_\mu(f) \leq I_\mu(g).$$

5. Für zwei integrierbare Funktionen f und g sind auch $\min\{f, g\}$ und $\max\{f, g\}$ integrierbar mit

$$I_\mu(\min\{f, g\}) \leq \min\{I_\mu(f), I_\mu(g)\}, \quad \max\{I_\mu(f), I_\mu(g)\} \leq I_\mu(\max\{f, g\}).$$

6. Ist f integrierbar und $E \in \mathcal{A}$ eine messbare Menge, so ist auch $\chi_E f$ integrierbar.

7. Ist f integrierbar, so gilt

$$\mu(\{x \in X : f(x) = \pm\infty\}) = 0.$$

8. Seien f und g zwei messbare Funktionen mit

$$\mu(\{x \in X : f(x) \neq g(x)\}) = 0.$$

Dann sind f und g entweder beide integrierbar oder beide nicht integrierbar, wobei im ersten Fall $I_\mu(f) = I_\mu(g)$ gilt.

Beweis Teil 1: Ist f integrierbar, so gilt $I_\mu(f_-) < \infty$ sowie $I_\mu(f_+) < \infty$ und die Formel $|f| = f_- + f_+$ garantiert $I_\mu(|f|) = I_\mu(f_-) + I_\mu(f_+) < \infty$. Umgekehrt impliziert $I_\mu(|f|) < \infty$ wegen $0 \leq f_\pm \leq |f|$ auch $I_\mu(f_\pm) < \infty$ und damit die Integrierbarkeit von f . Außerdem gilt stets $|I_\mu(f_+) - I_\mu(f_-)| \leq I_\mu(f_+) + I_\mu(f_-)$.

Teil 2: Punktweise Argumente zeigen, dass die messbare Funktion $g := f_2 - f_+$ nichtnegativ ist und mit den Eigenschaften des Integrals für nichtnegative Funktionen erhalten wir

$$f_1 = f_- + g, \quad f_2 = f_+ + g, \quad I_\mu(f_1) = I_\mu(f_-) + I_\mu(g), \quad I_\mu(f_2) = I_\mu(f_+) + I_\mu(g).$$

Hieraus folgt zum einen $I_\mu(f_-) \leq I_\mu(f_1) < \infty$ und $I_\mu(f_+) \leq I_\mu(f_2) < \infty$, zum anderen aber auch $I_\mu(f_+) - I_\mu(f_-) = I_\mu(f_2) - I_\mu(f_1)$.

Teil 3: Wir betrachten zunächst den Spezialfall $\beta = 0$ und setzen $\tilde{f} := \alpha f$. Für $\alpha = 0$ ist die Behauptung trivial und für $\alpha > 0$ bzw. $\alpha < 0$ ergibt sie sich aus $\tilde{f}_- = \alpha f_-$ und $\tilde{f}_+ = \alpha f_+$ bzw. $\tilde{f}_- = \alpha f_+$ und $\tilde{f}_+ = \alpha f_-$. Analog argumentieren wir für $\alpha = 0$. Als nächstes studieren wir den Spezialfall $\alpha = \beta = 1$. Dann gilt $h := f + g = (f_+ + g_+) - (f_- + g_-)$ und die Behauptung folgt aus dem zweiten Teil dieses Beweises. Im allgemeinen Fall setzen wir $h := \alpha f + \beta g = \tilde{f} + \tilde{g}$ mit $\tilde{f} := \alpha f$, $\tilde{g} := \beta g$ und können die Ergebnisse der Spezialfälle miteinander kombinieren.

Teil 4: Mit $h := g - f \geq 0$ folgt die Behauptung via $I_\mu(g) = I_\mu(f) + I_\mu(h)$ aus der Linearität des Integrals.

Teil 5: Mit einfachen punktweisen Rechnungen verifizieren wir³⁶

$$\min\{f, g\} = \frac{1}{2}(f + g) - \frac{1}{2}|f - g|, \quad \max\{f, g\} = \frac{1}{2}(f + g) + \frac{1}{2}|f - g|$$

und die Integrierbarkeitsaussagen folgen aus dem ersten und dem dritten Teil. Der vierte Teil impliziert außerdem $I_\mu(\min\{f, g\}) \leq I_\mu(f) \leq I_\mu(\max\{f, g\})$ sowie analoge Abschätzungen für $I_\mu(g)$ und die Ungleichungen für die Integrale ergeben sich aus elementaren Umformungen.

Teil 6: Dies folgt via

$$|\chi_E f| \leq \chi_E |f| \leq f$$

aus dem ersten und dem vierten Teil.

Teil 7: Sei $E_+ := f^{-1}(\{+\infty\})$. Dann gilt $f_+ \geq n \chi_{E_+}$ für alle $n \in \mathbb{N}$ und damit auch $\infty > I_\mu(f_+) \geq n \mu(E_+)$ bzw.

$$0 \leq \mu(E_+) \leq I_\mu(f_+)/n.$$

Der Grenzübergang $n \rightarrow \infty$ liefert die Behauptung und analog zeigen wir die Gültigkeit von $\mu(E_-) = 0$ mit $E_- := f^{-1}(\{-\infty\})$.

Teil 8: Wir setzen $h := f - g$ und bemerken, dass $\{x \in X : |h|(x) > 0\} = \emptyset$ gilt. Das Korollar zum Satz über die monotone Konvergenz nichtnegativer Funktionen liefert daher $I_\mu(|h|) = 0$ und der erste Teil garantiert, dass h integrierbar ist mit $I_\mu(h) = 0$. Die Behauptungen ergeben sich nun via $f = g + h$ bzw. $g = f - h$ aus dem dritten Teil. \square

Bemerkungen

1. Die dritte bzw. vierte Aussage garantiert die Linearität bzw. Monotonie des Integrals.
2. Die Menge aller integrierbaren Funktionen ist in natürlicher Weise ein reeller Vektorraum, den wir mit $\mathcal{L}_\mu^1(X)$ bezeichnen.

³⁶Diese Formeln sind auch in anderen Kontexten sehr nützlich.

Lemma (Satz über die monotone Konvergenz, Version 3) Sei $(f_n)_{n=1}^\infty$ eine monoton wachsende Folge integrierbarer Funktionen $f_n : X \rightarrow \overline{\mathbb{R}}$, sodass

$$I_\mu(f_n) \leq c$$

für eine geeignete Konstante $c \in \mathbb{R}$ sowie alle $n \in \mathbb{N}$ erfüllt ist. Dann ist der punktweise Limes $f : X \rightarrow \overline{\mathbb{R}}$ ebenfalls integrierbar und es gilt $I_\mu(f) = \lim_{n \rightarrow \infty} I_\mu(f_n)$.

Beweis O.B.d.A können wir $f_1 \geq 0$ annehmen, denn andernfalls betrachten wir die Funktionen $\tilde{f}_n := f_n - f_1$ sowie $\tilde{c} := c - I_\mu(f_1) \in \mathbb{R}$. Die vorherige Version dieses Lemmas garantiert nun die gewünschte Grenzwertformel sowie $I_\mu(f) \leq c < \infty$. Wegen $f_- = 0$ und $f_+ = f$ folgt auch die Integrierbarkeit von f . \square

Bemerkungen

1. Die optimale Konstante ist durch $c = \lim_{n \rightarrow \infty} I_\mu(f_n)$ gegeben, wobei dieser Grenzwert wegen der Monotonie existiert. Beachte auch, dass c im Lemma endlich (aber nicht unbedingt positiv) sein muss.
2. Das Lemma wird auch Satz von Beppo-Levi genannt.
3. Ein analoge Aussage gilt auch für monoton fallende Folgen integrierbarer Funktionen, wobei dann $I_\mu(f_n) \geq c$ für alle $n \in \mathbb{N}$ zu fordern ist.
4. Durch Betrachtung von $g_n = |f_n - f| = f - f_n$ können wir

$$\lim_{n \rightarrow \infty} I_\mu(|f_n - f|) = 0$$

zeigen, wobei dies weiter unten die Konvergenz in einer Halbnorm liefern wird.

Korollar (nützliches Integrierbarkeitskriterium) Sei $(E_n)_{n=1}^\infty$ eine paarweise disjunkte bzw. eine aufsteigende Folge von Mengen $E_n \in \mathcal{A}$ mit $X = \bigcup_{n=1}^\infty E_n$ und sei $f : X \rightarrow \overline{\mathbb{R}}$ eine messbare Funktion mit

$$\sum_{n=1}^{\infty} I_\mu(\chi_{E_n} |f|) < \infty \quad \text{bzw.} \quad \sup \{ I_\mu(\chi_{E_n} |f|) : n \in \mathbb{N} \} < \infty.$$

Dann ist f integrierbar.

Beweis Siehe die Hausaufgaben. \square

Korollar (Lemma von Fatou) für jede Folge $(f_n)_{n=1}^\infty$ nichtnegativer messbarer Funktionen $f_n : X \rightarrow \overline{\mathbb{R}}$ gilt

$$I_\mu\left(\liminf_{n \rightarrow \infty} f_n\right) \leq \liminf_{n \rightarrow \infty} I_\mu(f_n),$$

wobei beide Seiten auch den Wert ∞ annehmen können.³⁷

Beweis Siehe die Hausaufgaben. \square

³⁷Der einzig interessante Fall ist aber, dass die rechte Seite endlich ist, denn dies impliziert die Integrierbarkeit der Funktion $\liminf_{n \rightarrow \infty} f_n$.

1.7 Spezialfälle und Beispiele

Maße mit Dichten Ist $w : X \rightarrow \overline{\mathbb{R}}$ eine nichtnegative und integrierbare Funktion,³⁸ so wird durch

$$\mu_w(E) := I_\mu(\chi_E w)$$

ein weiteres Maß auf \mathcal{A} erzeugt, wobei w die Dichtefunktion genannt wird und wir im Lebesgueschen Fall vom Lebesgue-Stieltjes-Maß zu w sprechen. Für eine beschränkte Dichte w können wir (siehe die Hausaufgaben) außerdem zeigen, dass jede μ -integrierbare Funktion auch μ_w -integrierbar ist, wobei beide Integralbegriffe durch

$$I_{\mu_w}(f) = I_\mu(f w)$$

gekoppelt sind.³⁹

Beispiel: Die Funktion $w : \mathbb{R} \rightarrow \mathbb{R}$ mit

$$w(x) := \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$$

ist die Wahrscheinlichkeitsdichte einer normalisierten Gaußverteilung und genügt den Bedingungen

$$\int_{-\infty}^{+\infty} w(x) dx = 1, \quad \int_{-\infty}^{+\infty} x w(x) dx = 0, \quad \int_{-\infty}^{+\infty} x^2 w(x) dx = 1,$$

wobei die letzten beiden sicherstellen, dass eine entsprechende Zufallsgröße Ξ den Erwartungswert $\mathbf{E}(\Xi) = 0$ sowie die Varianz $\mathbf{Var}(\Xi) = 1$ besitzt. Das Maß λ_w liefert via

$$\lambda_w(M) = \mathbf{Prob}(\Xi \in M)$$

für jede Lebesgue-Menge $M \in \mathcal{L}$ die Wahrscheinlichkeit, dass Ξ einen Wert aus M annimmt. Für jede Lebesgue-integrierbare Funktion f gilt nun

$$I_{\lambda_w}(f) = \int_{-\infty}^{+\infty} f(x) w(x) dx = \mathbf{E}(f(\Xi)),$$

d.h. durch Integration von f bzgl. des Lebesgue-Stieltjes-Maßes λ_w können wir den Erwartungswert der Zufallsgröße $f(\Xi)$ berechnen.⁴⁰

Bemerkung: Das Maß μ_w ist absolut-stetig bzgl. μ , d.h. die Implikation

$$\mu(E) = 0 \quad \implies \quad \mu_w(E) = 0$$

³⁸Als Verallgemeinerung können wir sogar integrierbare Funktionen w zulassen, die auch negative Werte annehmen. Dann wird μ_w allerdings ein *signiertes Maß* sein und es kann $\mu_w(E) < 0$ gelten.

³⁹Wenn w nicht beschränkt ist, müssen wir etwas sorgfältiger formulieren. Insbesondere kann es dann Funktionen geben, die zwar bzgl. μ , aber nicht bzgl. μ_w integrierbar sind.

⁴⁰In der Stochastik wird X oft als Ω geschrieben und Ξ als X bezeichnet.

gilt für jedes $E \in \mathcal{A}$.

Ausblick*: Der Satz von Radon-Nikodým garantiert für ein σ -finites Maß μ , dass jedes andere Maß ν auf \mathcal{A} genau dann via $\nu = \mu_w$ durch eine geeignete Dichte w dargestellt werden kann, wenn ν absolut-stetig bzgl. μ ist. Ein Beweis findet sich zum Beispiel in [Bau, Abschnitt II.17]. Man nennt w dann auch die Radon-Nikodým-Ableitung von ν nach μ , denn symbolisch gilt

$$d\nu(x) = w(x) d\mu(x)$$

im Sinne einer Transformationsformel für Integrale (siehe dazu weiter unten). Beachte aber, dass diese Formel nicht im Sinne klassischer Ableitungen gilt (μ und ν sind ja Maße und keine Funktionen).

induzierte Maße und Integrale auf messbaren Teilmengen Für jedes $M \in \mathcal{A}$ wird durch

$$\mathcal{A}_M := \{M \cap E : E \in \mathcal{A}\}, \quad \mu_M(E \cap M) := \mu(E)$$

eine σ -Algebra $\mathcal{A}_M \subseteq \mathcal{P}(M)$ sowie ein Maß μ_M auf \mathcal{A}_M definiert, wobei wir dieses als das von μ auf M induzierte Maß bezeichnen. Insbesondere können wir zeigen, dass für jede μ_M -integrierbare Funktion $f : M \rightarrow \overline{\mathbb{R}}$ ihre triviale Fortsetzung

$$\tilde{f}(x) := \begin{cases} f(x) & \text{falls } x \in M \\ 0 & \text{sonst} \end{cases}$$

μ -integrierbar ist und das umgekehrt für jede μ -integrierbare Funktion $g : X \rightarrow \overline{\mathbb{R}}$ ihre Einschränkung auf M integrierbar bzgl. μ_M ist. Es gilt also

$$I_{\mu_M}(f) = I_{\mu}(\tilde{f}), \quad I_{\mu_M}(g|_M) = I_{\mu}(\chi_M g).$$

Bemerkungen:

1. Salopp gesprochen kann man sagen: Wenn wir Funktionen auf der Menge X integrieren können, so können wir auch Funktionen integrieren, die nur auf einer Teilmenge M von X definiert sind.
2. Wenn wir Funktionen auf M mit ihrer trivialen Fortsetzung auf X identifizieren, so entspricht μ_M (Menge als Index) gerade dem Maß μ_{χ_M} (Funktion als Index). Das von μ auf einer Menge M induzierte Maß/Integral kann also alternativ auch als das Maß/Integral betrachtet werden, dass der Dichtefunktion χ_M entspricht. In der Literatur werden beide Sichtweisen verwendet.

Beispiel: Die Quintessenz von induzierten Maßen und ihren Integralen kann aus der eindimensionalen Formel

$$\int_a^b f(x) dx = \int_{-\infty}^{+\infty} \chi_{[a,b]}(x) f(x) dx$$

abgelesen werden. Links lebt alles auf der Menge $M = [a, b] \subseteq X$, rechts alles auf $X = \mathbb{R}$.

Integrale zu Zählmaßen Als prototypisches Beispiel betrachten wir den Maßraum

$$X = \mathbb{R}, \quad \mathcal{A} = \mathcal{P}(\mathbb{R}), \quad \mu(E) = \#(E \cap \mathbb{N}),$$

wobei $\mu(E)$ für jedes $E \subseteq \mathbb{R}$ gerade die natürlichen Zahlen innerhalb von E zählt. Insbesondere ist bei diesem Zählmaß jede Teilmenge von \mathbb{R} und jede numerische Funktion auf \mathbb{R} messbar. Für jede nichtnegative Funktion $f : \mathbb{R} \rightarrow \overline{\mathbb{R}}$ ergibt sich dabei

$$I_\mu(f) = \sum_{n=1}^{\infty} f(n),$$

wobei alle Summanden auf der rechten Seite nichtnegativ sind und daher nur einer der beiden folgenden zwei Fälle eintreten kann:

1. Die Partialsummenfolge der Reihe konvergiert im eigentlichen Sinne gegen eine reelle Zahl und f ist integrierbar mit $I_\mu(f) < \infty$.
2. Die Partialsummenfolge der Reihe divergiert bzw. konvergiert im uneigentlichen Sinne. Dann gilt $I_\mu(f) = \infty$ und f ist nicht integrierbar.

Analoge Aussagen gelten im nichtpositiven Fall, aber wenn die Funktion f sowohl negative als auch positive Werte annimmt, müssen wir die obige Diskussion auf f_- , f_+ anwenden und erhalten das folgende Integrierbarkeitskriterium.

Theorem: Eine Funktion $f : \mathbb{R} \rightarrow \overline{\mathbb{R}}$ ist genau dann integrierbar bzgl. des Zählmaßes μ , wenn die Reihe der $f(n)$ *absolut konvergiert*, d.h. wenn $\sum_{n=1}^{\infty} |f(n)| < \infty$ gilt.

Bemerkungen:

1. Das Integral einer integrierbaren Funktion ist weiterhin durch die unendliche Summe der $f(n)$ gegeben, wohingegen die unendliche Summe der $|f(n)|$ das Integral von $|f|$ liefert.
2. Es gilt $|\sum_{n=1}^{\infty} f(n)| \leq \sum_{n=1}^{\infty} |f(n)| < \infty$, wobei dies mit der abstrakten Integrationstheorie begründet oder alternativ wie in *Analysis 1* mittels der entsprechenden Relation für die Partialsummen hergeleitet werden kann.
3. Wenn die Reihe der $f(n)$ konvergiert, aber nicht absolut konvergiert, so ist f *nicht* integrierbar bzgl. μ . Siehe dazu das Paradox der Umordnung.
4. Analoge Aussagen gelten für viele andere Zählmaße.
5. Ausblick*: Man kann das Konzept eines Zählmaßes auch mit der Idee einer Dichte w zusammenbringen, sofern man *verallgemeinerte Funktionen* bzw. sogenannte *Distributionen* zulässt. Im konkreten Fall gilt zum Beispiel

$$w(x) = \sum_{n=1}^{\infty} \delta_n(x),$$

wobei δ_n die Diracsche Delta-Distribution im Punkt n ist. Zählmaße können auch als singuläre Grenzwerte von Maßen mit regulären Dichten betrachtet werden.

Zum Beispiel gilt

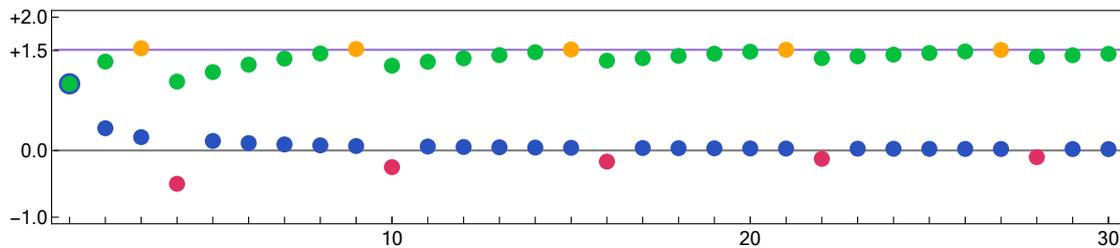
$$\delta_n(x) \approx \frac{\exp\left(-\frac{(x-n)^2}{\varepsilon}\right)}{\sqrt{\varepsilon\pi}}$$

für alle kleinen $\varepsilon > 0$, wobei wir die nichttriviale Frage, in welchem Sinne Maße und Distributionen konvergieren, hier nicht diskutieren können.

Beispiel: Für jeden Parameter $s > 0$ betrachten wir die Funktion $f_s : \mathbb{R} \rightarrow \overline{\mathbb{R}}$ mit

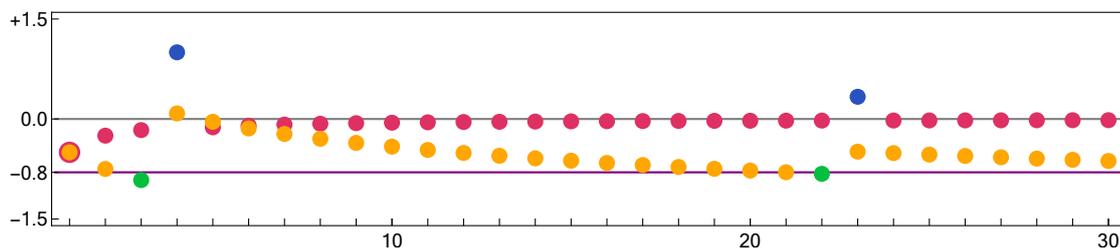
$$f_s(x) = \frac{\cos(\pi x)}{x^s} \quad \text{für } x > 0, \quad f_s(x) = 0 \quad \text{für } x \leq 0$$

und erhalten $f(n) = (-1)^n n^{-s}$ für alle $n \in \mathbb{N}$. Für $s > 1$ ist f_s integrierbar, aber dies gilt nicht für $0 < s \leq 1$, obwohl die Reihe der $f_s(n)$ nach dem Leibniz-Kriterium weiterhin konvergiert (aber eben nicht mehr absolut).



$$(y_{\varphi(n)})_{n \in \mathbb{N}} = \left(\frac{1}{1}, \frac{1}{3}, \frac{1}{5}, \frac{-1}{2}, \frac{1}{7}, \frac{1}{9}, \frac{1}{11}, \frac{1}{13}, \frac{1}{15}, \frac{-1}{4}, \frac{1}{17}, \frac{1}{19}, \frac{1}{21}, \frac{1}{23}, \frac{1}{25}, \frac{-1}{6}, \frac{1}{27}, \frac{1}{29}, \frac{1}{31}, \frac{1}{33}, \frac{1}{35}, \frac{-1}{8}, \frac{1}{37}, \frac{1}{39}, \frac{1}{41}, \frac{1}{43}, \frac{1}{45}, \frac{-1}{10}, \frac{1}{47}, \dots \right)$$

Abbildung Eine Umordnung der alternierenden harmonischen Folge $y_n = (-1)^{n-1}/n$, sodass $\sum_{n=1}^{\infty} y_{\varphi(n)} = r$ mit $r = 1.5$ gilt. Die positiven bzw. negativen Folgenglieder entsprechen den blauen bzw. roten Punkten und die grünen bzw. orangenen Punkte repräsentieren die Partialsummen, die kleiner bzw. größer als r sind. Erkennen Sie die Konstruktionsidee? Sie besteht — salopp gesprochen — darin, die gegebenen positiven und negativen Bausteine ihrer Größe nach übereinander zu stapeln, wobei bei jedem Über- bzw. Unterschreiten der vorgegebenen Schranke r die Sorte der Bausteine gewechselt wird.



$$(y_{\varphi(n)})_{n \in \mathbb{N}} = \left(\frac{-1}{2}, \frac{-1}{4}, \frac{-1}{6}, \frac{1}{1}, \frac{-1}{8}, \frac{-1}{10}, \frac{-1}{12}, \frac{-1}{14}, \frac{-1}{16}, \frac{-1}{18}, \frac{-1}{20}, \frac{-1}{22}, \frac{-1}{24}, \frac{-1}{26}, \frac{-1}{28}, \frac{-1}{30}, \frac{-1}{32}, \frac{-1}{34}, \frac{-1}{36}, \frac{-1}{38}, \frac{-1}{40}, \frac{-1}{42}, \frac{1}{3}, \frac{-1}{44}, \frac{-1}{46}, \dots \right)$$

Abbildung Die Umordnung für $r = -0.8$.

Paradox der Umordnung: Der *Umordnungssatz von Riemann* besagt, dass für jede reelle Zahlenfolge $(y_n)_{n=1}^{\infty}$ mit

$$\sum_{n=1}^{\infty} \max\{y_n, 0\} = +\infty, \quad \sum_{n=1}^{\infty} \min\{y_n, 0\} = -\infty$$

und jede (!) reelle Zahl $r \in \mathbb{R}$ eine bijektive Abbildung $\varphi : \mathbb{N} \rightarrow \mathbb{N}$ (die Umordnung der Indizes) existiert, sodass

$$\sum_{n=1}^{\infty} y_{\varphi(n)} = r.$$

Das Standardbeispiel ist die alternierende harmonische Folge und der Beweis dieser verblüffenden Aussage ist erstaunlich einfach und sogar konstruktiv (siehe dazu das Bild). Im Kontext des Zählmaßes μ meint dies, dass es für jede nicht integrierbare Funktion f mit $I_{\mu}(f_-) = I_{\mu}(f_+) = \infty$ und jeden vorgeschriebenen Wert r eine Folge von integrierbaren Funktionen $(f_m)_{m=1}^{\infty}$ gibt, sodass im Limes $m \rightarrow \infty$ die f_m punktweise gegen f und die Integrale der f_m gegen r konvergieren.

Zusatz

Lebesgue- versus Riemann-Integrierbarkeit in einer Dimension Sei $[a, b]$ ein endliches Intervall und sei $f : [a, b] \rightarrow \mathbb{R}$ eine Riemann-integrierbare (und damit auch beschränkte) Funktion (siehe dazu *Analysis 1*). Wir nehmen der Einfachheit halber an, dass f nichtnegativ ist, und setzen f via

$$\tilde{f}(x) := \begin{cases} f(x) & \text{falls } a < x \leq b \\ 0 & \text{sonst} \end{cases}$$

in trivialer Weise zu einer Funktion $\tilde{f} : \mathbb{R} \rightarrow \mathbb{R}$ fort. Für jedes $n \in \mathbb{N}$ sind

$$x_{n,j} = a + \frac{j}{2^n}(b-a) \quad \text{mit} \quad j = 0, \dots, 2^n$$

die Stützstellen einer äquidistanten Zerlegung des Intervalls $[a, b]$ in 2^n abgeschlossene Teilintervalle der Länge $(b-a)/2^n$, wobei beim Übergang von n zu $n+1$ jedes dieser Teilintervalle wiederum in zwei gleich große Teile zerlegt wird. Die Zahlen

$$\underline{S}_n := \frac{1}{2^n} \sum_{j=1}^{2^n} \underline{m}_{n,j} \quad \text{mit} \quad \underline{m}_{n,j} := \inf \{ f(x) : x \in [x_{n,j-1}, x_{n,j}] \}$$

bzw.

$$\overline{S}_n := \frac{1}{2^n} \sum_{j=1}^{2^n} \overline{m}_{n,j} \quad \text{mit} \quad \overline{m}_{n,j} := \sup \{ f(x) : x \in [x_{n,j-1}, x_{n,j}] \}$$

sind wegen der Beschränktheit von f wohldefiniert und liefern die entsprechende Unter- bzw. Obersumme von f (siehe wieder *Analysis 1*). Die Riemann-Integrierbarkeit von f garantiert

$$\lim_{n \rightarrow \infty} \underline{S}_n = \int_a^b f(x) dx = \lim_{n \rightarrow \infty} \overline{S}_n,$$

wobei in der Mitte das Riemann-Integral von f steht. Wir betrachten nun die einfachen Funktionen

$$\underline{f}_n := \sum_{j=1}^{2^n} \underline{m}_{n,j} \chi_{(x_{n,j-1}, x_{n,j}]}, \quad \overline{f}_n := \sum_{j=1}^{2^n} \overline{m}_{n,j} \chi_{(x_{n,j-1}, x_{n,j}]}$$

und bemerken, dass

$$I_\lambda(\underline{f}_n) = \underline{S}_n, \quad I_\lambda(\overline{f}_n) = \overline{S}_n$$

und damit auch $\lim_{n \rightarrow \infty} I_\mu(\underline{f}) = \int_a^b f(x) \, dx = \lim_{n \rightarrow \infty} I_\mu(\overline{f}_n)$ gilt. Unsere Definitionen implizieren (siehe die nachfolgende Abbildung) außerdem die Monotonie-Aussagen

$$\underline{f}_n \leq \underline{f}_{n+1} \leq \overline{f}_{n+1} \leq \overline{f}_n$$

für alle $n \in \mathbb{N}$ und deshalb existieren die punktweisen Grenzwerte

$$\underline{f}(x) := \lim_{n \rightarrow \infty} \underline{f}_n(x), \quad \overline{f}(x) := \lim_{n \rightarrow \infty} \overline{f}_n(x),$$

wobei unsere Konstruktion

$$\underline{f}(x) \leq f(x) \leq \overline{f}(x) \quad \text{für alle } x \in (a, b]$$

sicherstellt. Nach zweimaliger Anwendung des Satzes über die monotone Konvergenz — einmal auf die \underline{f}_n , das andere Mal auf die \overline{f}_n — erhalten wir

$$I_\lambda(\underline{f}) = I_\lambda(\overline{f}) = \int_a^b f(x) \, dx.$$

Insbesondere verschwindet das Integral der nichtnegativen Funktion $\overline{f} - \underline{f}$ und wir schließen insgesamt, dass eine Lebesgue-Nullmenge N existiert, sodass

$$\underline{f}(x) = \tilde{f}(x) = \overline{f}(x) \quad \text{für alle } x \notin N.$$

Beachte auch, dass $\underline{f}(x) = \tilde{f}(x) = \overline{f}(x) = 0$ für alle $x < a$ und alle $x > b$ nach Konstruktion gilt, d.h. N ist eine Teilmenge von $[a, b]$, und dass die Vollständigkeit von λ sowohl die Messbarkeit als auch die Integrierbarkeit von \tilde{f} impliziert.⁴¹

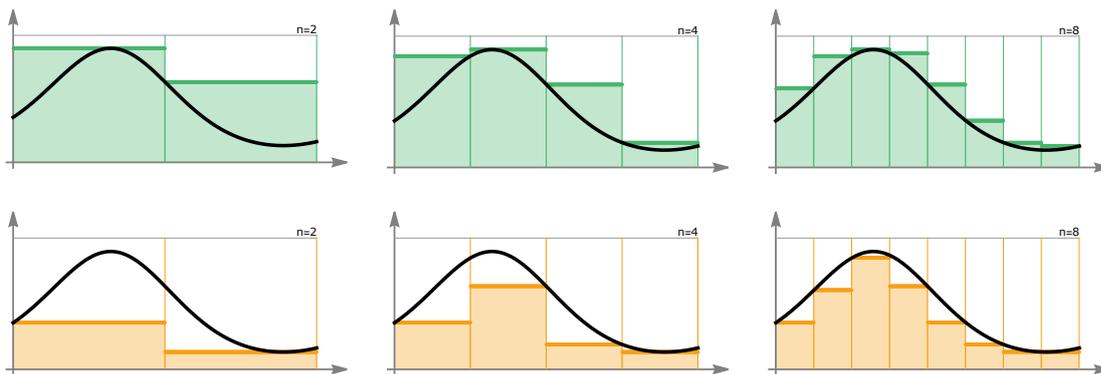


Abbildung Jede Riemann-integrierbare Funktion (schwarz) kann von *oben* (grün) bzw. *unten* (gelb) durch eine monoton *fallende* bzw. *wachsende* Folge einfacher Funktionen approximiert werden, wobei diese mittels spezieller Zerlegungen des Definitionsbereiches sowie den jeweiligen Ober- bzw. Untersummen konstruiert werden. Insbesondere liefern die Lebesguesche und die Riemannsches Theorie denselben Wert des Integrals. Es gibt aber Funktionen, die zwar Lebesgue-, aber nicht Riemann-integrierbar sind.

⁴¹Die Messbarkeit von $h := \overline{f} - \tilde{f}$ folgt aus der Tatsache, dass die Menge $\{x \in X : h(x) \neq 0\}$ eine Teilmenge von N und damit selbst messbar ist. Die Funktion $\tilde{f} = h + \overline{f}$ ist also als Summe messbarer Funktionen auch messbar. Die Integrierbarkeit von \tilde{f} ergibt sich außerdem nach dem achten Teil des Lemmas zu den elementaren Eigenschaften aus der Integrierbarkeit von \overline{f} .

Zusammenfassung: Jede Riemann-integrierbare Funktion ist nach trivialer Fortsetzung auch Lebesgue-integrierbar, wobei das Riemann-Integral und das Lebesgue-Integral denselben Wert annehmen.

Bemerkungen:

1. In diesem Sinne ist das eindimensionale Lebesgue-Integral eine Verallgemeinerung des Riemannsches Integralbegriffs. Eine analoge Aussage gilt für $d > 1$.
2. Wir können auf die triviale Fortsetzung (also den Übergang von f auf \tilde{f}) verzichten, sofern wir nicht mit dem Lebesgue-Maß auf \mathbb{R} , sondern mit dem induzierten Maß auf $[a, b]$ arbeiten.
3. Wir haben \underline{f} und \overline{f} mithilfe sehr spezieller Ober- und Untersummen konstruiert (äquidistante Zerlegung sowie eine Zweierpotenz als Anzahl der Teilintervalle) und brauchten daher als technisches Hilfsmittel nur den Satz über die monotone Konvergenz. Mit den allgemeineren Integralsätzen weiter unten können aber auch beliebige Riemann-Summen behandelt werden.
4. Ist f sogar stetig auf $[a, b]$, so gilt $\underline{f}(x) = f(x) = \overline{f}(x)$ für alle $x \in (a, b)$. Die obige Konstruktion wird zwar im Allgemeinen immer noch $0 = \underline{f}(a) = \overline{f}(a) \neq f(a)$ liefern, aber durch kleine Modifikationen in der Argumentation kann dieser Defekt behoben werden.
5. Die charakteristische Funktion der Dirichlet-Menge ist Lebesgue-integrierbar, aber nicht Riemann-integrierbar, denn sie kann nicht durch Treppenfunktionen approximiert werden.

Achtung: Es gibt auch Funktionen, die nicht Lebesgue-, aber Riemann-integrierbar im *uneigentlichen Sinne* sind. Ein klassisches Beispiel ist der *Sinus cardinalis*, also die Funktion $\text{sinc} : \mathbb{R} \rightarrow \mathbb{R}$ mit

$$\text{sinc}(x) := \frac{\sin(x)}{x},$$

wobei $\text{sinc}(0) = \lim_{\varepsilon \rightarrow 0} \text{sinc}(\varepsilon) = 1$ vereinbart sei. Es gilt

$$\int_{-\xi}^0 \text{sinc}(x) dx = \int_0^{+\xi} \text{sinc}(x) dx \xrightarrow{\xi \rightarrow \infty} \frac{1}{2} \pi,$$

aber auch

$$\int_{-\xi}^{+\xi} |\text{sinc}(x)| dx \xrightarrow{\xi \rightarrow \infty} \infty,$$

wobei beide Aussagen in den Hausaufgaben diskutiert werden.

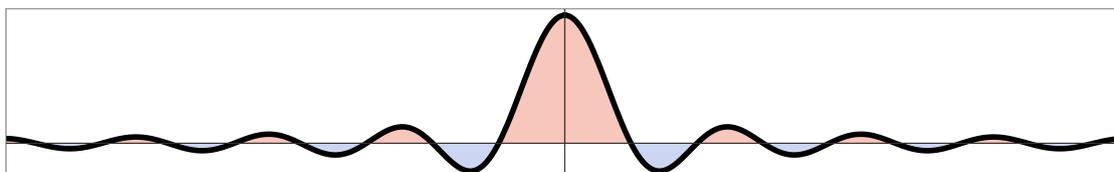


Abbildung Der Sinus cardinalis ist auf ganz \mathbb{R} nicht Lebesgue-integrierbar, da sich die positiven (rot) und negativen (blau) Flächen zwischen seinem Graphen und der horizontalen Achse jeweils zu ∞ addieren. Er ist aber uneigentlich Riemann-integrierbar.

1.8 Weitere Integralsätze

Annahme Wir wollen im Folgenden immer voraussetzen, dass das Maß μ vollständig ist, obwohl viele der folgenden Aussagen auch ohne diese Annahme gelten, sofern die konkrete Formulierung etwas angepasst wird.

Sprechweise Wir sagen, eine punktweise Eigenschaft gelte für fast alle $x \in X$ (oder kürzer: fast überall), wenn sie höchstens auf einer Menge vom Maß 0 nicht erfüllt ist.

Beispiele:

1. Ist f integrierbar, so gilt $f(x) < \infty$ für fast alle $x \in X$.
2. Das Integral einer nichtnegativen und messbaren Funktion $f(x)$ verschwindet genau dann, wenn $f(x) = 0$ für fast alle $x \in X$ gilt (bzw. wenn $f = 0$ fast überall).
3. Sind $f, g : X \rightarrow \mathbb{R}$ zwei integrierbare Funktionen, die fast überall gleich sind (d.h. $f(x) = g(x)$ für fast alle $x \in X$), so gilt $I_\mu(f) = I_\mu(g)$.

Bemerkung: Da μ vollständig ist, gelten die folgenden Aussagen für je zwei Funktionen $f, g : X \rightarrow \overline{\mathbb{R}}$, die fast überall gleich sind:

1. f ist genau dann messbar, wenn g messbar ist.
2. f ist genau dann integrierbar, wenn g integrierbar ist.

Bei nicht-vollständigen Mäßen muss die erste Aussage nicht richtig sein und dies impliziert, dass auch die zweite im Allgemeinen falsch ist (denn jede integrierbare Funktion ist ja insbesondere auch messbar).

1-Halbnorm bzgl. des Maßes μ Für jede messbare Funktion $f : X \rightarrow \overline{\mathbb{R}}$ ist

$$\|f\|_{1,\mu} := I_\mu(|f|) \in [0, \infty]$$

wohldefiniert, wobei f genau dann integrierbar ist, wenn $\|f\|_{1,\mu} < \infty$ gilt. Aus dem Lemma über die elementaren Eigenschaften des Integrals können wir die folgenden Aussagen herleiten:

1. $\|f\|_{1,\mu} \geq 0$,
2. $\|\alpha f\|_{1,\mu} = |\alpha| \|f\|_{1,\mu}$,
3. $\|f + g\|_{1,\mu} \leq \|f\|_{1,\mu} + \|g\|_{1,\mu}$.

Achtung: $\|\cdot\|_{1,\mu}$ ist keine Norm auf $\mathcal{L}_\mu^1(X)$, sondern nur eine Halbnorm. Aus $\|f\|_{1,\mu} = 0$ folgt nämlich nicht $f = 0$, sondern nur $f = 0$ fast überall. Siehe dazu auch die Definition der Lebesgue-Räume weiter unten, wobei dort auch das Analogon $\|\cdot\|_{p,\mu}$ für $p \in (1, \infty]$ erklärt wird.

Theorem (Satz über die majorisierte Konvergenz) Sei $(f_n)_{n=1}^\infty$ eine Folge von integrierbaren Funktionen $f_n : X \rightarrow \overline{\mathbb{R}}$, die punktweise fast überall gegen eine Funktion $f : X \rightarrow \overline{\mathbb{R}}$ konvergiert. Außerdem sei $h : X \rightarrow \overline{\mathbb{R}}$ eine integrierbare Funktion, sodass

$$|f_n(x)| \leq h(x)$$

für alle $n \in \mathbb{N}$ und jedes $x \in X$ erfüllt ist. Dann gilt

$$\|f_n - f\|_{1,\mu} \xrightarrow{n \rightarrow \infty} 0, \quad I_\mu(f_n) \xrightarrow{n \rightarrow \infty} I_\mu(f)$$

und f ist insbesondere integrierbar.

Beweis Spezialfall, Teil 1: Wir nehmen zunächst an, dass $f(x) = \lim_{n \rightarrow \infty} f_n(x)$ für jedes x in X erfüllt ist und betrachten für alle $m, k \in \mathbb{N}$ die messbaren Funktionen

$$g_{m,k} := \min \{f_n : m \leq n \leq m+k\}, \quad g_m := \inf \{f_n : n \geq m\},$$

wobei $g_{m,k}$ als Minimum endlich vieler integrierbarer Funktionen selbst integrierbar ist (siehe das Lemma oben). Nach Konstruktion gilt außerdem

$$f(x) = \liminf_{n \rightarrow \infty} f_n(x) = \lim_{m \rightarrow \infty} g_m(x), \quad g_m(x) = \lim_{k \rightarrow \infty} g_{m,k}(x)$$

für jedes $x \in X$ sowie

$$-h \leq g_m \leq g_{m,k+1} \leq g_{m,k} \leq +h, \quad -h \leq g_m \leq g_{m+1} \leq +h$$

und damit auch

$$-c \leq I_\mu(g_{m,k}) \leq +c, \quad c := I_\mu(h) < +\infty.$$

Die Folge $(g_{m,k})_{k=1}^\infty$ ist für jedes feste m monoton fallend und konvergiert punktweise gegen g_m , wobei auch $\lim_{k \rightarrow \infty} I_\mu(g_{m,k}) \geq -c$ gilt. Der Satz über die monotone Konvergenz liefert daher sowohl die Integrierbarkeit von g_m als auch

$$I_\mu(g_m) = \lim_{k \rightarrow \infty} I_\mu(g_{m,k}), \quad I_\mu(g_m) \leq c.$$

Wir können schließlich den Satz über die monotone Konvergenz auch auf die monoton wachsende Folge $(g_m)_{m=1}^\infty$ anwenden und erhalten die Integrierbarkeit von f sowie

$$I_\mu(f) = \lim_{m \rightarrow \infty} I_\mu(g_m)$$

als Zwischenergebnis.

Spezialfall, Teil 2: Wir führen zusätzlich die Funktionen

$$\tilde{g}_{m,k} := \max \{f_n : m \leq n \leq m+k\}, \quad \tilde{g}_m := \sup \{f_n : n \geq m\},$$

ein und zeigen

$$f(x) = \limsup_{n \rightarrow \infty} f_n(x) = \lim_{m \rightarrow \infty} \tilde{g}_m(x), \quad I_\mu(f) = \lim_{m \rightarrow \infty} I_\mu(\tilde{g}_m)$$

analog zum ersten Teil. Außerdem gilt

$$|f_n - f| \leq \tilde{g}_n - g_n$$

und wir erhalten

$$0 \leq |I_\mu(f_n) - I_\mu(f)| \leq |I_\mu(f_n - f)| \leq I_\mu(|f_n - f|) = \|f_n - f\|_{1,\mu} \leq I_\mu(\tilde{g}_n) - I_\mu(g_n).$$

Die gewünschten Grenzwertformeln ergeben sich nun im Grenzübergang $n \rightarrow \infty$.

Allgemeiner Fall: Nach Voraussetzung gibt es eine Menge $E \in \mathcal{A}$ mit $\mu(E') = 0$, sodass $f(x) = \lim_{n \rightarrow \infty} f_n(x)$ für alle $x \in E$ gilt. Wir definieren

$$\tilde{f}_n := \chi_E f_n, \quad \check{f}_n := \chi_{E'} f_n, \quad \tilde{f} := \chi_E f, \quad \check{f} := \chi_{E'} f,$$

wobei \tilde{f} als punktweiser Limes von \tilde{f}_n messbar ist. Die Messbarkeit von \check{f} und $|\check{f}|$ ergibt sich aus der Vollständigkeit von μ und der Tatsache, dass E' eine Nullmenge ist. Mit kleinen Nebenrechnungen (siehe dazu auch das Lemma zu den elementaren Eigenschaften sowie das Korollar zur zweiten Version des Satzes über die monotone Konvergenz) zeigen wir außerdem

$$I_\mu(\check{f}_n) = I_\mu(\check{f}) = 0, \quad \|\check{f}_n - \check{f}\|_{1,\mu} = 0.$$

Für die Funktionen \tilde{f}_n und \tilde{f} können wir hingegen die Ergebnisse des Spezialfalls verwenden und erhalten insgesamt die Behauptung. \square

Bemerkungen

1. Die Funktion h wird die Majorante genannt, wobei immer $h \geq 0$ sowie $I_\mu(h) < \infty$ gilt.
2. Die Voraussetzung $|f_n| \leq h$ kann leicht abgeschwächt werden. Es reicht zu fordern, dass für jedes $n \in \mathbb{N}$ die punktweise Abschätzung $|f_n(x)| \leq h(x)$ für fast alle $x \in X$ erfüllt ist, wobei die Ausnahmemenge sogar von n abhängen darf.
3. Das Theorem gilt auch für nicht-vollständige Maße μ , sofern wir zulassen, dass alle Funktionen auf Nullmengen abgeändert werden dürfen.
4. Ganz allgemein (und nicht nur unter den Voraussetzungen des Theorems) gilt

$$|I_\mu(f_n) - I_\mu(f)| \leq \|f_n - f\|_{1,\mu},$$

d.h. aus $\lim_{n \rightarrow \infty} \|f_n - f\|_{1,\mu} = 0$ (Konvergenz in der Halbnorm) folgt auch immer $\lim_{n \rightarrow \infty} I_\mu(f_n) = I_\mu(f)$ (Konvergenz der Integrale).

5. Ohne Majorante können wir aus der punktweisen Konvergenz weder auf die Integrierbarkeit der Limesfunktion noch auf die Konvergenz der Halbnorm schließen.

Gegenbeispiele

1. Im eindimensionalen Lebesgue-Fall betrachten wir

$$f_n(x) := \max \left\{ -n^3, \min \left\{ +n^3, 1/x^3 \right\} \right\} \xrightarrow{n \rightarrow \infty} f(x) := 1/x^3,$$

wobei $f_n(0) = f(0) = \infty$ vereinbart sei. Jede Funktion $f_n : \mathbb{R} \rightarrow \mathbb{R}$ ist ungerade und integrierbar mit

$$I_\lambda(f_n, -) = I_\lambda(f_n, +) = \int_0^{1/n} n^3 dx + \int_{1/n}^{\infty} 1/x^3 dx = 3n^2, \quad I_\lambda(f_n) = 0,$$

aber f ist nicht integrierbar wegen

$$I_\lambda(f_\pm) = \lim_{n \rightarrow \infty} I_\lambda(f_{n,\pm}) = \infty.$$

Beachte, dass es hier eben keine integrierbare Majorante gibt.

2. Mit der Wahl

$$f_n(x) := \min\{n^2, 1/x^2\} \xrightarrow{n \rightarrow \infty} f(x) := 1/x^2$$

berechnen wir

$$\|f_n - f\|_{1,\lambda} = \int_{-\infty}^{+\infty} |f_n(x) - f(x)| dx = 2 \int_0^{1/n} (1/x^2 - n^2) dx = 2(\infty - n) = \infty,$$

d.h. $\|f_n - f\|_{1,\lambda}$ wird im Limes $n \rightarrow \infty$ nicht klein. Beachte, dass jede Funktion f_n nichtnegativ und integrierbar ist und dass f zwar nichtnegativ, aber wegen $I_\mu(f) = \lim_{n \rightarrow \infty} I_\lambda(f_n) = \infty$ wieder nicht integrierbar ist.

Theorem (Konvergenz von Cauchy-Folgen) Sei $(f_n)_{n=1}^\infty$ eine Cauchy-Folge bzgl. $\|\cdot\|_{1,\mu}$, d.h.

$$\sup \left\{ \|f_n - f_l\|_{1,\mu} : l \geq n \right\} \xrightarrow{n \rightarrow \infty} 0.$$

Dann existiert eine integrierbare Funktion $f : X \rightarrow \overline{\mathbb{R}}$ mit

$$\|f_n - f\|_{1,\mu} \xrightarrow{n \rightarrow \infty} 0.$$

Außerdem existiert eine strikt monotone Indexfolge $(n_j)_{j=1}^\infty$ sowie eine nichtnegative und integrierbare Funktion $h : X \rightarrow \overline{\mathbb{R}}$, sodass

$$|f_{n_j}(x)| \leq h(x)$$

für jedes $x \in X$ und alle $j \in \mathbb{N}$ sowie

$$f_{n_j}(x) \xrightarrow{j \rightarrow \infty} f(x)$$

für fast alle $x \in X$.

Beweis Wir wählen zunächst eine strikt monotone Indexfolge $(n_j)_{j=1}^\infty$, sodass

$$\sup \left\{ \|f_{n_j} - f_l\|_{1,\mu} : l \geq n_j \right\} \leq \left(\frac{1}{2}\right)^j$$

und setzen

$$g_m := f_{n_{m+1}} - f_{n_m}, \quad s_j := \sum_{m=1}^j g_m, \quad t_j := \sum_{m=1}^j |g_m|$$

für alle $m \in \mathbb{N}$ und $j \in \mathbb{N}$. Die Funktionen g_m sowie s_j und t_j sind nach Voraussetzung integrierbar, wobei

$$I_\mu(g_m) \leq I_\mu(|g_m|) = \|f_{n_{m+1}} - f_{n_m}\|_{1,\mu} \leq \left(\frac{1}{2}\right)^m$$

für alle $m \in \mathbb{N}$ sowie

$$0 \leq t_j \leq t_{j+1}, \quad I_\mu(t_j) = \sum_{m=1}^j I_\mu(|g|_m) \leq \sum_{m=1}^j \left(\frac{1}{2}\right)^m \leq \sum_{m=1}^{\infty} \left(\frac{1}{2}\right)^m = 1$$

für alle $j \in \mathbb{N}$ gilt. Nach dem Satz über monotone Konvergenz ist

$$t(x) := \lim_{j \rightarrow \infty} t_j(x)$$

integrierbar mit $I_\mu(t) = \lim_{j \rightarrow \infty} I_\mu(t_j) \leq 1$ und die Menge

$$N := \{x \in X : t(x) = \infty\}$$

muss eine Nullmenge sein (da andernfalls $I_\mu(t) = \infty$ gelten würde). Für jedes $x \notin N$ gilt

$$\sum_{m=1}^{\infty} |g_m(x)| = t(x) < \infty,$$

d.h. die Reihe zur reellen Zahlenfolge $(g_m(x))_{m=1}^{\infty}$ konvergiert absolut. Insbesondere wird durch

$$s(x) := \lim_{j \rightarrow \infty} s_j(x) = \sum_{m=1}^{\infty} g_m(x) \quad \text{für } x \in N', \quad g(x) := 0 \quad \text{für } x \in N$$

eine Funktion $s : X \rightarrow \mathbb{R}$ definiert, die als punktweiser Grenzwert der Folge $(\chi_{N'} s_j)_{j=1}^{\infty}$ messbar ist (siehe das entsprechende Lemma im Abschnitt zur Messbarkeit). Da nach Konstruktion außerdem

$$|s_j| \leq t_j \leq t, \quad I_\mu(s_j) \leq I_\mu(t_j) \leq I_\mu(t) = 1$$

für alle $j \in \mathbb{N}$ erfüllt ist, ergibt sich die Integrierbarkeit von s sowie

$$\|s_j - s\|_{1,\mu} \xrightarrow{j \rightarrow \infty} 0$$

aus dem Satz über majorisierte Konvergenz. Wir setzen nun

$$f(x) := f_{n_1}(x) + s(x), \quad h(x) := |f_{n_1}(x)| + t(x)$$

und bemerken, dass beide Funktionen integrierbar sind. Unsere Definitionen implizieren weiterhin

$$s_{j-1} = \sum_{m=1}^{j-1} g_m = (f_{n_2} - f_{n_1}) + \dots + (f_{n_j} - f_{n_{j-1}}) = f_{n_j} - f_{n_1}, \quad s_{j-1} - s = f_{n_j} - f$$

für alle $j \in \mathbb{N}$ und alle Konvergenzbehauptungen für die Teilfolge $(f_{n_j})_{j=1}^{\infty}$ ergeben sich unmittelbar. Sei nun $\varepsilon > 0$ fixiert. Dann wählen wir zuerst $k_* \in \mathbb{N}$ mit

$$\sup \left\{ \|f_k - f_l\|_{1,\mu} : l \geq k \in \mathbb{N} \right\} \leq \frac{1}{2} \varepsilon$$

für $k \geq k_*$ und anschließend für jedes $k \geq k_*$ ein $j_k \in \mathbb{N}$ mit

$$n_{j_k} \geq k, \quad \|f_{n_{j_k}} - f\|_{1,\mu} \leq \frac{1}{2} \varepsilon.$$

Insgesamt erhalten wir

$$0 \leq \|f_k - f\|_{1,\mu} \leq \|f_k - f_{n_{j_k}}\|_{1,\mu} + \|f_{n_{j_k}} - f\|_{1,\mu} \leq \frac{1}{2} \varepsilon + \frac{1}{2} \varepsilon \leq \varepsilon,$$

und weil $\varepsilon > 0$ beliebig war, folgt $\lim_{k \rightarrow \infty} \|f_k - f\|_{1,\mu} = 0$. □

Bemerkung Das Theorem und sein Beweis können alternativ auch in die folgenden beiden Teile aufgespalten werden:

1. Jede Cauchy-Folge bzgl. $\|\cdot\|_{1,\mu}$ konvergiert in dieser Halbnorm gegen eine integrierbare Limesfunktion.
2. Jede Folge, die bzgl. der Halbnorm konvergiert, besitzt mindestens eine Teilfolge, die punktweise fast überall gegen die Limesfunktion konvergiert und außerdem majorisiert ist.⁴²

Die erste Aussage wird später die Vollständigkeit des entsprechenden Lebesgue-Raumes garantieren. Die zweite ist in gewisser Weise die Umkehrung zum Satz über die majorisierte Konvergenz. Beachte aber, dass im Allgemeinen wirklich nur Teilfolgen punktweise konvergieren (siehe dazu das Bild).

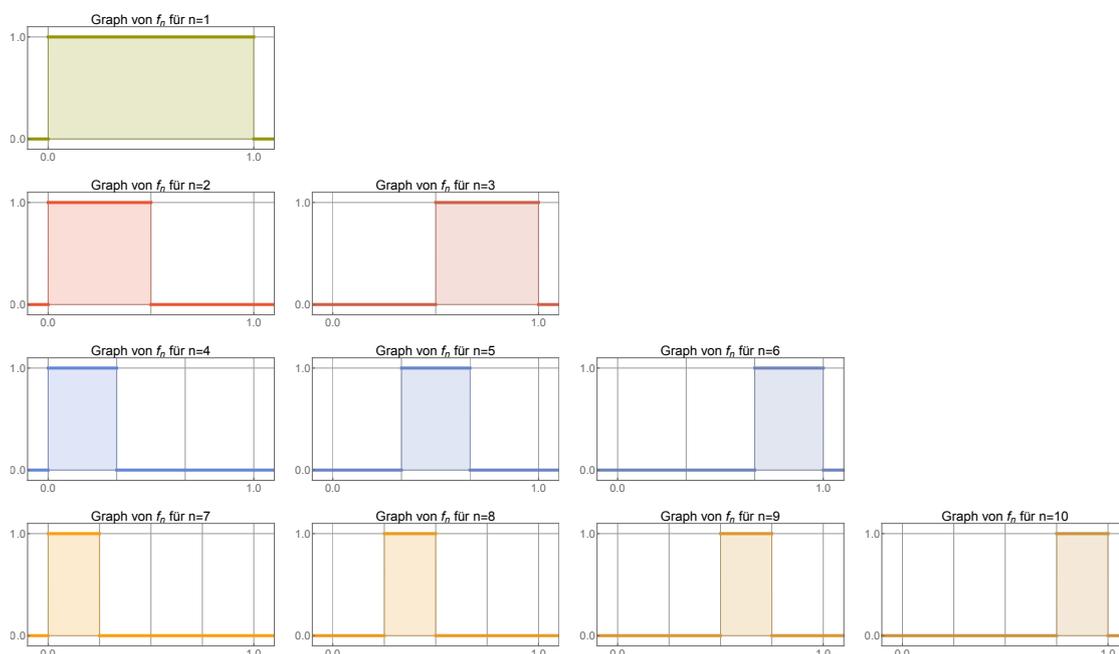


Abbildung Eine Variante der Haarschen Funktionenfolge $(f_n)_{n=1}^\infty$ auf $X = \mathbb{R}$ mit

$$f_1 := \chi_{(0,1]}, \quad f_2 := \chi_{(0,1/2]}, \quad f_3 := \chi_{(1/2,1]}, \quad f_4 := \chi_{(0,1/3]}, \quad f_5 := \chi_{(1/3,2/3]}, \quad f_6 := \chi_{(2/3,1]}, \quad \dots$$

Diese konvergiert in der 1-Halbnorm (bzgl. des Lebesgue-Maßes) gegen die Nullfunktion $f = 0$, aber für jedes $x \in [0, 1]$ nicht punktweise, da die entsprechende Zahlenfolge $(f_n(x))_{n=1}^\infty$ jeweils unendlich viele Nullen und Einsen enthält. Die Teilfolge in der ersten Spalte konvergiert jedoch punktweise gegen die Nullfunktion (sogar überall), wobei die entsprechende Indexfolge $(n_j)_{j=1}^\infty$ durch $n_j = 1 + (j-1)j/2$ gegeben ist.

Zusatz*

Räume integrierbarer Funktionen Zwei messbare Funktionen $f, \tilde{f} : X \rightarrow \overline{\mathbb{R}}$ heißen äquivalent bzgl. μ , falls beide fast überall gleich sind, und wir schreiben $f \sim \tilde{f}$. Dies definiert eine Äquivalenzrelation mit den Klassen

$$[f] := \{ \tilde{f} : \tilde{f} \sim f \},$$

⁴²Beachte, dass jede Folge $(f_n)_{n=1}^\infty$, die in der 1-Halbnorm konvergiert, die Cauchy-Eigenschaft besitzt (Übungsaufgabe).

und wir können für jeden Parameter $1 \leq p < \infty$ zeigen, dass die Definition

$$\|[f]\|_{p,\mu} = \left(\int_X |f(x)|^p dx \right)^{1/p}$$

sinnvoll ist, d.h. dass die rechte Seite sich nicht ändert, wenn wir f durch \tilde{f} mit $f \sim \tilde{f}$ ersetzen (Übungsaufgabe). Analoges gilt für

$$\|[f]\|_{\infty,\mu} = \lim_{p \rightarrow \infty} \|f\|_{p,\mu} = \inf \{c \in \overline{\mathbb{R}} : |f(x)| \leq c \text{ für fast alle } x \in X\},$$

wobei $\|[f]\|_{\infty,\mu}$ auch das essentielle Supremum von f genannt wird.

Die Lebesgue-Räume werden durch

$$\mathbb{L}_\mu^p(X) := \{[f] : \|f\|_{p,\mu} < \infty\}$$

eingeführt und können via

$$\alpha [f] + \beta [g] := [\alpha f + \beta g]$$

in natürlicher Weise mit einer Addition sowie einer skalaren Multiplikation ausgestattet werden. Insbesondere besitzt nun $\|\cdot\|_{p,\mu}$ alle Eigenschaften einer Norm, denn aus $\|[f]\|_{p,\mu} = 0$ folgt $f \sim 0$ und damit $[f] = [0]$.

Achtung Die Elemente des Raumes $\mathbb{L}_\mu^p(X)$ sind per Definition keine Funktionen auf X , sondern Mengen von Funktionen auf X (nämlich die Äquivalenzklassen). In der mathematischen Analysis werden aber Funktionen oftmals mit ihrer Klasse identifiziert, d.h. wir schreiben $f \in \mathbb{L}_\mu^p(X)$ anstelle von $[f] \in \mathbb{L}_\mu^p(X)$. Diese Sichtweise hat viele Vorteile, bringt aber auch einige gewöhnungsbedürftige Besonderheiten mit sich. Zum Beispiel können wir nicht sagen, welchen Wert eine \mathbb{L}_μ^p -Funktion f in einem einzelnen Punkt $x \in X$ annimmt, da dieser Wert sich beim Übergang von f zu \tilde{f} mit $\tilde{f} \sim f$ ändern kann (sofern $\{x\}$ eine Nullmenge bzgl. μ ist). Wir können aber sagen, was das Integral von f über einer messbaren Menge E ist, da die Antwort auf diese Frage nicht von der Wahl des Repräsentanten abhängt.

Ausblick

1. $\mathbb{L}_\mu^p(X)$ ist ein vollständiger normierter Raum und damit ein *Banach-Raum*. $\mathbb{L}_\mu^2(X)$ ist sogar ein *Hilbert-Raum*, wobei

$$\langle f, g \rangle = \int_X f(x) g(x) dx$$

das zugrunde liegende Skalarprodukt ist.

2. Es gilt die Hölder-Ungleichung

$$\left| \int_X f(x) g(x) dx \right| \leq \|f\|_{p,\mu} \|g\|_{p',\mu},$$

wobei p' der konjugierte Exponent zu p ist und durch

$$p' := \frac{p}{p-1} \quad \text{bzw.} \quad \frac{1}{p} + \frac{1}{p'} = 1$$

bestimmt werden kann. Mit direkten Rechnungen zeigen wir

$$1' = \infty, \quad \left(\frac{5}{3}\right)' = \frac{5}{2}, \quad 2' = 2, \quad \left(\frac{7}{3}\right)' = \frac{7}{4}, \quad \infty' = 1$$

sowie $(p)' = p$ für alle $p \in [1, \infty]$.

3. Für $p < \infty$ gilt

$$(\mathbb{L}_\mu^p(X))^* \cong \mathbb{L}_\mu^{p'}(X)$$

im Sinne einer kanonischen Isomorphie, d.h. $\mathbb{L}_\mu^{p'}(X)$ kann als der duale Vektorraum zu $\mathbb{L}_\mu^p(X)$ angesehen werden.⁴³

Achtung: $\mathbb{L}_\mu^\infty(X)$ ist der Dualraum von $\mathbb{L}_\mu^1(X)$, aber der Dualraum von $\mathbb{L}_\mu^\infty(X)$ ist im Allgemeinen größer als $\mathbb{L}_\mu^1(X)$.

4. Für $1 < p < \infty$ ergibt sich insbesondere

$$(\mathbb{L}_\mu^p(X))^{**} = (\mathbb{L}_\mu^{p'}(X))^* = \mathbb{L}_\mu^p(X),$$

d.h. $\mathbb{L}_\mu^p(X)$ ist ein *reflexiver* Banach-Raum.

5. Ausgehend von den Lebesgue-Räumen werden die *Sobolev-Räume* durch

$$\mathbb{W}_\mu^{k,p}(X) := \{f \in \mathbb{L}_\mu^p(X) : \partial_x f, \dots, \partial_x^k f \in \mathbb{L}_\mu^p(X)\}$$

eingeführt, wobei die Ableitungen in einem verallgemeinerten bzw. im *schwachen Sinne* zu verstehen sind. Diese und andere Funktionenräume spielen eine zentrale Rolle in der modernen Theorie partieller Differentialgleichungen.

⁴³Dualräume werden in der *Funktionalanalysis* eingeführt.

Kapitel 2

Integration in den Anwendungen

Vorlesung 06-1 : 30. November

Vorbemerkung In diesem Kapitel diskutieren wir inner- sowie außermathematische Anwendungen der Integralrechnung, wobei wir uns zunächst auf n -dimensionale Lebesgue-Maß auf Teilmengen des \mathbb{R}^n konzentrieren und andere, nicht verschiebungs-invariante Maße — zum Beispiel Lebesgue-Stieltjes-Maße mit Dichten — ignorieren. Aus praktischer Sicht können wir dabei auch fast immer mit dem n -dimensionalen Riemann-Integral arbeiten, aber aus theoretischer Sicht ist der Unterschied erheblich, da das Lebesgue-Maß die deutlich besseren Eigenschaften aufweist.

Klarstellung

1. Wir studieren in diesem Abschnitt *Gebietsintegrale* (oft auch *Bereichsintegrale* genannt) als Verallgemeinerung der bestimmten Integrale aus *Analysis 1*. Das Konzept eines unbestimmten Integrals sowie der Hauptsatz der Differential- und Integralrechnung ist sehr viel schwieriger zu verallgemeinern.
2. Wir werden Spezialfälle der Integralsätze von Gauß und Stokes herleiten. Die allgemeine Form dieser Sätze braucht die Integration über gekrümmten Flächen, die wir erst im nächsten Kapitel kennenlernen.

Notation Wir bezeichnen im Folgenden Punkte¹ im \mathbb{R}^n mit

$$\mathbf{x} = (x_1, \dots, x_n)$$

obwohl die traditionelle mathematische Notation $x = (x_1, \dots, x_n)$ ist, und meinen mit

$$\langle \mathbf{x}, \mathbf{y} \rangle = x_1 y_1 + \dots + x_n y_n \quad \text{bzw.} \quad \|\mathbf{x}\| = \|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} = \sqrt{x_1^2 + \dots + x_n^2}$$

das entsprechende Skalarprodukt bzw. die euklidische Norm. Analog schreiben wir $f : D \rightarrow \mathbb{R}$ bzw. $\mathbf{f} : D \rightarrow \mathbb{R}^m$ für skalare bzw. vektorielle Funktionen auf einer Menge $D \subseteq \mathbb{R}^n$. Insbesondere gilt

$$\mathbf{f}(\mathbf{x}) = \begin{pmatrix} f_1(x_1, \dots, x_n) \\ \vdots \\ f_m(x_1, \dots, x_n) \end{pmatrix},$$

d.h. die vektorielle Funktion \mathbf{f} besitzt die Komponenten f_1, \dots, f_m , die jeweils skalare Funktionen auf D sind.

¹An einigen Stellen werden wir sauber zwischen Spalten- und Zeilenvektoren unterscheiden, aber im Moment ist das noch nicht wichtig.

2.1 Exkurs*: das Riemann-Integral auf Quadern

Ziel Wir führen zunächst das Analogon zum bestimmten Integral für skalare Funktionen $f : Q \rightarrow \mathbb{R}$ ein, die auf einem abgeschlossenen Quader Q definiert sind.

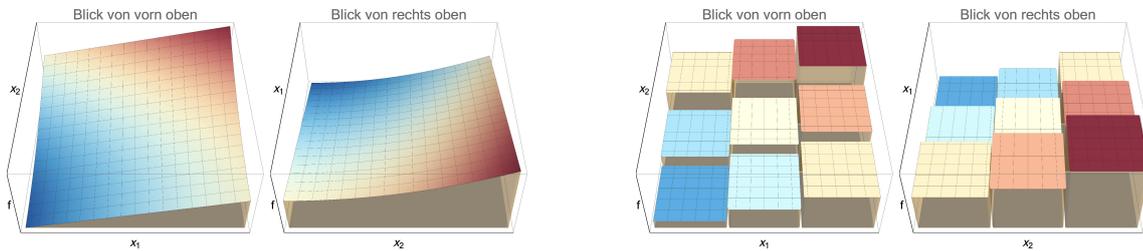


Abbildung Links: Das bestimmte Integral einer Funktion f auf einem n -dimensionalen Quader Q ist das (vorzeichenbehaftete) $n+1$ -dimensionale Volumen (Braun) unter dem Graphen von f , wobei Bereiche mit $f > 0$ bzw. $f < 0$ positiv bzw. negativ gezählt werden. Die dargestellte Funktion mit $n = 2$ nimmt nur positive Werte an und wird daher ein positives Integral liefern. Rechts: Das Integral kann als Summe von Produkten berechnet werden, sofern die Funktion konstant auf n -dimensionalen Teilquadern ist.

Zerlegungen von Quadern Eine Menge der Bauart

$$Q = [a_1, b_1] \times \dots \times [a_n, b_n] \subset \mathbb{R}^n$$

wird n -dimensionaler Quader genannt, wobei ein eindimensionaler Quader gerade ein abgeschlossenes Intervall ist. Wir wollen immer $a_j < b_j$ voraussetzen und nennen

$$\text{vol}(Q) := |b_1 - a_1| \cdot \dots \cdot |b_n - a_n|,$$

das n -dimensionale Volumen von Q . Ein eindimensionales Volumen ist dabei eine Länge, ein zweidimensionales ein Flächeninhalt.

Eine Zerlegung Z von Q zerlegt das j -te Koordinaten-Intervall $[a_i, b_i]$ des Quaders via

$$a_j = x_{j,0} < x_{j,1} < x_{j,2} < \dots < x_{j,K_j-1} < x_{j,K_j} = b_j,$$

in K_j Teilintervalle. Insgesamt wird der Quader Q damit in viele Teilquader

$$Q_{k_1, \dots, k_n} := [x_{1,k_1-1}, x_{1,k_1}] \times \dots \times [x_{n,k_n-1}, x_{n,k_n}], \quad k_j = 1 \dots K_j$$

zerlegt, wobei $K_1 \cdot \dots \cdot K_n$ gerade die Anzahl dieser Teilquader ist. Die Zahl

$$\|Z\| = \max_{j=1 \dots n} \max_{k=1 \dots K_j} |x_{j,k} - x_{j,k-1}|$$

wird die Feinheit der Zerlegung genannt. Sie beschreibt die maximale Kantenlänge eines Teilquaders.

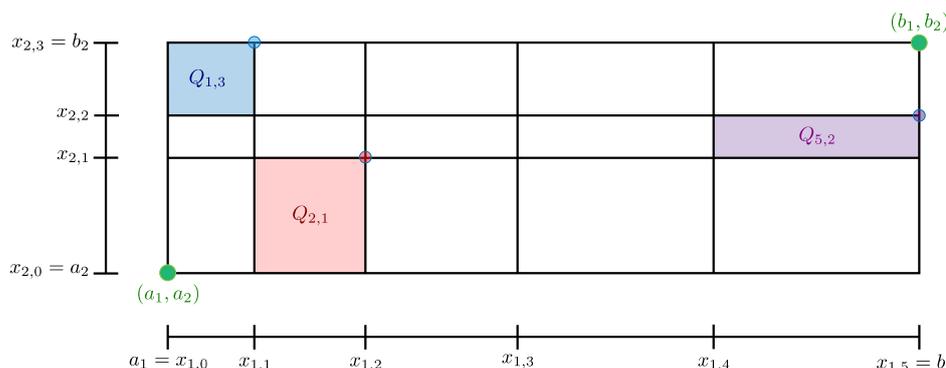


Abbildung Zerlegung eines Quaders in Teilquader, hier dargestellt für $n = 2$ und $K_1 = 5$ und $K_2 = 3$. Die genauen Details der Indizierung sind dabei gar nicht so wichtig.

Verfeinerung Eine weitere Zerlegung \tilde{Z} des Quaders Q mit

$$a_j = \tilde{x}_{j,0} < \tilde{x}_{j,1} < \tilde{x}_{j,2} < \dots < \tilde{x}_{j,\tilde{K}_j-1} < \tilde{x}_{j,\tilde{K}_j} = b_j,$$

wird Verfeinerung von Z genannt, wenn jeder Teilquader von \tilde{Z} ganz in einem Teilquader von Z liegt. Alternativ kann man sagen, dass jede der \tilde{Z} zugrunde liegenden Intervall-Zerlegungen eine Verfeinerung der entsprechenden Intervall-Zerlegung von Z ist.

Beispiel Für jedes $K \in \mathbb{N}$ wird durch

$$x_{j,k} = a_j + \frac{k}{K} (b_j - a_j), \quad k = 0 \dots K_j := K$$

eine äquidistante Intervall-Zerlegung von $[a_j, b_j]$ definiert. Insgesamt entsteht eine Zerlegung von Q in K^n Teilquader, die alle paarweise kongruent zueinander sind. Diese wollen wir uniforme Zerlegung von Q nennen und mit $Z^{(K)}$ bezeichnen.

Bemerkung* Für zwei gegebene Zerlegungen Z und \tilde{Z} kann man eine dritte dadurch erzeugen, dass man für jedes $j = 1 \dots n$ die Mengen $\{x_{j,0}, \dots, x_{j,K_j}\}$ sowie $\{\tilde{x}_{j,0}, \dots, \tilde{x}_{j,\tilde{K}_j}\}$ vereinigt und die Elemente anschließend der Größe nach neu durchnummeriert. Die dritte Zerlegung wird auch als gemeinsame Verfeinerung der ersten beiden bezeichnet.

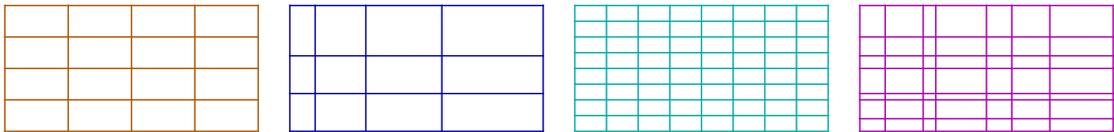


Abbildung Vier verschiedene Zerlegungen eines zweidimensionalen Quaders, wobei die orange und die türkise jeweils uniform sind, da sie aus äquidistanten Intervall-Zerlegungen gewonnen wurden. Die türkise Zerlegung ist dabei eine Verfeinerung der orangen und die lilane ist die gemeinsame Verfeinerung der orangen und der blauen Zerlegung.

Strategie Mithilfe von Quaderzerlegungen können wir die Theorie des Riemann-Integrals ganz analog zum eindimensionalen Fall aufbauen. Der einzige Unterschied ist, dass wir nun mit n -dimensionalen Teilquadern von Q sowie $n+1$ -dimensionalen Teilquadern arbeiten, wobei letztere benutzt werden, um die Fläche unter dem Graphen von f zu approximieren. Beachte dabei, dass das Volumen eines $n+1$ -dimensionalen Quaders als das Produkt seiner Höhe sowie dem Volumen seiner n -dimensionalen Grundfläche berechnet werden kann.

Spezielle und allgemeine Riemann-Summen Analog zu *Analysis 1* führen wir durch

$$U(Z) = \sum_{k_1=1}^{K_1} \dots \sum_{k_n=1}^{K_n} \inf f|_{Q_{k_1, \dots, k_n}} \cdot \text{vol}(Q_{k_1, \dots, k_n})$$

bzw.

$$O(Z) = \sum_{k_1=1}^{K_1} \dots \sum_{k_n=1}^{K_n} \sup f|_{Q_{k_1, \dots, k_n}} \cdot \text{vol}(Q_{k_1, \dots, k_n})$$

die Untersumme bzw. die Obersumme ein. Hierbei bezeichnen die Zahlen

$$\inf f|_M := \inf\{f(\mathbf{x}) : \mathbf{x} \in M\} \quad \text{bzw.} \quad \sup f|_M := \sup\{f(\mathbf{x}) : \mathbf{x} \in M\}$$

für stetige Funktionen f und kompakte Mengen M gerade das Minimum bzw. Maximum von f auf M , aber im Allgemeinen müssen sie als Infimum bzw. als Supremum von f auf M betrachtet werden. Etwas allgemeiner ist die Riemann-Summe

$$R(Z) = \sum_{k_1=1}^{K_1} \cdots \sum_{k_n=1}^{K_n} f(\boldsymbol{\xi}_{k_1, \dots, k_n}) \cdot \text{vol}(Q_{k_1, \dots, k_n}),$$

aber hier muss f für jeden Teilquader in einer Stützstelle $\boldsymbol{\xi}_{k_1, \dots, k_n} \in Q_{k_1, \dots, k_n}$ ausgewertet werden, deren Wahl natürlich den Wert von $R(Z)$ beeinflussen wird.

Alternative Notation Man schreibt auch gerne $Q_{\mathbf{k}}$ statt Q_{k_1, \dots, k_n} mit Multi-Index $\mathbf{k} = (k_1, \dots, k_n)$ sowie

$$U(Z) = \sum_{Q_{\mathbf{k}} \in Z} \inf f|_{Q_{\mathbf{k}}} \cdot \text{vol}(Q_{\mathbf{k}}), \quad O(Z) = \sum_{Q_{\mathbf{k}} \in Z} \sup f|_{Q_{\mathbf{k}}} \cdot \text{vol}(Q_{\mathbf{k}}),$$

wobei die Multi-Summe auf der jeweiligen rechten Seite über alle Teilquader $Q_{\mathbf{k}}$ der Zerlegung gebildet wird und damit eigentlich eine n -fache Summe ist.

Beispiel Wir betrachten für $n = 2$ das Einheitsquadrat $Q = [0, 1] \times [0, 1]$ sowie die Funktion $f : Q \rightarrow \mathbb{R}$ mit

$$f(x_1, x_2) = x_1 + x_2^2$$

Für jedes $K \in \mathbb{N}$ wählen wir die entsprechende uniforme Zerlegung $Z^{(K)}$ von Q mit Stützstellen

$$x_{1,k} = \frac{k_1}{K}, \quad k_1 = 0 \dots K_1 = K, \quad x_{2,k} = \frac{k_2}{K}, \quad k_2 = 0 \dots K_2 = K$$

und Teilquadern

$$Q_{k_1, k_2} = \left[\frac{k_1 - 1}{K}, \frac{k_1}{K} \right] \times \left[\frac{k_2 - 1}{K}, \frac{k_2}{K} \right], \quad \text{vol}(Q_{k_1, k_2}) = \frac{1}{K^2}, \quad k_1, k_2 = 1 \dots K.$$

Die (sehr einfache) Funktion f nimmt auf jedem Teilquader in der jeweiligen linken unteren bzw. rechten oberen Ecke ihr Minimum bzw. Maximum an und wir erhalten daher

$$\inf f|_{Q_{k_1, k_2}} = \frac{(k_1 - 1)}{K} + \frac{(k_2 - 1)^2}{K^2} \quad \text{bzw.} \quad \sup f|_{Q_{k_1, k_2}} = \frac{k_1}{K} + \frac{k_2^2}{K^2}.$$

Dies impliziert

$$\begin{aligned} U(Z^{(K)}) &= \sum_{k_1=1}^K \sum_{k_2=1}^K \left(\frac{(k_1 - 1)}{K} + \frac{(k_2 - 1)^2}{K^2} \right) \cdot \frac{1}{K^2} \\ &= \left(K \sum_{k_1=1}^K \frac{k_1 - 1}{K^3} \right) + \left(K \sum_{k_2=1}^K \frac{(k_2 - 1)^2}{K^4} \right) \\ &= \left(\frac{1}{K^2} \sum_{i=0}^{K-1} i \right) + \left(\frac{1}{K^3} \sum_{i=0}^{K-1} i^2 \right) \\ &= \left(\frac{K(K-1)}{2K^2} \right) + \frac{K \cdot (K-1 + 2(K-1)^2)}{6K^3} = \frac{5}{6} - \frac{1}{K} + \frac{1}{6K^2} \end{aligned}$$

sowie analog

$$O(Z^{(K)}) = \sum_{k_1=1}^K \sum_{k_2=1}^K \left(\frac{k_1}{K} + \frac{k_2^2}{K^2} \right) \cdot \frac{1}{K^2} = \left(\frac{1}{K^2} \sum_{i=1}^K i \right) + \left(\frac{1}{K^3} \sum_{i=1}^K i^2 \right) = \frac{5}{6} + \frac{1}{6K^2},$$

wobei wir Indexverschiebung sowie die allgemeinen Summenformeln

$$\sum_{i=I_1}^{I_2} i = \frac{(I_2 - I_1 + 1)(I_1 + I_2)}{2}, \quad \sum_{i=I_1}^{I_2} i^2 = \frac{(I_2 - I_1 + 1)(I_2 - I_1 + 2I_1^2 + 2I_1I_2 + 2I_2^2)}{6}$$

verwendet haben. Insbesondere erhalten wir

$$0 \leq O(Z^{(K)}) - U(Z^{(K)}) = \frac{1}{K} \xrightarrow{K \rightarrow \infty} 0,$$

das heißt der Unterschied zwischen Ober- und Untersumme wird kleiner, je feiner die Zerlegung wird.

Bemerkung: Dieses einfache Beispiel illustriert, wie man Funktionen auf Quadern durch Approximation mit Ober- und Untersummen integrieren kann. Dieses Verfahren ist aber nur auf Computern wirklich praktikabel, da es in den meisten Fällen keine einfache Summenformel geben wird. Unten werden wir eine andere Möglichkeiten der Integralberechnung kennenlernen, zum Beispiel den Satz von Fubini.

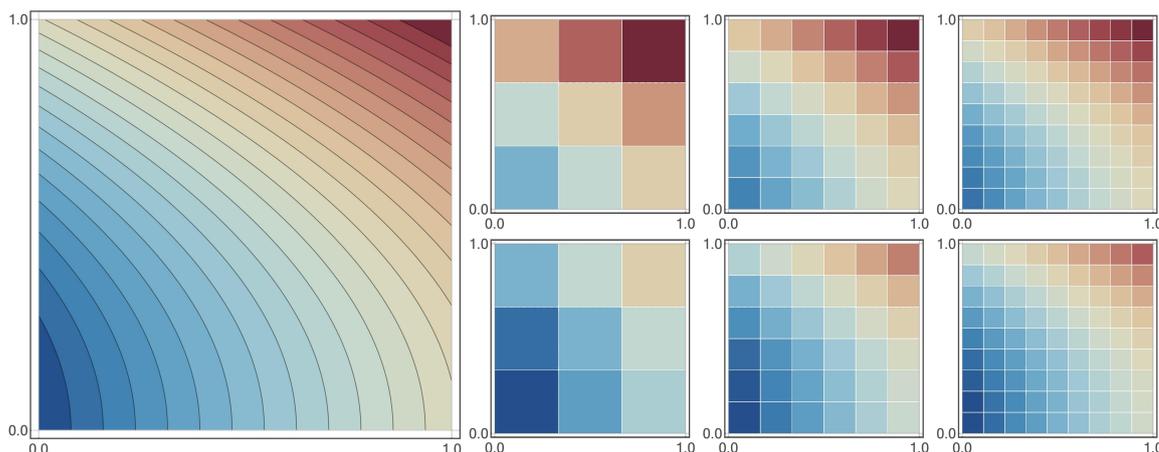


Abbildung Illustration des eben gerechneten Beispiels, wobei ganz links die Funktion dargestellt ist. Bei der Berechnung der Ober- (oben) und Untersummen (unten) wird die Funktion f auf jedem Teilquader durch eine Konstante ersetzt, nämlich das Maximum bzw. das Minimum von f auf diesem Quader (hier dargestellt für die uniformen Zerlegungen mit $K = 3$, $K = 6$ und $K = 9$).

Monotonie-Eigenschaften Ganz analog zu *Analysis 1* können wir aus unseren Definitionen die folgenden Aussagen abgeleitet werden:

1. Es gilt

$$U(Z) \leq R(Z) \leq O(Z)$$

für jede Zerlegung Z .

2. Ist \tilde{Z} eine beliebige Verfeinerung von Z , so folgt

$$U(Z) \leq U(\tilde{Z}) \leq O(\tilde{Z}) \leq O(Z).$$

3. Sind $Z^{(1)}$ und $Z^{(2)}$ zwei beliebige Zerlegungen, so gilt

$$U(Z^{(1)}) \leq O(Z^{(2)}).$$

Folgerung Die reellen Zahlen

$$\bar{U} := \sup\{U(Z) : Z \text{ ist Zerlegung von } Q\}$$

bzw.

$$\underline{O} := \inf\{O(Z) : Z \text{ ist Zerlegung von } Q\}$$

sind wohldefiniert und werden das Unterintegral bzw. das Oberintegral von f genannt. Insbesondere gilt

$$U(Z) \leq \bar{U} \leq \underline{O} \leq O(Z),$$

für jede Zerlegung Z , d.h. $\text{Untersumme} \leq \text{Unterintegral} \leq \text{Oberintegral} \leq \text{Obersumme}$.

Definition Gilt

$$\bar{U} = \underline{O} =: \int_Q f(\mathbf{x}) \, d\mathbf{x}$$

so heißt die Funktion f (Riemann-)integrierbar und das Symbol auf der rechten Seite wird das (Riemann-)Integral von f genannt.

Alternative Notation Für $n = 2$ bzw. $n = 3$ schreibt man oftmals auch

$$\iint_Q f(x_1, x_2) \, dx_1 \, dx_2 \quad \text{bzw.} \quad \iiint_Q f(x_1, x_2, x_3) \, dx_1 \, dx_2 \, dx_3$$

für das Integral von f über Q . Siehe dazu auch den Satz von Fubini weiter unten. Außerdem gilt

$$\int_{[a,b]} f(x) \, dx = \int_a^b f(x) \, dx$$

für $n = 1$.

Lemma (Integrale und Grenzwert) Ist $(Z^{(m)})_{m \in \mathbb{N}}$ eine beliebige Folge von Zerlegungen und gilt

$$\lim_{m \rightarrow \infty} U(Z^{(m)}) = \lim_{m \rightarrow \infty} O(Z^{(m)}),$$

so ist f integrierbar und beide Grenzwerte liefern das Integral $\int_Q f(\mathbf{x}) \, d\mathbf{x}$.

Beweis Auf Grund unserer Beobachtungen gilt

$$U(Z^{(m)}) \leq \bar{U} \leq \underline{O} \leq O(Z^{(m)})$$

und die Behauptung ergibt sich nach Grenzübergang $m \rightarrow \infty$ aus dem Sandwich-Prinzip für reelle Zahlenfolgen. Letzteres kann hier angewendet werden, weil Ober- und Untersummen selbst für $n > 1$ immer Zahlen sind. \square

Bemerkung

1. In aller Regel wird $\lim_{m \rightarrow \infty} \|Z^{(m)}\| = 0$ gelten, d.h. die Feinheit von $Z^{(m)}$ wird mit wachsendem m immer kleiner werden.
2. Um die Voraussetzung im Lemma zu überprüfen, reicht es, *eine* Folge von Zerlegungen mit der gewünschten Eigenschaft zu finden. Wir werden meist annehmen, dass $Z^{(m+1)}$ für jedes $m \in \mathbb{N}$ eine Verfeinerung von $Z^{(m)}$ ist oder dass jedes $Z^{(m)}$ eine uniforme Zerlegung beschreibt.
3. Für stetige Funktionen f gilt auch

$$\int_Q f(\mathbf{x}) \, d\mathbf{x} = \lim_{m \rightarrow \infty} R(Z^{(m)}),$$

sofern für jede Zerlegung $Z^{(m)}$ und jeden ihrer Teilquader eine entsprechende Stützstelle gewählt wurde.

4. Es gilt

$$\iint_{[0,1] \times [0,1]} f(x_1, x_2) \, dx_1 \, dx_2 = \frac{5}{6}$$

für die Daten aus dem letzten Beispiel.

Gegenbeispiel Die Funktion $f : [0, 1] \times \dots \times [0, 1] \rightarrow \mathbb{R}$ mit

$$f(\mathbf{x}) = \begin{cases} 1 & \text{falls alle Komponenten } x_j \text{ von } \mathbf{x} \text{ rationale Zahlen sind,} \\ 0 & \text{sonst} \end{cases}$$

ist *nicht* Riemann-integrierbar, denn es gilt

$$U(Z) = 0, \quad O(Z) = 1$$

für jede Zerlegung Z . Für $n = 1$ ist das gerade die *Dirichlet-Funktion*.

Theorem (eine hinreichende Bedingung für Integrierbarkeit) Jede stetige Funktion auf einem Quader ist integrierbar.

Beweisskizze Wir nehmen der Einfachheit halber an, dass $Q = [0, 1] \times \dots \times [0, 1]$ der n -dimensionale Einheitswürfel ist und dass $f : Q \rightarrow \mathbb{R}$ sogar stetig differenzierbar ist. Wir betrachten außerdem für jedes $K \in \mathbb{N}$ die uniforme Zerlegung $Z^{(K)}$, in der jeder Teilwürfel $Q_{\mathbf{k}}$ das Volumen

$$\text{vol}(Q_{\mathbf{k}}) = \frac{\text{vol}(Q)}{K^n} = \frac{1}{K^n}$$

sowie die Kantenlängen $1/K$ besitzt. Der Satz von Taylor – angewendet auf einen Teilquader $Q_{\mathbf{k}}$ – impliziert

$$|f(\mathbf{x}) - f(\mathbf{x}_*)| \leq \sum_{j=1}^n |\partial_{x_j} f(\mathbf{x}_*)| |x_{*,j} - x_j| \leq \frac{C}{K}$$

für alle $\mathbf{x}, \mathbf{x}_* \in Q_{\mathbf{k}}$, wobei die Konstante C via

$$C := \sum_{j=1}^n \max_{\mathbf{y} \in Q} |\partial_j f(\mathbf{y})|.$$

sogar global, d.h. für alle Teilwürfel gleich, gewählt werden kann. Wählen wir \mathbf{x} als Minimierer und \mathbf{x}_* als Maximierer von f in $Q_{\mathbf{k}}$ – diese existieren immer, weil f stetig und jeder Quader kompakt ist — ergibt sich

$$0 \leq \max f|_{Q_{\mathbf{k}}} - \min f|_{Q_{\mathbf{k}}} \leq \frac{C}{K},$$

und nach Multiplikation mit $\text{vol}(Q_{\mathbf{k}})$ sowie Summation über alle Teilquader erhalten wir

$$0 \leq O(Z^{(K)}) - U(Z^{(K)}) \leq K^n \frac{C}{K} \frac{1}{K^n} \leq \frac{C}{K}.$$

Die Behauptung folgt nun nach Grenzübergang $K \rightarrow \infty$ aus dem obigen Lemma. \square

Bemerkung Es gibt auch integrierbare Funktionen, die nicht stetig sind, zum Beispiel stückweise stetige Funktionen. Im Rahmen dieser Vorlesung werden solche Funktionen aber erstmal keine Rolle spielen.

2.2 Satz von Fubini auf Quadern

Ziel Wir leiten nun die Spezialfälle von drei wichtigen Theoremen auf Quadern her, nämlich die Integralsätze von Fubini, Gauß und Stokes. Das erste wird es uns erlauben, Integrale effektiv auszurechnen. Das zweite und das dritte Theorem können als eine Verallgemeinerung des Hauptsatzes der Differential- und Integralrechnung betrachtet werden und wir werden dies im Fortgang dieser Vorlesung immer wieder thematisieren.

Theorem (Satz von Fubini für 2- und für n -dimensionale Quader) Für jede stetige Funktion $f : Q \rightarrow \mathbb{R}$ auf einem zweidimensionalen Quader $Q = [a_1, b_1] \times [a_2, b_2]$ gilt

$$\int_Q f(\mathbf{x}) \, d\mathbf{x} = \int_{a_2}^{b_2} \left(\int_{a_1}^{b_1} f(x_1, x_2) \, dx_1 \right) dx_2 = \int_{a_1}^{b_1} \left(\int_{a_2}^{b_2} f(x_1, x_2) \, dx_2 \right) dx_1$$

d.h. das zweidimensionale Integral kann durch Verschachtelung zweier eindimensionaler Integrale berechnet werden. Analoge Aussagen gelten für n -dimensionale Quader, wobei dann n eindimensionale Integrale in beliebiger Reihenfolge zu verschachteln sind.

Beweisskizze Wir betrachten der Einfachheit halber nur den zweidimensionalen Einheitswürfel $Q = [0, 1] \times [0, 1]$, nehmen f als stetig differenzierbar an und wählen wieder für jedes $K \in \mathbb{N}$ die uniforme Zerlegung $Z^{(K)}$ bestehend aus K^2 gleichgroßen Teilwürfeln der Bauart

$$Q_{k_1, k_2} = I_{k_1} \times I_{k_2}, \quad I_k = \left[\frac{k-1}{K}, \frac{k}{K} \right], \quad \text{vol}(Q_{k_1, k_2}) = \frac{1}{K^2}.$$

Mit Approximation durch Riemann-Summen erhalten wir

$$\int_Q f(\mathbf{x}) \, d\mathbf{x} \approx \frac{1}{K^2} \sum_{k_1=1}^K \sum_{k_2=1}^K f\left(\frac{k_1}{K}, \frac{k_2}{K}\right),$$

wobei wir in jedem Teilwürfel den jeweils rechten oberen Punkt als Stützstelle gewählt haben und der Approximationsfehler mit einem Taylor-Argument – analog zum vorherigen Beweis – durch C/K abgeschätzt werden kann. Durch Umstellung der Terme erhalten wir

$$\begin{aligned} \int_Q f(\mathbf{x}) \, d\mathbf{x} &\approx \frac{1}{K} \sum_{k_1=1}^K \left(\frac{1}{K} \sum_{k_2=1}^K f\left(\frac{k_1}{K}, \frac{k_2}{K}\right) \right) \approx \frac{1}{K} \sum_{k_1=1}^K \left(\int_0^1 f\left(\frac{k_1}{K}, x_2\right) \, dx_2 \right) \\ &\approx \int_0^1 \left(\int_0^1 f(x_1, x_2) \, dx_2 \right) \, dx_1, \end{aligned}$$

wobei wir zweimal die Riemann-Summen-Approximation eindimensionaler Integrale verwendet haben. Im Limes $K \rightarrow \infty$ werden die Fehlerterme immer kleiner und wie erhalten die erste behauptete Formel. Die zweite ergibt sich dann analog bzw. aus Symmetriegründen. \square

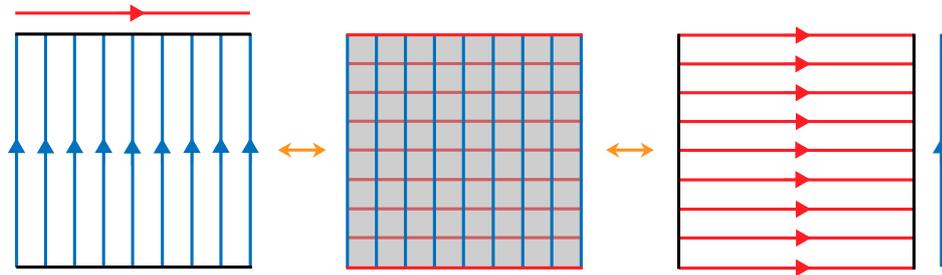


Abbildung Schematische Darstellung des Satzes von Fubini zur Berechnung von Integralen auf einem zweidimensionalen Quader (Grau), wobei Blau bzw. Rot die x_1 - bzw. x_2 -Richtung darstellen und die Pfeile jeweils eine eindimensionale Integration andeuten.

Bemerkungen

1. In dieser Vorlesung werden wir den Satz von Fubini nur für stetige Integranden f verwenden. Er gilt sinngemäß auch für jede andere integrierbare Funktion, obwohl dann gewisse Subtilitäten bei der Verschachtelungsformel auftreten können, die wir hier aber nicht diskutieren wollen.
2. Man schreibt auch gerne

$$\int_Q f(\mathbf{x}) \, d\mathbf{x} = \int_{a_2}^{b_2} \int_{a_1}^{b_1} f(x_1, x_2) \, dx_1 \, dx_2 = \int_{a_1}^{b_1} \int_{a_2}^{b_2} f(x_1, x_2) \, dx_2 \, dx_1.$$

Beachte aber, dass die Reihenfolge der Integralzeichen $\int_{a_j}^{b_j}$ und die Reihenfolge der dx_j konsistent (also gerade entgegengesetzt) sein müssen: Das Integralzeichen wirkt wie „Klammer auf“, das entsprechende dx_j wie „Klammer zu“.

3. Im Fall einer Produktdarstellung $f(x_1, x_2) = g_1(x_1) g_2(x_2)$ vereinfacht sich die Formel zu

$$\int_Q f(\mathbf{x}) \, d\mathbf{x} = \left(\int_{a_1}^{b_1} g_1(x_1) \, dx_1 \right) \cdot \left(\int_{a_2}^{b_2} g_2(x_2) \, dx_2 \right),$$

aber im Allgemeinen kann die Verschachtelung zweier Integrale *nicht* als das Produkt zweier Integrale geschrieben werden.

Beispiel Für $n = 2$, $Q = [a_1, b_1] \times [a_2, b_2]$ sowie

$$f(x_1, x_2) = \sin(x_1) + \cos(x_2)$$

gilt

$$\int_Q f(\mathbf{x}) \, d\mathbf{x} = \int_{a_1}^{b_1} \left(\int_{a_2}^{b_2} (\sin(x_1) + \cos(x_2)) \, dx_2 \right) dx_1.$$

Das innere Integral kann nun mit dem eindimensionalen Hauptsatz der Differential- und Integralrechnung — angewendet bzgl. der Integrationsvariable x_2 — zu

$$\begin{aligned} \int_{a_2}^{b_2} (\sin(x_1) + \cos(x_2)) \, dx_2 &= \left[\sin(x_1) x_2 + \sin(x_2) \right]_{x_2=a_2}^{x_2=b_2} \\ &= (b_2 - a_2) \sin(x_1) + \sin(b_2) - \sin(a_2) \end{aligned}$$

berechnet werden. Durch Einsetzen und nochmaliges Anwenden des Hauptsatzes — aber diesmal bzgl. x_1 — erhalten wir

$$\begin{aligned} \int_Q f(\mathbf{x}) \, d\mathbf{x} &= \int_{a_1}^{b_1} ((b_2 - a_2) \sin(x_1) + \sin(b_2) - \sin(a_2)) \, dx_1 \\ &= \left[-(b_2 - a_2) \cos(x_1) + (\sin(b_2) - \sin(a_2)) x_1 \right]_{x_1=a_1}^{x_1=b_1} \\ &= -(b_2 - a_2) (\cos(b_1) - \cos(a_1)) + (b_1 - a_1) (\sin(b_2) - \sin(a_2)). \end{aligned}$$

Alternativ können wir via

$$\int_Q f(\mathbf{x}) \, d\mathbf{x} = \int_{a_2}^{b_2} \left(\int_{a_1}^{b_1} (\sin(x_1) + \cos(x_2)) \, dx_1 \right) dx_2$$

auch x_1 als die innere und x_2 als die äußere Variable betrachten und erhalten, mit anderen Zwischenschritten, am Ende dasselbe Ergebnis.

Beispiele

1. Für den additiven Integranden

$$f(x, y, z) = 3x + 2y + z$$

liefert der Satz von Fubini

$$\begin{aligned} \iiint_{[0,1] \times [2,4] \times [5,8]} f(x, y, z) \, dx \, dy \, dz &= \int_5^8 \left(\int_2^4 \left(\int_0^1 (3x + 2y + z) \, dx \right) dy \right) dz \\ &= \int_5^8 \left(\int_2^4 \left(\frac{3}{2} + 2y + z \right) dy \right) dz \\ &= \int_5^8 (15 + 2z) \, dz = 45 + 39 = 84. \end{aligned}$$

Andere Verschachtelungen führen auch zum Ziel, zum Beispiel

$$\begin{aligned} \iiint_{[0,1] \times [2,4] \times [5,8]} f(x, y, z) \, dx \, dy \, dz &= \int_0^1 \left(\int_2^4 \left(\int_5^8 (3x + 2y + z) \, dz \right) dy \right) dx \\ &= \int_0^1 \left(\int_2^4 \left(9x + 6y + \frac{39}{2} \right) dy \right) dx \\ &= \int_0^1 (18x + 75) \, dx = 9 + 75 = 84. \end{aligned}$$

2. Für den multiplikativen Integranden

$$f(x, y, z) = x^3 y^2 z$$

erhalten wir

$$\begin{aligned} \iiint_{[0,1] \times [2,4] \times [5,8]} f(x, y, z) \, dx \, dy \, dz &= \left(\int_0^1 x^3 \, dx \right) \cdot \left(\int_2^4 y^2 \, dy \right) \cdot \left(\int_5^8 z \, dz \right) \\ &= \frac{1}{4} \cdot \frac{56}{3} \cdot \frac{39}{2} = 91 \end{aligned}$$

nach Vereinfachung der Verschachtelung und Berechnung der drei Integrale.

2.3 Exkurs*: Kurven und Kurvenintegrale

Ziel In diesem Abschnitt stellen wir einige Resultate über Kurven zusammen, da diese in dieser Vorlesung immer wieder eine wichtige Rolle spielen werden. Die meisten Konzepte und Formeln hatten wir schon in *Analysis 1+2* kennengelernt und müssen sie uns nur in Erinnerung rufen.

Definition Eine parametrisierte Kurve ist eine (meist mindestens zweimal stetig differenzierbare) Abbildung $\mathbf{c} : I \rightarrow \mathbb{R}^n$ auf einem Intervall $I \subset \mathbb{R}$. Die Punktmenge

$$\text{im}(\mathbf{c}) := \{\mathbf{c}(t) : t \in I\} \subset \mathbb{R}^n$$

wird auch Bild (oder *Image*) der parametrisierten Kurve genannt und ist das geometrische Objekt, das man landläufig *Kurve* nennt. Für jedes $t \in I$ wird

$$\frac{d}{dt}\mathbf{c}(t) = \dot{\mathbf{c}}(t) = \begin{pmatrix} \dot{c}_1(t) \\ \vdots \\ \dot{c}_n(t) \end{pmatrix}$$

der Tangentialvektor im Kurvenpunkt $\mathbf{c}(t)$ genannt und der Parameter t kann oftmals als Zeit interpretiert werden. Man nennt den Tangentialvektor dann den momentanen Geschwindigkeitsvektor.

Beispiel Die Abbildung $\mathbf{c} : \mathbb{R} \rightarrow \mathbb{R}^2$ mit

$$\mathbf{c}(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}$$

ist eine parametrisierte Kurve. Ihr Bild

$$\text{im}(\mathbf{c}) = \{\mathbf{x} = (x_1, x_2) : x_1^2 + x_2^2 = 1\} \subset \mathbb{R}^2$$

ist gerade die Einheitskreislinie in der Ebene. Wir sagen auch, die Abbildung \mathbf{c} parametrisiert die Kreislinie.

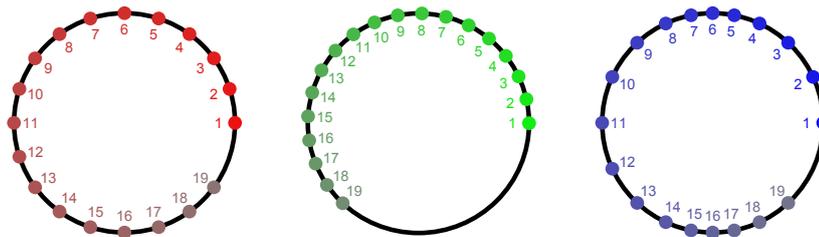


Abbildung Drei verschiedene Parametrisierungen \mathbf{c} der Kreislinie, die jeweils unterschiedlichen Durchlaufgeschwindigkeiten entsprechen. Die Punkte markieren den Kurvenpunkt $\mathbf{c}(t_k)$ zu äquidistant gewählten Zeiten $t_k = k \Delta t$.

Sprechweise In der Mathematik lässt man häufig das Attribut „parametrisiert“ weg und nennt sowohl die Abbildung \mathbf{c} als auch die Punktmenge $\text{im}(\mathbf{c})$ schlicht *Kurve*. Das kann gerade am Anfang einige Verwirrung stiften, aber in aller Regel wird durch den Kontext klar, ob man Kurve in dem einen oder dem anderen Sinne meint. Eine ähnliche Zweideutigkeit werden wir beim Studium der Flächen antreffen.

Alternative Notationen In den Anwendungswissenschaften und in der Geometrie bezeichnet man die Abbildung \mathbf{c} oftmals nicht explizit, sondern benutzt dieselben Buchstaben wie für Komponenten von Vektoren in \mathbb{R}^n . Man schreibt also einfach

$$\mathbf{x}(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} \quad \text{und} \quad \dot{\mathbf{x}}(t) = \begin{pmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \end{pmatrix}$$

um deutlich zu machen, dass die räumlichen Koordinaten x_j sich entlang der Kurve mit dem Parameter t ändern. Diese Notation ist sehr intuitiv, benutzt aber die x_j in zweifacher Bedeutung, nämlich einmal als Koordinate in \mathbb{R}^n und einmal als Komponenten einer Funktion $I \rightarrow \mathbb{R}^n$). Eine Mischform der Notation ist $x_j = c_j(t)$. Wie immer kann man immer mit beiden Notationen arbeiten, wobei aber je nach Kontext und Vorliebe mal die eine und mal die andere „besser geeignet“ ist.

Reparametrisierung von Kurven Eine parametrisierte Kurve $\tilde{\mathbf{c}} : \tilde{I} \rightarrow \mathbb{R}^n$ wird Reparametrisierung von $\mathbf{c} : I \rightarrow \mathbb{R}^n$ genannt, wenn es eine bijektive, d.h. invertierbare Abbildung $h : \tilde{I} \rightarrow I$ (der sogenannte Parameterwechsel) gibt, so dass

$$\tilde{\mathbf{c}}(\tilde{t}) = \mathbf{c}(h(\tilde{t}))$$

für alle $\tilde{t} \in \tilde{I}$ gilt. Insbesondere gilt $\text{im}(\tilde{\mathbf{c}}) = \text{im}(\mathbf{c})$, d.h. $\tilde{\mathbf{c}}$ und \mathbf{c} beschreiben bzw. parametrisieren dasselbe geometrische Objekt, aber mittels der verschiedenen Parameter \tilde{t} und $t = h(\tilde{t})$.

Beispiel Die Abbildung $\tilde{\mathbf{c}} : \mathbb{R} \rightarrow \mathbb{R}^2$ mit

$$\tilde{\mathbf{c}}(\tilde{t}) = \begin{pmatrix} \cos(\omega \tilde{t}) \\ \sin(\omega \tilde{t}) \end{pmatrix}$$

ist eine Reparametrisierung der im ersten Beispiel angegebenen Kurve \mathbf{c} , wobei die Kreislinie nun mit konstanter Winkelgeschwindigkeit ω durchlaufen wird. Der entsprechende Reparameterwechsel ist durch $t = h(\tilde{t}) = \omega \tilde{t}$ gegeben. Die alternative Formel

$$\tilde{\mathbf{c}}(\tilde{t}) = \begin{pmatrix} \cos(t + \mu_1 \sin(\mu_2 \tilde{t})) \\ \sin(t + \mu_1 \sin(\mu_2 \tilde{t})) \end{pmatrix}$$

mit den zwei Konstanten $-1 < \mu_1 < +1$ und $\mu_2 \in \mathbb{R}$ liefert eine andere Parametrisierung der Kreislinie bzw. eine Reparametrisierung von \mathbf{c} , wobei diesmal die Winkelgeschwindigkeit nicht mehr konstant ist, sondern selbst variiert.

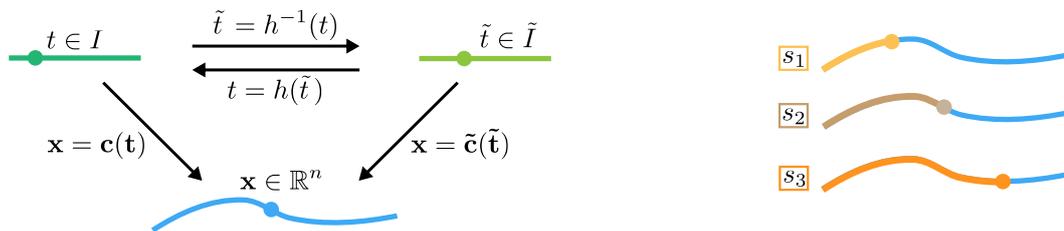


Abbildung Links: Illustration der Konzepte *Parametrisierung* und *Reparametrisierung* von Kurven. Rechts: Zum Bogenlängenparameter (meist s genannt): Jeder Kurvenpunkt kann eindeutig durch eine entsprechende Länge charakterisiert werden.

Länge von Kurven und Kurvenintegrale Wir hatten in *Analysis 2* gesehen, dass die Länge einer parametrisierten Kurve $\mathbf{c} : [a, b] \rightarrow \mathbb{R}^n$ durch

$$L(\mathbf{c}) = \int_a^b \left| \frac{d}{dt} \mathbf{c}(t) \right| dt$$

berechnet werden kann. Ist $\tilde{\mathbf{c}} : [\tilde{a}, \tilde{b}] \rightarrow \mathbb{R}$ eine Reparametrisierung von \mathbf{c} mit bijektivem Parameterwechsel $h : [\tilde{a}, \tilde{b}] \rightarrow [a, b]$, so implizieren die Kettenregel sowie die Transformationsformel für Integrale die Identität

$$L(\tilde{\mathbf{c}}) = \int_{\tilde{a}}^{\tilde{b}} \left| \frac{d}{d\tilde{t}} \tilde{\mathbf{c}}(\tilde{t}) \right| d\tilde{t} = \int_a^b \left| \frac{d}{dt} \mathbf{c}(t) \right| dt = L(\mathbf{c}).$$

Insbesondere ist die Länge *invariant unter Reparametrisierung* und damit eine *geometrische Größe*, d.h. eigentlich eine Eigenschaft der Punktmenge $\text{im}(\mathbf{c})$. Man braucht aber die Abbildung bzw. Parametrisierung \mathbf{c} , um die Länge des geometrischen Objektes überhaupt ausrechnen zu können.

Bemerkung Mit einem analogen Beweis zeigen wir, dass das erste Kurvenintegral

$$\int_{\mathbf{c}} f(\mathbf{x}) \, ds := \int_a^b f(\mathbf{c}(t)) \left| \frac{d}{dt} \mathbf{c}(t) \right| dt$$

sowie das zweite Kurvenintegral

$$\int_{\mathbf{c}} f(\mathbf{x}) \cdot d\mathbf{x} := \int_a^b \mathbf{f}(\mathbf{c}(t)) \cdot \frac{d}{dt} \mathbf{c}(t) dt$$

invariant unter Reparametrisierung sind. Hierbei ist f bzw. \mathbf{f} ein stetiges Skalarfeld bzw. Vektorfeld auf \mathbb{R}^n und beim Kurvenintegral zweiter Art muss zusätzlich gefordert werden, dass die Funktion h monoton wachsend ist, denn andernfalls ändert sich der Durchlaufsinne sowie das Vorzeichen des Integrals. Beachte, dass in beiden Fällen auf der rechten Seite eine skalare Funktion auf einem Intervall integriert wird und wie daher alle Erkenntnisse aus *Analysis 1* verwenden dürfen.

Bogenlängenparameter Ist $\mathbf{c} : [a, b] \rightarrow \mathbb{R}^n$ eine gegebene stetig differenzierbare parametrisierte Kurve, so wird durch

$$l(t) := \int_a^t |\dot{\mathbf{c}}(\tau)| \, d\tau$$

eine strikt monoton wachsende und stetig differenzierbare Funktion l definiert. Diese bildet das Intervall $[a, b]$ bijektiv auf das Intervall $[0, L(\mathbf{c})]$ ab und wird Bogenlängenfunktion von \mathbf{c} genannt. Ihre Ableitung

$$\ell(t) := \dot{l}(t) = |\dot{\mathbf{c}}(t)|$$

heißt infinitesimales Längenelement von \mathbf{c} . Gilt $\ell(t) = 1$ für alle $t \in [a, b]$, so sagt man, \mathbf{c} ist nach Bogenlänge parametrisiert. Parametrisierte Kurven mit dieser Eigenschaft sind besonders wichtig und nützlich.

Lemma (spezielle Parametrisierung) Für jede parametrisierte Kurve existiert eine Reparametrisierung nach Bogenlänge.

Beweis Für eine parametrisierte Kurve $\mathbf{c} : [a, b] \rightarrow \mathbb{R}^n$ kann die Bogenlängenfunktion l invertiert werden und mit Hilfe der Umkehrabbildung l^{-1} definieren wir

$$\mathbf{d} : [0, L(\mathbf{c})] \rightarrow \mathbb{R}^n \quad \text{mit} \quad \mathbf{d}(s) = \mathbf{c}(l^{-1}(s))$$

als Reparametrisierung von \mathbf{c} , wobei s mit $0 \leq s \leq L(\mathbf{c})$ gerade der Bogenlängenparameter ist. Mit $t = l^{-1}(s)$ bzw. $s = l(t)$ und der symbolischen Kettenregel gilt

$$\frac{d\mathbf{d}}{ds} = \frac{dt}{ds} \frac{d\mathbf{c}}{dt} = \left(\frac{ds}{dt} \right)^{-1} \frac{d\mathbf{c}}{dt} = \ell^{-1} \frac{d\mathbf{c}}{dt}$$

und damit $\left| \frac{d}{ds} \mathbf{d}(s) \right| = 1$ für alle s . Die Reparametrisierung \mathbf{d} ist also in der Tat nach Bogenlänge parametrisiert. \square

Bemerkung: In Physiknotation ist $ds = \ell dt = |\dot{\mathbf{x}}| dt$ das Gesetz für die zeitliche Änderung des Bogenlängenparameters s und taucht in der symbolischen Schreibweise des ersten Kurvenintegrals auf.

2.4 Integralsätze auf zweidimensionalen Quadern

Rand zweidimensionaler Quader Der Rand des Quaders

$$Q = [a_1, b_1] \times [a_2, b_2]$$

besteht aus vier Streckensegmenten und kann auch als Bild der parametrisierten Kurve

$$\mathbf{c}(t) = \begin{pmatrix} c_1(t) \\ c_2(t) \end{pmatrix} := \begin{cases} \begin{pmatrix} a_1 + t - t_0 \\ a_2 \end{pmatrix} & \text{für } t_0 \leq t \leq t_1 \\ \begin{pmatrix} b_1 \\ a_2 + t - t_1 \end{pmatrix} & \text{für } t_1 \leq t \leq t_2 \\ \begin{pmatrix} b_1 - t + t_2 \\ b_2 \end{pmatrix} & \text{für } t_2 \leq t \leq t_3 \\ \begin{pmatrix} a_1 \\ b_2 - t + t_3 \end{pmatrix} & \text{für } t_3 \leq t \leq t_4 \end{cases}$$

interpretiert werden kann. Die *Zeitpunkte* t_0, t_1, t_2, t_3, t_4 genügen dabei den Formeln

$$t_1 - t_0 = t_3 - t_2 = b_1 - a_1, \quad t_2 - t_1 = t_4 - t_3 = b_2 - a_2$$

bzw.

$$\mathbf{c}(t_0) = \mathbf{c}(t_4) = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \quad \mathbf{c}(t_1) = \begin{pmatrix} b_1 \\ a_2 \end{pmatrix}, \quad \mathbf{c}(t_2) = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \quad \mathbf{c}(t_3) = \begin{pmatrix} a_1 \\ b_2 \end{pmatrix}$$

und wir verifizieren

$$\frac{d}{dt} \mathbf{c}(t) = \dot{\mathbf{c}}(t) = \begin{cases} \begin{pmatrix} +1 \\ 0 \end{pmatrix} & \text{für } t_0 \leq t \leq t_1 \\ \begin{pmatrix} 0 \\ +1 \end{pmatrix} & \text{für } t_1 \leq t \leq t_2 \\ \begin{pmatrix} -1 \\ 0 \end{pmatrix} & \text{für } t_2 \leq t \leq t_3 \\ \begin{pmatrix} 0 \\ -1 \end{pmatrix} & \text{für } t_3 \leq t \leq t_4 \end{cases}$$

durch einfache Rechnungen.

Bemerkung: Diese Parametrisierungen sehen auf den ersten Blick kompliziert aus, beschreiben aber eine sehr einfache Idee: Ein gedachter Punkt bewegt sich mit konstantem Geschwindigkeitsbetrag auf dem Rand des Quaders. Er startet zur Zeit t_0 in der linken unteren Ecke mit Koordinaten (a_1, a_2) und bewegt sich zunächst nach rechts, bis er zur Zeit t_1 in der rechten unteren Ecke (b_1, a_2) ankommt. Danach läuft er nach oben bis zur Zeit t_2 , zu der er die obere rechte Ecke (b_1, b_2) erreicht hat. Anschließend bewegt er sich solange nach links, bis er zur Zeit t_3 in der linken oberen Ecke (a_1, b_2) zu finden ist. Dort wendet es sich nach unten und kehrt zur Zeit t_4 schließlich zu seinem Ausgangspunkt in der linken unteren Ecke zurück.

Vektorfelder auf dem Rand zweidimensionaler Quader In jedem Punkt \mathbf{x} des Randes von Q (mit Ausnahme der vier Eckpunkte) können wir nun durch

$$\boldsymbol{\tau}(\mathbf{x}) := \begin{cases} \begin{pmatrix} +1 \\ 0 \end{pmatrix} & \text{für } x_2 = a_2 \\ \begin{pmatrix} 0 \\ +1 \end{pmatrix} & \text{für } x_1 = b_1 \\ \begin{pmatrix} -1 \\ 0 \end{pmatrix} & \text{für } x_2 = b_2 \\ \begin{pmatrix} 0 \\ -1 \end{pmatrix} & \text{für } x_1 = a_1 \end{cases} \quad \boldsymbol{\nu}(\mathbf{x}) := \begin{cases} \begin{pmatrix} 0 \\ -1 \end{pmatrix} & \text{für } x_2 = a_2 \\ \begin{pmatrix} +1 \\ 0 \end{pmatrix} & \text{für } x_1 = b_1 \\ \begin{pmatrix} 0 \\ +1 \end{pmatrix} & \text{für } x_2 = b_2 \\ \begin{pmatrix} -1 \\ 0 \end{pmatrix} & \text{für } x_1 = a_1 \end{cases}$$

einen (positiv orientierten und normierten) Tangentialvektor $\boldsymbol{\tau}(\mathbf{x})$ sowie einen (nach außen zeigenden und ebenfalls normierten) Normalenvektor $\boldsymbol{\nu}(\mathbf{x})$ definieren. Diese beiden Vektorfelder leben auf dem Rand des Quaders und können unabhängig von einer Parametrisierung eingeführt werden. Für jede Parametrisierung werden allerdings die *Konsistenzbedingungen*

$$\dot{\mathbf{c}}(t) \parallel \boldsymbol{\tau}(\mathbf{c}(t)), \quad \dot{\mathbf{c}}(t) \perp \boldsymbol{\nu}(\mathbf{c}(t))$$

für (fast) alle t erfüllt sein (mit Ausnahme der Zeiten, die Eckpunkten entsprechen), sodass $\boldsymbol{\tau}(\mathbf{c}(t))$, $\boldsymbol{\nu}(\mathbf{c}(t))$ mit dem Frenetschen Zweibein der parametrisierten Kurve zusammenfallen. Für die oben angegebene Parametrisierung gilt sogar $\dot{\mathbf{c}}(t) = \boldsymbol{\tau}(\mathbf{c}(t))$.

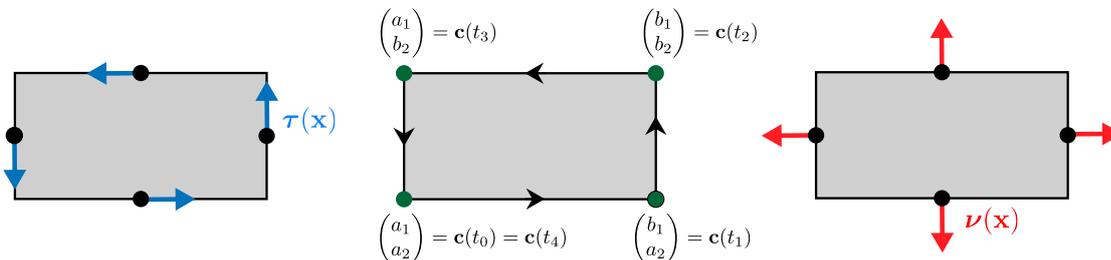


Abbildung *Mitte:* Bei der angegebenen Parametrisierung des Randes läuft ein gedachter Punkt mit konstantem Geschwindigkeitsbetrag die vier Seiten des Rechtecks entlang, wobei die Parametrisierung in den Ecken (grün) nicht differenzierbar ist. *Links:* Der Tangentialvektor in vier ausgewählten Punkten des Randes. *Rechts:* Der Normalenvektor in denselben Randpunkten.

Divergenz und Rotation in zwei Dimensionen Für ein zweidimensionales Vektorfeld \mathbf{f} gilt

$$\begin{aligned} \operatorname{div} \mathbf{f}(x_1, x_2) &:= \partial_{x_1} f_1(x_1, x_2) + \partial_{x_2} f_2(x_1, x_2), \\ \operatorname{rot} \mathbf{f}(x_1, x_2) &:= \partial_{x_1} f_2(x_1, x_2) - \partial_{x_2} f_1(x_1, x_2). \end{aligned}$$

Bemerkung: Wir werden uns die geometrische und physikalische Bedeutung dieser Größen schrittweise erarbeiten, wobei den Integralsätzen von Gauß und Stokes eine wichtige Rolle zufallen wird.

Theorem (Satz von Gauß für zweidimensionale Quader) Für jedes stetig differenzierbare Vektorfeld $\mathbf{f} : Q \subset \mathbb{R}^2 \rightarrow \mathbb{R}^2$ gilt

$$\int_Q \operatorname{div} \mathbf{f}(\mathbf{x}) \, d\mathbf{x} = \int_{\mathbf{c}} \langle \mathbf{f}(\mathbf{x}), \boldsymbol{\nu}(\mathbf{x}) \rangle \, ds$$

d.h. das Gebietsintegral des Skalarfeldes $\operatorname{div} \mathbf{f}$ über Q kann als ein Kurvenintegral erster Art berechnet werden. Dabei ist \mathbf{c} irgendeine Parametrisierung des Randes von Q (zum Beispiel die oben angegebene), $\boldsymbol{\nu}$ das nach außen zeigende normale Vektorfeld auf dem Rand von Q und $\langle \cdot, \cdot \rangle$ das Skalarprodukt im \mathbb{R}^2 .

Beweis Berechnung der linken Seite: Mit dem Satz von Fubini sowie dem Hauptsatz der Differential- und Integralrechnung gilt

$$\begin{aligned} \int_Q \partial_{x_1} f_1(x_1, x_2) \, dx_1 \, dx_2 &= \int_{a_2}^{b_2} \left(\int_{a_1}^{b_1} \partial_{x_1} f_1(x_1, x_2) \, dx_1 \right) dx_2 \\ &= \int_{a_2}^{b_2} \left(f_1(b_1, x_2) - f_1(a_1, x_2) \right) dx_2 \\ &= \int_{a_2}^{b_2} f_1(b_1, x_2) \, dx_2 - \int_{a_2}^{b_2} f_1(a_1, x_2) \, dx_2 \end{aligned}$$

und analog (bzw. durch konsequente Vertauschung der Indizes 1 und 2) zeigen wir

$$\int_Q \partial_{x_2} f_2(x_1, x_2) \, dx_1 \, dx_2 = \int_{a_1}^{b_1} f_2(x_1, b_2) \, dx_1 - \int_{a_1}^{b_1} f_2(x_1, a_2) \, dx_1.$$

Durch Addition beider Formeln können wir das Gebietsintegral von $\operatorname{div} \mathbf{f}$ als Summe von vier eindimensionalen Integralen darstellen.

Berechnung der rechten Seite: Das Kurvenintegral über den Rand von Q kann via

$$\int_{\mathbf{c}} \dots \, ds = \int_{\mathbf{c}_u} \dots \, ds + \int_{\mathbf{c}_r} \dots \, ds + \int_{\mathbf{c}_o} \dots \, ds + \int_{\mathbf{c}_l} \dots \, ds$$

als Summe von vier Teilintegralen dargestellt werden, die dem unteren (u), rechten (r), oberen (o) und linken (l) Geradenstück entsprechen und für die wir jeweils die entsprechenden Teilformeln von oben verwenden können. Für das untere Stück gilt

$$\mathbf{c}_u(t) = \begin{pmatrix} a_1 + t - t_0 \\ a_2 \end{pmatrix}, \quad \dot{\mathbf{c}}_u(t) = \begin{pmatrix} +1 \\ 0 \end{pmatrix}, \quad \boldsymbol{\nu}(\mathbf{c}_u(t)) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \quad t \in [t_0, t_1]$$

und die Definition des Kurvenintegrals impliziert

$$\begin{aligned} \int_{\mathbf{c}_u} \langle \mathbf{f}(\mathbf{x}), \boldsymbol{\nu}(\mathbf{x}) \rangle \, ds &= \int_{t_0}^{t_1} \langle \mathbf{f}(\mathbf{c}_u(t)), \boldsymbol{\nu}(\mathbf{c}_u(t)) \rangle |\dot{\mathbf{c}}_u(t)| \, dt \\ &= - \int_{t_0}^{t_1} f_2(a_1 + t - t_0, a_2) \, dt = - \int_{a_1}^{b_1} f_2(x_1, a_2) \, dx_1, \end{aligned}$$

wobei die letzte Gleichheit aus dem Transformationssatz für eindimensionale Integrale mittels der Substitution $x_1 = a_1 + t - t_0$ sowie der Bedingung $t_1 - t_0 = b_1 - a_1$ folgt. Für das rechte Randstück folgt wegen

$$\mathbf{c}_r(t) = \begin{pmatrix} b_1 \\ a_2 + t - t_1 \end{pmatrix}, \quad \dot{\mathbf{c}}_r(t) = \begin{pmatrix} 0 \\ +1 \end{pmatrix}, \quad \boldsymbol{\nu}(\mathbf{c}_r(t)) = \begin{pmatrix} +1 \\ 0 \end{pmatrix}, \quad t \in [t_1, t_2]$$

mit ähnlichen Argumenten die Formel

$$\begin{aligned} \int_{\mathbf{c}_r} \langle \mathbf{f}(\mathbf{x}), \boldsymbol{\nu}(\mathbf{x}) \rangle ds &= \int_{t_1}^{t_2} \langle \mathbf{f}(\mathbf{c}_r(t)), \boldsymbol{\nu}(\mathbf{c}_r(t)) \rangle |\dot{\mathbf{c}}_r(t)| dt \\ &= + \int_{t_1}^{t_2} f_1(b_1, a_2 + t - t_1) dt = + \int_{a_2}^{b_2} f_1(b_1, x_2) dx_2. \end{aligned}$$

Analog zeigen wir schließlich

$$\int_{\mathbf{c}_o} \langle \mathbf{f}(\mathbf{x}), \boldsymbol{\nu}(\mathbf{x}) \rangle ds = + \int_{t_2}^{t_3} f_2(b_1 - t + t_2, b_2) dt = + \int_{a_1}^{b_1} f_2(x_1, b_2) dx_1.$$

bzw.

$$\int_{\mathbf{c}_l} \langle \mathbf{f}(\mathbf{x}), \boldsymbol{\nu}(\mathbf{x}) \rangle ds = - \int_{t_3}^{t_4} f_1(a_1, b_2 - t + t_3) dt = - \int_{a_2}^{b_2} f_1(a_1, x_2) dx_2.$$

für den oberen bzw. den linken Teilweg. Insgesamt erhalten wir hier dieselben vier Terme wie bei der Berechnung des Gebietsintegrals im ersten Beweisschritt. \square

Theorem (Satz von Stokes für zweidimensionale Quader) Für jedes stetig differenzierbare Vektorfeld $\mathbf{f} : Q \subset \mathbb{R}^2 \rightarrow \mathbb{R}^2$ gilt

$$\int_Q \operatorname{rot} \mathbf{f}(\mathbf{x}) dx = \int_{\mathbf{c}} \langle \mathbf{f}(\mathbf{x}), \boldsymbol{\tau}(\mathbf{x}) \rangle ds,$$

wobei $\boldsymbol{\tau}$ das positiv orientierte tangentialen Vektorfeld auf dem Rand von Q ist.

Beweis Analog zum Beweis des Satzes von Gauß, siehe Hausaufgabe. \square

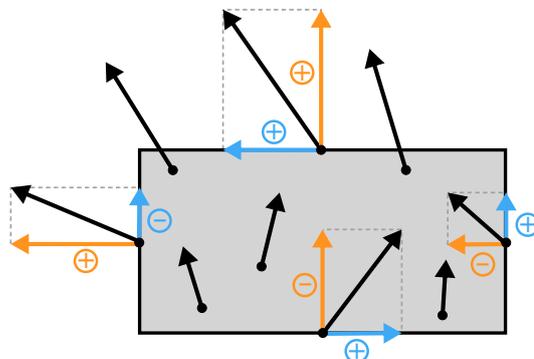


Abbildung Aufspaltung des Vektors $\mathbf{f}(\mathbf{x})$ (schwarz) am Rand in einen tangentialen (hellblau, parallel zu $\boldsymbol{\tau}(\mathbf{x})$) und einen normalen Anteil (orange, parallel zu $\boldsymbol{\nu}(\mathbf{x})$), dargestellt für vier Punkten des Randes. Die Kurvenintegrale im Satz von Stokes bzw. Gauß quantifizieren die mittleren vorzeichenbehafteten Längen dieser Anteile, die durch die Skalarprodukte $\langle \mathbf{f}(\mathbf{x}), \boldsymbol{\tau}(\mathbf{x}) \rangle$ und $\langle \mathbf{f}(\mathbf{x}), \boldsymbol{\nu}(\mathbf{x}) \rangle$ gegeben sind.

Interpretation Jedes Vektorfeld auf einem zweidimensionalen Quader kann in jedem Randpunkt \mathbf{x} (mit Ausnahme der Eckpunkte) via

$$\mathbf{f}(\mathbf{x}) = \langle \mathbf{f}(\mathbf{x}), \boldsymbol{\tau}(\mathbf{x}) \rangle \boldsymbol{\tau}(\mathbf{x}) + \langle \mathbf{f}(\mathbf{x}), \boldsymbol{\nu}(\mathbf{x}) \rangle \boldsymbol{\nu}(\mathbf{x})$$

eindeutig in einen tangentialen und einen normalen Anteil aufgespalten werden (wobei wir für diese Formel benutzt haben, dass $\boldsymbol{\tau}(\mathbf{x})$ und $\boldsymbol{\nu}(\mathbf{x})$ zueinander senkrecht stehen und beide die Länge 1 haben). Das Kurvenintegral im Satz von Gauß gibt nun den effektiven, d.h. über den Rand integrierten, „Ausfluss“ bzw. „Einfluss“ des Vektorfeldes an, wobei auslaufende bzw. einlaufende Randvektoren positive bzw. negative Beiträge liefern. Das Kurvenintegral im Satz von Stokes gibt hingegen einen „Umlaufwert“ an, wobei auch wieder Vorzeichen eine Rolle spielen. Die Details der Interpretation werden natürlich vom physikalischen Kontext abhängen. Zum Beispiel entspricht ein Vektorfeld in der Hydrodynamik dem Geschwindigkeitsfeld einer Flüssigkeit. In der Maxwellschen Theorie beschreiben Vektorfelder jedoch die elektrischen oder magnetischen Kräfte, die in jedem Punkt der Ebene ($n = 2$) oder des Raumes ($n = 3$) auf ein gedachtes Probeteilchen mit Einheitsladung wirken.

Spezialfall horizontale Vektorfelder Wir betrachten die zwei Vektorfelder

$$\text{(HH)} \quad \mathbf{f}(\mathbf{x}) = \begin{pmatrix} \phi(x_1) \\ 0 \end{pmatrix}, \quad \text{(HV)} \quad \mathbf{f}(\mathbf{x}) = \begin{pmatrix} \phi(x_2) \\ 0 \end{pmatrix},$$

wobei $\phi : \mathbb{R} \rightarrow \mathbb{R}$ eine gegebene und stetig differenzierbare Funktion ist (vielleicht $\phi(s) = c_1 s + c_0$). Beide Vektorfelder sind *horizontal* (da f_2 jeweils verschwindet) und (im Allgemeinen) nicht-konstant, wobei das erste bzw. zweite Vektorfeld sich nur in horizontaler bzw. vertikaler Richtung ändert (da f_1 nur von x_1 bzw. x_2 abhängt). Für diese einfachen Vektorfelder können die Kurvenintegrale in den Sätzen von Gauß und Stokes problemlos berechnet werden (siehe Hausaufgaben) und vermitteln ein intuitives Verständnis für die zu Grunde liegenden Prinzipien.

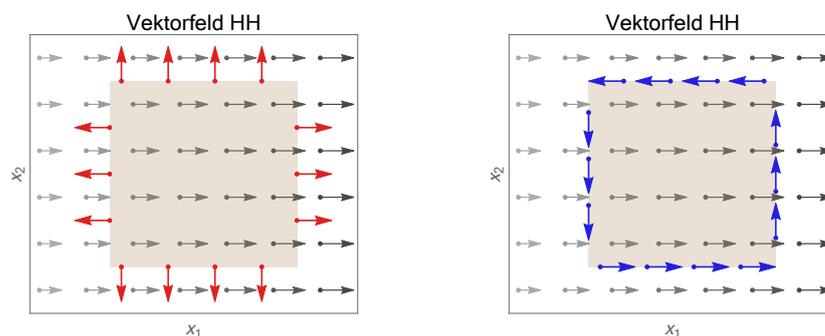


Abbildung *Links:* Der untere und der obere Rand des braunen Quaders liefern in diesem Beispiel beide keinen (bzw. einen verschwindenden) Beitrag zum Gaußschen Kurvenintegral. Die Beiträge vom linken und vom rechten Rand unterscheiden sich aber sowohl im Vorzeichen als auch im Betrag — der „Ausfluss“ rechts übersteigt den „Einfluss“ links — und das Gebietsintegral der Divergenz liefert daher einen positiven Wert („Quelle“). Die roten Vektoren repräsentieren $\boldsymbol{\nu}(\mathbf{x})$. *Rechts:* Die nicht-verschwindenden Einzelbeiträge zum Stokeschen Kurvenintegral stammen vom oberen und vom unteren Rand und addieren sich hier zu 0, da sie den gleichen Betrag, aber entgegengesetzte Vorzeichen besitzen. Das Gebietsintegral der Rotation wird daher auch verschwinden. Die blauen Vektoren stellen $\boldsymbol{\tau}(\mathbf{x})$ dar.

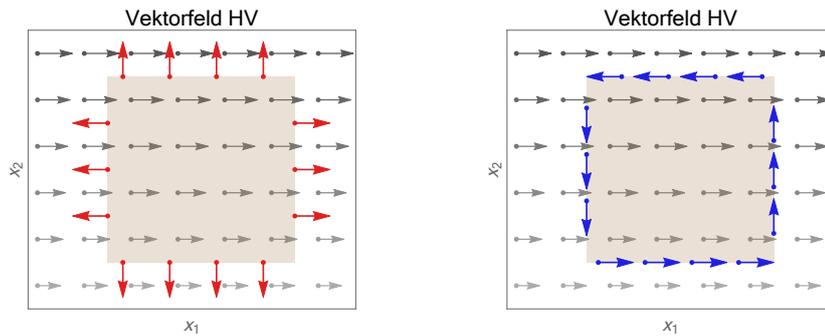


Abbildung Links: Auch in diesem Beispiel tragen der obere und untere Rand des Quaders nicht zum Gaußschen Kurvenintegral bei, aber diesmal heben sich auch die Beiträge vom linken und vom rechten Rand gegenseitig auf („Einfluss“ gleich „Ausfluss“). Das Gebietsintegral der Divergenz muss daher verschwinden. Rechts: Der linke und der rechte Rand liefern wieder keinen Beitrag zum Stokesschen Kurvenintegral, aber der negative Beitrag vom oberen Rand hat diesmal einen größeren Betrag als der positive Beitrag vom unteren Rand. Das Gebietsintegral der Rotation wird daher negativ sein und eine effektive „Umwirbelung“ des Quaders durch das Vektorfeld im Uhrzeigersinn beschreiben.

Bemerkungen

1. Wir haben in diesem Abschnitt nur die jeweils einfachste Variante der Integral-sätze von Gauß und Stokes kennengelernt. Beide gelten in viel allgemeineren Zusammenhängen (von denen wir einige noch studieren werden) und spielen eine fundamentale Rolle in allen Naturwissenschaften, zum Beispiel in der Maxwell'schen Theorie der elektromagnetischen Felder.
2. Die Sätze von Gauß und Stokes erlauben es uns, die Bedeutung der Differentialoperatoren div und rot bzw. **rot** besser zu verstehen. Auch das werden wir weiter unten genauer diskutieren. Sie stellen außerdem jeweils eine höherdimensionale Verallgemeinerung des Hauptsatzes der Differential- und Integralrechnung dar.
3. Die oben angegebene zweidimensionale Version des Satzes von Stokes wird auch Satz von Green genannt. Er wird in den Anwendungswissenschaften oftmals in der Form

$$\iint_Q (\partial_{x_1} f_2(x_1, x_2) - \partial_{x_2} f_1(x_1, x_2)) dx_1 dx_2 = \int_{\mathbf{c}} (f_1(x_1, x_2) dx_1 + f_2(x_1, x_2) dx_2)$$

angegeben, wobei die linke Seite gerade das Gebietsintegral von $\operatorname{rot} \mathbf{f}$ ist. Die rechte Seite stellt eine alternative Notation für das Stokessche Kurvenintegral dar, wobei bei der Berechnung des Integrals die konsistenten Substitutionen

$$x_j = c_j(t), \quad dx_j = \dot{c}_j(t) dt, \quad f_j(x_1, x_2) dx_j = f_j(c_1(t), c_2(t)) \dot{c}_j(t) dt$$

einzusetzen sind und anschließend die t -Integration über das Parametrisierungsintervall durchgeführt wird.

Theorem (partielle Integration auf zweidimensionalen Quadern) Für zwei stetig differenzierbare skalare Funktionen $f, \tilde{f} : Q \rightarrow \mathbb{R}$ gilt

$$\int_Q (\partial_{x_j} f(\mathbf{x})) \tilde{f}(\mathbf{x}) d\mathbf{x} = - \int_Q f(\mathbf{x}) (\partial_{x_j} \tilde{f}(\mathbf{x})) d\mathbf{x} + \int_{\mathbf{c}} f(\mathbf{x}) \tilde{f}(\mathbf{x}) \nu_j(\mathbf{x}) ds$$

für $j = 1$ und $j = 2$, wobei ν_j die j -te Komponente des nach außen zeigenden normalen Vektorfeldes $\boldsymbol{\nu}$ ist.

Beweis Man kann dies wieder im Detail nachrechnen. Oder etwas eleganter wie folgt ableiten: Für $j = 1$ betrachten wir das Vektorfeld

$$\mathbf{f}(x_1, x_2) = \begin{pmatrix} f_1(x_1, x_2) \\ f_2(x_1, x_2) \end{pmatrix} := \begin{pmatrix} f(x_1, x_2) \tilde{f}(x_1, x_2) \\ 0 \end{pmatrix}$$

mit verschwindender zweiter Komponente und erhalten

$$\int_Q \operatorname{div} \mathbf{f}(\mathbf{x}) \, d\mathbf{x} = \int_Q \partial_{x_1} (f(\mathbf{x}) \tilde{f}(\mathbf{x})) \, d\mathbf{x}$$

sowie

$$\int_{\mathbf{c}} \langle \mathbf{f}(\mathbf{x}), \boldsymbol{\nu}(\mathbf{x}) \rangle \, ds = \int_{\mathbf{c}} f(\mathbf{x}) \tilde{f}(\mathbf{x}) \nu_1(\mathbf{x}) \, ds.$$

Die erste Behauptung folgt nun direkt aus dem Satz von Gauß sowie der Produktregel für partielle Ableitungen. Die Formel für $j = 2$ kann analog abgeleitet werden. \square

2.5 Gebietsintegrale auf kompakten Mengen

Ziel Wir wollen Gebietsintegrale nicht nur für Quader, sondern für allgemeinere Mengen einführen. Ein Teilaspekt wird dabei die Berechnung des n -dimensionalen Volumens von Mengen sein, wobei der Volumenbegriff in der Riemannschen Theorie gerade dem Jordan-Inhalt entspricht.

Erinnerung Eine Menge $D \subset \mathbb{R}^n$ ist kompakt, falls sie beschränkt und abgeschlossen ist. Letzteres meint, dass jeder Randpunkt von D Element von D (und nicht von $\mathbb{R}^n \setminus D$) ist.

Bemerkung Eine Menge $D \subset \mathbb{R}^n$ ist genau dann beschränkt, wenn $D \subseteq Q$ für jeden hinreichend großen Quader Q gilt.

Definition Eine Funktion $f : D \rightarrow \mathbb{R}$ heißt (Riemann)-integrierbar (auf der Menge D), falls die triviale Fortsetzung von f , d.h. die Funktion $f^* : \mathbb{R}^n \rightarrow \mathbb{R}$ mit

$$f^*(\mathbf{x}) = \begin{cases} f(\mathbf{x}) & \text{für } \mathbf{x} \in D \\ 0 & \text{für } \mathbf{x} \notin D \end{cases}$$

auf jedem Quader Q mit $D \subset Q$ (Riemann)-integrierbar ist. Wir schreiben dann

$$\int_D f(\mathbf{x}) \, d\mathbf{x} = \int_Q f^*(\mathbf{x}) \, d\mathbf{x},$$

wobei der Wert des Integrals unabhängig von der Wahl von Q sein wird (solange $D \subseteq Q$ gilt).

Bemerkung Diese Definition ist zwar aus theoretischer Sicht sehr wichtig, in der Praxis aber oftmals unbrauchbar, weil die Berechnung von Ober- und Untersummen zu f^* in der Regel sehr kompliziert ist. Wir werden unten hinreichende Kriterien und Berechnungsvorschriften kennenlernen, die wesentlich handhabbarer sind. Zunächst müssen wir aber das n -dimensionale Volumen von Teilmengen des \mathbb{R}^n besser verstehen.

Charakteristische Funktion Für jede Menge $D \subset \mathbb{R}^n$ wird durch

$$\chi_D(\mathbf{x}) = \begin{cases} 1 & \text{für } \mathbf{x} \in D \\ 0 & \text{für } \mathbf{x} \notin D \end{cases}$$

ihre charakteristische Funktion $\chi_D : \mathbb{R}^n \rightarrow \mathbb{R}$ definiert.

Exkurs*: Jordan-Inhalt von Teilmengen des \mathbb{R}^n

Definition Die kompakte Menge $D \subset \mathbb{R}^n$ heißt (im Riemannschen Sinne) messbar, wenn χ_D auf jedem Quader Q mit $D \subseteq Q$ integrierbar ist. In diesem Fall nennen wir

$$\text{vol}(D) := \int_Q \chi_D(\mathbf{x}) \, d\mathbf{x}$$

das n -dimensionale Volumen von D , wobei der Wert nicht von der Wahl von Q abhängen wird. Gilt $\text{vol}(D) = 0$, so wird D auch Nullmenge genannt.

Bemerkung

1. Ein ein- bzw. zweidimensionales Volumen wird auch *Länge* bzw. *Flächeninhalt* genannt.
2. Messbare Mengen im Sinne der Definition immer kompakt, also beschränkt und abgeschlossen. Man kann die Konzepte *Messbarkeit* und *Volumen* aber auch für allgemeinere Mengen einführen (siehe die Lebesguesche Theorie).

Beispiele

1. Quader sind messbar und es gilt

$$\text{vol}([a_1, b_1] \times \dots \times [a_n, b_n]) = (b_1 - a_1) \cdot \dots \cdot (b_n - a_n).$$

Entartete Quader mit $a_j = b_j$ für mindestens ein $j = 1 \dots n$ sind Nullmengen.

2. Endliche Vereinigungen von Quadern sind messbar. Genauer gesagt: Gilt

$$D = \bigcup_{m=1}^M Q_m$$

für endlich viele Quader, die sich höchstens in Randpunkten überlappen, so gilt auch

$$\text{vol}(D) = \sum_{m=1}^M \text{vol}(Q_m).$$

Überlappen sich die Quader jedoch in inneren Punkten, so kann diese Formel nicht verwendet werden, aber man kann D dann immer als Vereinigung anderer Quader darstellen, für die die Formel wieder gilt (siehe Bild).

3. Die n -dimensionale, abgeschlossene Kugel

$$\overline{B}_r(\mathbf{x}_*) = \{\mathbf{x} \in \mathbb{R}^n : |\mathbf{x} - \mathbf{x}_*| \leq r\}$$

mit Radius r und Mittelpunkt $\mathbf{x}_* \in \mathbb{R}^n$ ist messbar, wobei die aus der Schule bekannten Formeln

$$\text{vol}(B_r(\mathbf{x}_*)) = \pi r^2 \quad (\text{für } n = 2) \quad \text{bzw.} \quad \text{vol}(B_r(\mathbf{x}_*)) = \frac{4}{3} \pi r^3 \quad (\text{für } n = 3)$$

gelten. In der mathematischen Theorie ist die offene Kugel

$$B_r(\mathbf{x}_*) = \{\mathbf{x} \in \mathbb{R}^n : |\mathbf{x} - \mathbf{x}_*| < r\}$$

übrigens auch messbar, aber die offene und die abgeschlossene Kugel haben immer dasselbe Volumen, weil die Differenzmenge (Sphäre vom Radius r um \mathbf{x}_* eine Nullmenge ist.

4. Alle in einfacher Weise aus endlich vielen elementaren geometrischen Objekten (Kugeln, Zylinder, Kegel, Quader) zusammengesetzte Teilmengen des \mathbb{R}^3 sind messbar. Analoge Aussagen gelten für alle $n \in \mathbb{N}$.

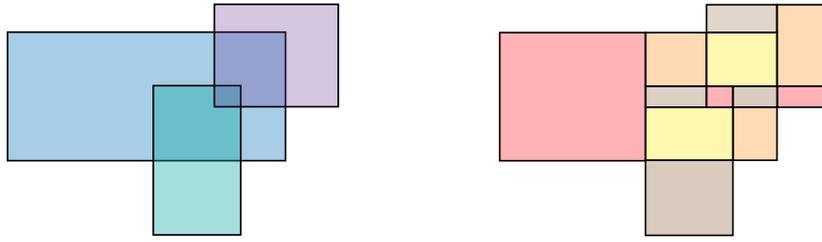


Abbildung Eine Vereinigung von sich im Inneren überlappenden Quadern kann auch als Vereinigung von Quadern dargestellt werden, die sich nur in Randpunkten überlappen.

Gegenbeispiel Die Dirichlet-Menge

$$D = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 \text{ und } x_2 \text{ sind rationale Zahlen aus dem Intervall } [0, 1]\}$$

ist eine seltsame Punktwolke, die nicht im (Riemannschen Sinne) messbar ist. Denn wenn wir die Ober- und Untersummen der entsprechenden charakteristischen Funktion auf einer (noch so feinen) Zerlegung Z des Einheitsquaders $[0, 1] \times [0, 1]$ berechnen, so erhalten wir stets $U(Z) = 0$ und $O(Z) = 1$.

Theorem (Approximation messbarer Mengen) Eine kompakte Menge D ist genau dann messbar, falls es für jedes $\varepsilon > 0$ zwei messbare Mengen $\underline{D}_\varepsilon$ und \overline{D}_ε mit den folgenden Eigenschaften gibt:

1. $\underline{D}_\varepsilon$ und \overline{D}_ε sind jeweils die Vereinigung endlich vieler Quader.
2. Es gilt $\underline{D}_\varepsilon \subset D \subset \overline{D}_\varepsilon$.
3. Es gilt $0 \leq \text{vol}(\overline{D}_\varepsilon) - \text{vol}(\underline{D}_\varepsilon) \leq \varepsilon$.

Insbesondere gilt dann immer

$$\text{vol}(\underline{D}_\varepsilon) \leq \text{vol}(D) \leq \text{vol}(\underline{D}_\varepsilon) + \varepsilon, \quad \text{vol}(\overline{D}_\varepsilon) - \varepsilon \leq \text{vol}(D) \leq \text{vol}(\overline{D}_\varepsilon),$$

d.h. die Volumina von $\underline{D}_\varepsilon$, D und \overline{D}_ε unterscheiden sich untereinander nur um Terme der Ordnung ε .

Beweisidee Wir könnten den Beweis mit unseren Mitteln im Prinzip führen, müssten aber viele und relativ unübersichtliche Notationen einführen. Wir appellieren lieber an die Intuition und wollen bemerken, dass es für jede hinreichend feine Zerlegung eines Quaders Q mit $D \subset Q$ nur drei Arten von Teilquadern $Q_{\mathbf{k}}$ geben wird:

$$\begin{aligned} \text{Typ 1: } \inf \chi_D|_{Q_{\mathbf{k}}} &= 0, \quad \sup \chi_D|_{Q_{\mathbf{k}}} = 0 && \text{(d.h. } Q_{\mathbf{k}} \subset Q \setminus D) \\ \text{Typ 2: } \inf \chi_D|_{Q_{\mathbf{k}}} &= 0, \quad \sup \chi_D|_{Q_{\mathbf{k}}} = 1 && \text{(d.h. } Q_{\mathbf{k}} \cap D \neq \emptyset \neq Q_{\mathbf{k}} \cap (\mathbb{R}^n \setminus D)) \\ \text{Typ 3: } \inf \chi_D|_{Q_{\mathbf{k}}} &= 1, \quad \sup \chi_D|_{Q_{\mathbf{k}}} = 1 && \text{(d.h. } Q_{\mathbf{k}} \subset D) \end{aligned}$$

Die Idee ist, dass die Vereinigung aller Quader vom Typ 3 die innere Approximation der Menge D liefert, wohingegen alle Quader vom Typ 2 und Typ 3 zusammen die äußere Approximation darstellen. Die Vereinigung aller Quader vom Typ 2 überdeckt dabei gerade den Rand von D . \square

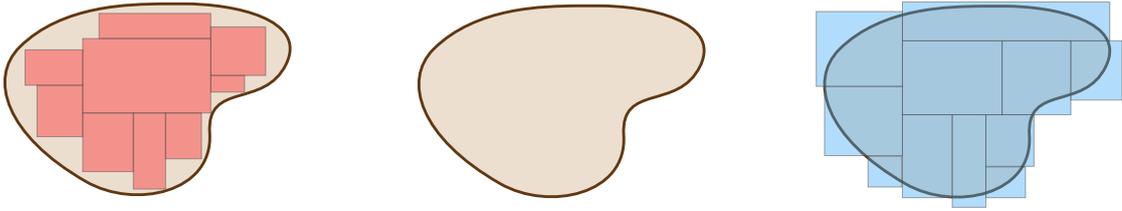


Abbildung Approximation einer Menge (braun) von innen (rot) und von außen (blau), wobei sich hier alle Quader höchstens in Randpunkten überlappen. Man könnte auch mit gleichgroßen Quadern approximieren (wie beim *Millimeterpapier* in der Schule). Die Menge ist genau dann messbar, wenn man den Unterschied zwischen den Volumina der inneren und äußeren Approximation beliebig klein machen kann.

Bemerkung Der Fall $\underline{D}_\varepsilon = \emptyset$ — die leere Menge als Vereinigung von null Quadern — ist zugelassen, siehe zum Beispiel das Kurvenbild weiter unten. Es gilt natürlich $\text{vol}(\emptyset) = 0$.

Approximation von π^* An Stelle von Quadern können wir im Prinzip auch andere geometrische Figuren verwenden. Die zweidimensionale Einheitskreisscheibe kann zum Beispiel von innen und von außen durch K kongruente, gleichschenklige Dreiecke approximiert werden (siehe Bild). Der Winkel am Schenkel eines Dreiecks ist dabei gerade $2\pi/K$ und die Scheitellänge nimmt den Wert 1 bzw. $1/\sqrt{1 - \sin^2(\pi/K)}$ an. Die entsprechenden Flächeninhalte können mit schulmathematischen Methoden zu

$$\text{vol}(\underline{D}_\varepsilon) = \frac{\sin(\pi\varepsilon) \sqrt{1 - \sin^2(\pi\varepsilon)}}{\varepsilon}, \quad \text{vol}(\overline{D}_\varepsilon) = \frac{\sin(\pi\varepsilon)}{\varepsilon \sqrt{1 - \sin^2(\pi\varepsilon)}}$$

berechnet werden, wobei wir $\varepsilon = 1/K$ substituiert haben. Mit Hilfe einer Taylor-Entwicklung bzgl. der kleinen Größe ε kann man nun die effektiven Formeln

$$\begin{aligned} \text{vol}(\underline{D}_\varepsilon) &= \pi - \frac{2}{3} \pi^3 \varepsilon^2 + \frac{2}{15} \pi^5 \varepsilon^4 + O(\varepsilon^6), \\ \text{vol}(\overline{D}_\varepsilon) &= \pi + \frac{1}{3} \pi^3 \varepsilon^2 + \frac{2}{15} \pi^5 \varepsilon^4 + O(\varepsilon^6) \end{aligned}$$

ableiten, wobei O das Landau-Symbol ist. Insbesondere gilt

$$0 \leq \text{vol}(\overline{D}_\varepsilon) - \text{vol}(\underline{D}_\varepsilon) \leq \varepsilon, \quad |\text{vol}(\overline{D}_\varepsilon) - \pi| \leq \varepsilon, \quad |\pi - \text{vol}(\underline{D}_\varepsilon)| \leq \varepsilon$$

für alle hinreichend großen K .

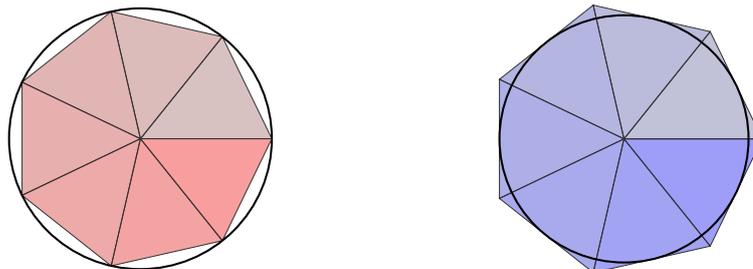


Abbildung Zur Approximation der Kreisscheibe mit K gleichschenkligen Dreiecken, hier dargestellt für $K = 7$. Für nicht allzu kleine Werte von K erhält man durch Berechnung der entsprechenden Flächen brauchbare Näherungsformeln für π , denen aber im Zeitalter des Computers keine praktische Bedeutung mehr zukommt.

Bemerkung Kurven in der Ebene oder im Raum sowie Flächen im Raum sind Nullmengen, sofern sie hinreichend regulär sind (siehe die Bilder).

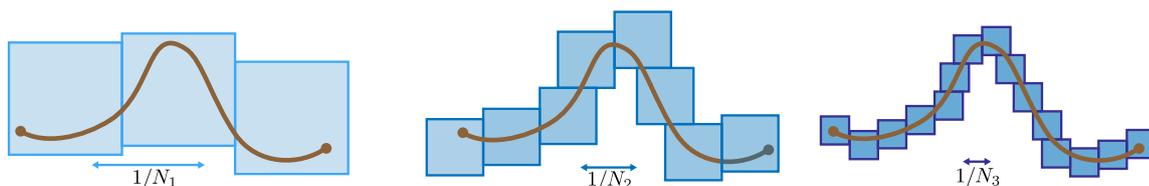


Abbildung Das Bild einer stetig differenzierbaren, parametrisierten Kurve in der Ebene mit Länge L ist eine zweidimensionale Nullmenge, weil wir die Bildmenge durch ungefähr $L \cdot N$ Quadrate mit Kantenlänge $1/N$ und Volumen $1/N^2$ überdecken, d.h. von außen approximieren, können. Hier dargestellt für eine Kurve (braun) und drei Werte $N_1 < N_2 < N_3$. Die entsprechende innere Approximation besteht aus null Quadern. Analog kann man zeigen, dass auch das Bild einer stetig differenzierbaren, parametrisierten Fläche eine dreidimensionale Nullmenge ist, wobei wir ungefähr $A \cdot N^2$ Würfel mit Kantenlänge $1/N$ und Volumen $1/N^3$ benötigen, sofern A die Oberfläche ist.

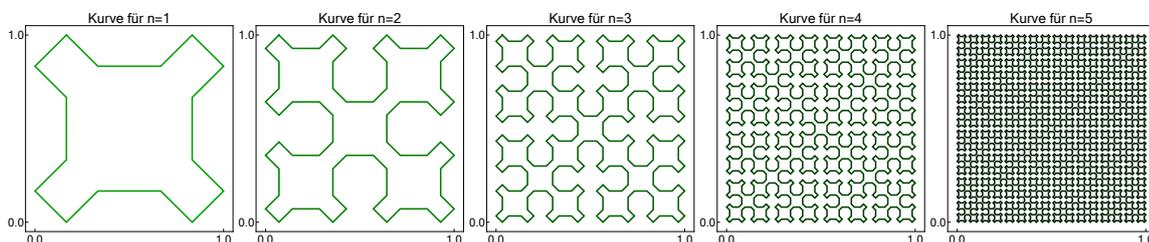


Abbildung Die planare *Sierpinski-Kurve* wird aus einer Folge von selbstähnlichen, stückweise affinen Kurven durch Grenzübergang $n \rightarrow \infty$ gewonnen und ist ein Beispiel für eine *flächenfüllende Kurve*. Das Limesobjekt ($n = \infty$) entspricht einer parametrisierten Kurve, die aber nicht mehr stetig differenzierbar ist, sodass das Überdeckungsargument aus dem letzten Bild nicht greift. Das Bild jeder approximierenden Kurve ($n < \infty$) ist aber eine endliche Vereinigung von Geradenstücken und damit eine Nullmenge. Ähnliche Beispiel sind die *Hilbert-Kurve* sowie die *Peano-Kurve*, siehe jeweils WIKIPEDIA.

Theorem (Zweite äquivalente Charakterisierung messbarer Mengen) Eine kompakte Menge D ist genau dann messbar, wenn die Menge ihrer Randpunkte eine Nullmenge ist.

Beweisidee Auch hier wollen wir den rigorosen Beweis nicht führen (siehe etwa [AORS, Satz 19.1.3]), aber die zugrundeliegenden Ideen hatten wir schon im Beweis des letzten Theorems kennengelernt: Betrachten wir wieder einen Quader Q mit $D \subseteq Q$ sowie eine entsprechende Zerlegung, so bilden die Teilquader von Typ 2 gerade eine Überdeckung der Menge der Randpunkte, wohingegen die Quader vom Typ 3 die Menge Q von innen approximieren. \square

Integrale stetiger Funktionen

Theorem (Hauptsatz über Integration auf messbaren Mengen) Ist D messbar, so ist jede stetige Funktion $f : D \rightarrow \mathbb{R}$ integrierbar. Insbesondere gilt das Minimum- und Maximumprinzip

$$\min f \cdot \text{vol}(D) \leq \int_D f(\mathbf{x}) \, d\mathbf{x} \leq \max f \cdot \text{vol}(D),$$

wobei wir hier ausgenutzt haben, dass stetige Funktionen auf kompakten Mengen ihr Minimum und Maximum annehmen.

Beweisidee Eine erste Beweisstrategie geht wie folgt. Wir approximieren D durch eine endliche Vereinigung von Quadern (die sich nur in Randpunkten überlappen) und approximieren anschließend das Integral von f auf jedem Teilquader durch Ober- oder Untersummen bzgl. einer hinreichend feinen Zerlegung des Teilquaders. Die handwerkliche Schwierigkeit liegt dabei in der Buchhaltung aller Teilquader sowie der lokalen Approximationsfehler. Eine elegantere — aber für den Laien undurchsichtigere — Beweisstrategie ist in [AORS, Satz 19.1.4] beschrieben. Die Doppelungleichung fällt in jedem Beweis als Nebenprodukt ab. \square

Bemerkungen

1. Es gilt $\text{vol}(D) = \int_D 1 \, d\mathbf{x}$ für jede messbare Menge D .
2. Ist D ein Nullmenge, so gilt $\int_D f(\mathbf{x}) \, d\mathbf{x} = 0$ für jede stetige Funktion f auf D .
3. Wie schon auf Quadern gilt: Nicht jede integrierbare Funktion ist stetig. Zum Beispiel sind stückweise stetige Funktionen auch integrierbar.
4. Man schreibt oftmals auch $\min f|_D$ bzw. $\max f|_D$ statt $\min f$ und $\max f$. Das ist vor allem dann sinnvoll, wenn f nicht nur auf D , sondern sogar auf einer größeren Menge definiert ist, aber nur über D integriert werden soll.

Zusammenhang und Mittelwertsatz der Integralrechnung* Eine kompakte Menge $D \subset \mathbb{R}^n$ heißt zusammenhängend, falls je zwei Punkte aus D durch eine Kurve in D miteinander verbunden werden können (oder, salopp gesprochen, falls D nicht in zwei oder mehr separate Teile zerfällt). Auf einer Menge D , die messbar mit $\text{vol}(D) > 0$ und außerdem zusammenhängend ist, existiert für jede stetige Funktion f (mindestens) ein Punkt $\mathbf{x}_* \in D$, so dass

$$f(\mathbf{x}_*) = \frac{1}{\text{vol}(D)} \int_D f(\mathbf{x}) \, d\mathbf{x},$$

wobei man die rechte Seite das Integralmittel von f nennt.

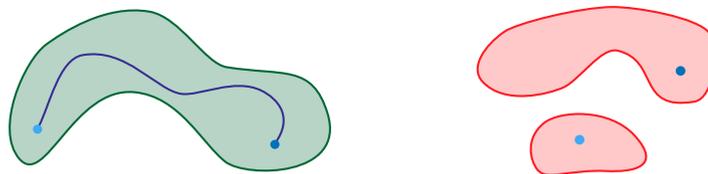


Abbildung Links: Die grüne Menge ist zusammenhängend und je zwei Punkte (blau) können immer durch eine Kurve verbunden werden, die ganz in dieser Menge verläuft. Rechts: Die rote Menge ist nicht zusammenhängend. Sie besteht aus zwei Teilen (den *Zusammenhangskomponenten*), die für sich betrachtet jeweils zusammenhängend sind.

Bemerkung Im Allgemeinen kann man den Punkt \mathbf{x}_* im Mittelwertsatz nur sehr schwer bestimmen, aber es ist trotzdem oftmals sehr nützlich zu wissen, dass es ihn gibt. Für seine Existenz sind der Zusammenhang von D sowie die Stetigkeit von f wichtig. Übungsaufgabe: Geben Sie ein Beispiel für D und f an, so dass es kein \mathbf{x}_* gibt.

Rechenregeln für Gebietsintegrale Unsere bisherigen Definitionen und Resultate implizieren — analog zu den entsprechenden Ergebnissen aus *Analysis 1* — die folgenden Aussagen:

1. (Linearität) Es gilt

$$\int_D \lambda \cdot f(\mathbf{x}) + \tilde{\lambda} \cdot \tilde{f}(\mathbf{x}) \, d\mathbf{x} = \lambda \cdot \int_D f(\mathbf{x}) \, d\mathbf{x} + \tilde{\lambda} \cdot \int_D \tilde{f}(\mathbf{x}) \, d\mathbf{x},$$

wobei λ und $\tilde{\lambda}$ reelle Zahlen sind.

2. (Gebietsadditivität) Für zwei messbare Mengen D und \tilde{D} gilt

$$\int_{D \cup \tilde{D}} f(\mathbf{x}) \, d\mathbf{x} = \int_D f(\mathbf{x}) \, d\mathbf{x} + \int_{\tilde{D}} f(\mathbf{x}) \, d\mathbf{x},$$

sofern sich beide Mengen nur in Randpunkten überlappen.

3. (Monotonie und Positivität) Es gelten die Implikationen

$$f(\mathbf{x}) \geq \tilde{f}(\mathbf{x}) \quad \text{für alle } \mathbf{x} \in D \quad \Longrightarrow \quad \int_D f(\mathbf{x}) \, d\mathbf{x} \geq \int_D \tilde{f}(\mathbf{x}) \, d\mathbf{x}$$

sowie

$$f(\mathbf{x}) \geq 0 \quad \text{für alle } \mathbf{x} \in D \quad \Longrightarrow \quad \int_D f(\mathbf{x}) \, d\mathbf{x} \geq 0.$$

Die zweite Aussage ist dabei eigentlich ein Spezialfall der ersten (via $\tilde{f}(\mathbf{x}) = 0$ für alle $\mathbf{x} \in D$).

4. („Dreiecksungleichung“ für Integrale) Es gilt

$$\left| \int_D f(\mathbf{x}) \, d\mathbf{x} \right| \leq \int_D |f(\mathbf{x})| \, d\mathbf{x}.$$

sowie

$$\left| \int_D (f(\mathbf{x}) - \tilde{f}(\mathbf{x})) \, d\mathbf{x} \right| \leq \int_D |f(\mathbf{x}) - \tilde{f}(\mathbf{x})| \, d\mathbf{x}.$$

Achtung: Es gibt hier kein Dreieck. Der Name dieser Ungleichungen hat sich aus abstrakten Analogiegründen ergeben.

Weitere Ungleichungen für Integrale* Sehr wichtig ist die Hölder-Ungleichung

$$\left| \int_D f(\mathbf{x}) \tilde{f}(\mathbf{x}) \, d\mathbf{x} \right| \leq \left(\int_D |f(\mathbf{x})|^p \, d\mathbf{x} \right)^{1/p} \left(\int_D |\tilde{f}(\mathbf{x})|^{p'} \, d\mathbf{x} \right)^{1/p'},$$

wobei immer stillschweigend vorausgesetzt wird, dass die Parameter p und q durch

$$\frac{1}{p} + \frac{1}{p'} = 1$$

miteinander gekoppelt sind. Im Spezialfall $p = p' = 2$ spricht man auch von der Cauchy-Schwartz-Ungleichung für Integrale. Die Jensensche Ungleichung garantiert

$$\psi \left(\frac{\int_D f(\mathbf{x}) w(\mathbf{x}) \, d\mathbf{x}}{\int_D w(\mathbf{x}) \, d\mathbf{x}} \right) \leq \frac{\int_D \psi(f(\mathbf{x})) w(\mathbf{x}) \, d\mathbf{x}}{\int_D w(\mathbf{x}) \, d\mathbf{x}}$$

für alle integrierbaren Funktionen $f : D \rightarrow \mathbb{R}$, wobei $\psi : \mathbb{R} \rightarrow \mathbb{R}$ eine beliebige *konvexe* Funktion bezeichnet und $w : D \rightarrow \mathbb{R}_+$ eine stetige und nicht-negative Funktion ist (die meist *Gewichtsfunktion* genannt wird). Mit $w(\mathbf{x}) = 1$ für alle \mathbf{x} ergibt sich der Spezialfall

$$\psi \left(\frac{1}{\text{vol}(D)} \int_D f(\mathbf{x}) \, d\mathbf{x} \right) \leq \frac{1}{\text{vol}(D)} \int_D \psi(f(\mathbf{x})) \, d\mathbf{x}$$

als entsprechende Ungleichung für Integralmittelwerte.

Funktionen mit überall verschwindendem Integral Ist $f : D \rightarrow \mathbb{R}$ eine stetige Funktion mit

$$\int_{\tilde{D}} f(\mathbf{x}) \, d\mathbf{x} = 0$$

für alle messbaren Mengen $\tilde{D} \subset D$, so gilt

$$f(\mathbf{x}) = 0 \quad \text{für alle inneren Punkte } \mathbf{x} \in D.$$

Dieses Ergebnis kann relativ leicht durch einen Widerspruchsbeweis abgeleitet werden (Übungsaufgabe; man wähle \tilde{D} geeignet als Quader oder Kugel). Die analoge Aussage für integrierbare Funktionen (die mangels Stetigkeit deutlich schwieriger zu zeigen ist) wird oftmals *Hauptsatz der Variationsrechnung* genannt.

Integrale vektorwertiger Funktionen Funktionen $\mathbf{f} : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ werden komponentenweise integriert, d.h. es gilt

$$\int_D \begin{pmatrix} f_1(\mathbf{x}) \\ \vdots \\ f_m(\mathbf{x}) \end{pmatrix} d\mathbf{x} = \begin{pmatrix} \int_D f_1(\mathbf{x}) \, d\mathbf{x} \\ \vdots \\ \int_D f_m(\mathbf{x}) \, d\mathbf{x} \end{pmatrix}$$

sofern alle skalaren Integrale auf der rechten Seite existieren. Dies gilt insbesondere, wenn D messbar und wenn \mathbf{f} stetig ist.

Schwerpunkte messbarer Mengen Der n -dimensionale Vektor

$$\text{bar}(D) := \frac{1}{\text{vol}(D)} \int_D \mathbf{x} \, d\mathbf{x}$$

wird Schwerpunkt (oder Baryzentrum) der messbaren Menge $D \subset \mathbb{R}^n$ genannt. Er beschreibt den physikalischen Schwerpunkt, sofern wir annehmen, dass die Menge D mit einem homogenen Material der (normierten) Massendichte 1 gefüllt ist. Insgesamt müssen wir also $1 + n$ skalare Integrale über der Menge D berechnen, nämlich

$$\int_D 1 \, d\mathbf{x}, \quad \int_D x_1 \, d\mathbf{x}, \quad \dots, \quad \int_D x_n \, d\mathbf{x}.$$

Ein zweidimensionales Beispiel werden wir dazu weiter unten studieren. Beachte, dass der Schwerpunkt einer nicht-konvexen Menge außerhalb liegen kann und dass bei inhomogenen Materialien die (nicht-konstante) Massendichte in den Integralen auftaucht.

Parameterabhängige Integrale Hängt der Integrand f nicht nur von $\mathbf{x} \in \mathbb{R}^n$, sondern auch stetig differenzierbar von Parametern $\mathbf{p} \in \mathbb{R}^k$ ab, so gilt

$$\partial_{p_i} \int_D f(\mathbf{x}, \mathbf{p}) \, d\mathbf{x} = \int_D \partial_{p_i} f(\mathbf{x}, \mathbf{p}) \, d\mathbf{x}$$

für alle $i = 1 \dots k$, sofern das Integrationsgebiet D unabhängig von \mathbf{p} ist. Sollte D jedoch auch von den Parametern abhängen, so muss die rechte Seite in dieser Formel um weitere Terme ergänzt werden.

Beispiel Auf jeder messbaren Menge $D \subset \mathbb{R}^2$ gilt zum Beispiel

$$\partial_{p_1} \iint_D (p_1 x_1 + \sin(p_2 x_2)) \, dx_1 \, dx_2 = \iint_D x_1 \, dx_1 \, dx_2,$$

sowie

$$\partial_{p_2} \iint_D (p_1 x_1 + \sin(p_2 x_2)) \, dx_1 \, dx_2 = \iint_D x_2 \cos(p_2 x_2) \, dx_1 \, dx_2,$$

wobei wir für jedes konkrete D die Formeln natürlich auch durch explizite Berechnung aller Integrale begründen können.

2.6 Integration über spezielle Gebietsklassen

Ziel Wir wollen in diesem Abschnitt für gewisse Gebietsklassen in zwei und drei Dimensionen effektive Integrationsformeln angeben.

Zweidimensionale Normalbereiche

Definition Eine Menge der Bauart

$$D = \left\{ (x_1, x_2) : a \leq x_1 \leq b, \quad g(x_1) \leq x_2 \leq h(x_1) \right\}$$

bzw.

$$D = \left\{ (x_1, x_2) : a \leq x_2 \leq b, \quad g(x_2) \leq x_1 \leq h(x_2) \right\}$$

wird x_1 -Normalbereich bzw. x_2 -Normalbereich genannt, wobei a, b zwei Zahlen und g, h zwei stetige skalare Funktionen in einer Variablen sind. Statt *Bereich* kann man immer auch *Gebiet* sagen.

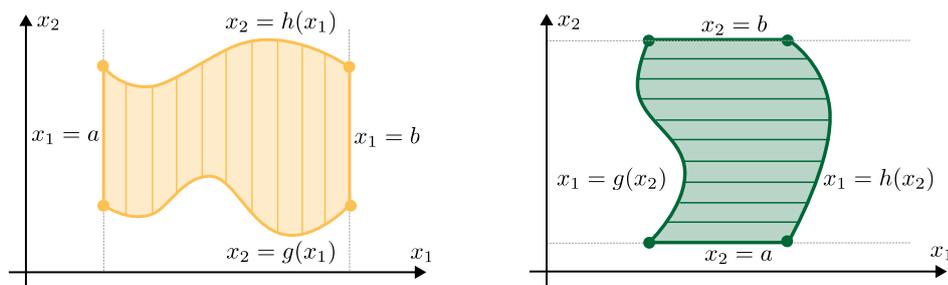


Abbildung Ein x_1 -Normalbereich (gelb) sowie ein x_2 -Normalbereich (grün), wobei jeder genau vier verschiedene Eckpunkte besitzt.

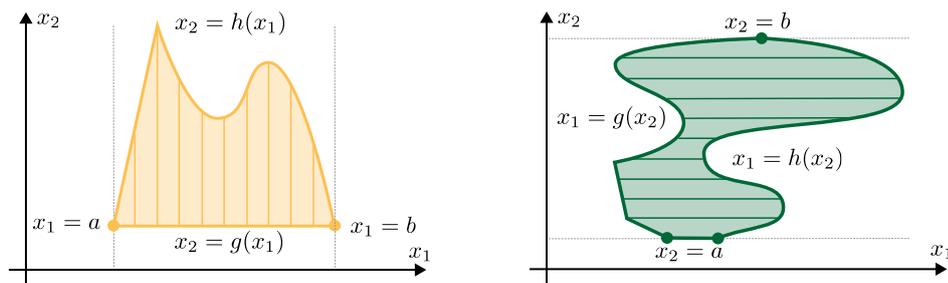


Abbildung Weitere Beispiele für Normalbereiche, wobei diesmal einige Eckpunkte zusammenfallen und die Funktionen g und h auch nur stückweise differenzierbar sind. Salopp gesprochen gilt: Einen x_1 - bzw. x_2 -Normalbereich können wir vertikal bzw. horizontal so schraffieren, dass die Schraffurlinien niemals unterbrochen werden. Rechtecke und Kreisscheiben gehören damit zu beiden Klassen.

Theorem (Fubini für Normalbereiche in 2D) Jeder x_1 -Normalbereich D ist messbar und für jede stetige Funktion $f : D \rightarrow \mathbb{R}$ gilt

$$\int_D f(\mathbf{x}) \, d\mathbf{x} = \int_a^b \left(\int_{g(x_1)}^{h(x_1)} f(x_1, x_2) \, dx_2 \right) dx_1.$$

Eine analoge Formel mit konsequent vertauschten Indizes gilt für x_2 -Normalbereiche.

Beweisidee Man kann dieses Resultat analog zum Satz von Fubini beweisen, muss aber zusätzlich auch noch das Gebiet mit Quadern approximieren. Alternativ kann man die Formel aus dem Transformationssatz für Integrale ableiten, den wir weiter unten kennenlernen werden. Siehe auch die Übungsaufgaben. \square

Bemerkung Die Reihenfolge der Integralschachtelung darf diesmal — im Gegensatz zu Quadern — im Allgemeinen nicht vertauscht werden, da bei x_1 -Normalbereichen die Integrationsgrenzen der inneren Integrationsvariablen x_2 von der äußeren Integrationsvariablen x_1 abhängen. Oder anders gesagt: Bei Quadern darf man die Verschachtelung vertauschen, weil Quader sowohl x_1 -Normalbereiche als auch x_2 -Normalbereiche sind.

Beispiele

1. Für

$$a = -1 \quad b = +1, \quad g(x_1) = -x_1^2 - 1, \quad h(x_1) = +x_1^2 + 1$$

wird D durch zwei Geradenstücke sowie zwei Parabelstücke begrenzt. Mit

$$f(x_1, x_2) = 1 + x_1 + x_2 + x_1 x_2$$

erhalten wir

$$\begin{aligned} \int_D f(\mathbf{x}) \, d\mathbf{x} &= \int_{-1}^{+1} \left(\int_{-x_1^2-1}^{x_1^2+1} (1 + x_1 + x_2 + x_1 x_2) \, dx_2 \right) dx_1 \\ &= \int_{-1}^{+1} \left[x_2 + x_1 x_2 + \frac{1}{2} x_2^2 + \frac{1}{2} x_1 x_2^2 \right]_{x_2=-x_1^2-1}^{x_2=+x_1^2+1} dx_1 \\ &= \int_{-1}^{+1} \left((2x_1^2 + 2) + x_1(2x_1^2 + 2) + 0 + 0 \right) dx_1 \\ &= \int_{-1}^{+1} \left(2x_1^3 + 2x_1^2 + 2x_1 + 2 \right) dx_1 \\ &= 2 \left[\frac{1}{4} x_1^4 + \frac{1}{3} x_1^3 + \frac{1}{2} x_1^2 + x_1 \right]_{x_1=-1}^{x_1=+1} \\ &= 0 + \frac{4}{3} + 0 + 4 = \frac{16}{3}, \end{aligned}$$

wobei wir bei der Berechnung des inneren und des äußeren Integrals jeweils den eindimensionalen Hauptsatz der Differential- und Integralrechnung angewendet haben.

2. Um den Flächeninhalt der Einheitskreisscheibe zu berechnen, setzen wir

$$a = -1, \quad b = +1, \quad g(x_1) = -\sqrt{1 - x_1^2}, \quad h(x_1) = +\sqrt{1 - x_1^2}$$

und berechnen

$$\begin{aligned}
 \text{vol}(D) &= \iint_D 1 \, dx_1 \, dx_2 = \int_{-1}^{+1} \left(\int_{-\sqrt{1-x_1^2}}^{+\sqrt{1+x_1^2}} 1 \, dx_2 \right) dx_1 \\
 &= \int_{-1}^{+1} 2\sqrt{1-x_1^2} \, dx_1 \\
 &= \left[x_1 \sqrt{1-x_1^2} + \arcsin(x_1) \right]_{x_1=-1}^{x_1=+1} \\
 &= \arcsin(+1) - \arcsin(-1) = \pi,
 \end{aligned}$$

wobei $\arcsin : [-1, +1] \rightarrow [-\pi/2, +\pi/2]$ die lokale Umkehrfunktion des Sinus ist und wir die Stammfunktion beim letzten Integral aus einer Tabelle abgelesen bzw. via

$$\begin{aligned}
 \frac{d}{dx_1} \left(x_1 \sqrt{1-x_1^2} + \arcsin(x_1) \right) &= \sqrt{1-x_1^2} - \frac{1}{2} \frac{x_1(-2x_1)}{\sqrt{1-x_1^2}} + \frac{1}{\sqrt{1-x_1^2}} \\
 &= \frac{1-x_1^2 + 2x_1^2 + 1}{\sqrt{1-x_1^2}} = 2\sqrt{1-x_1^2}
 \end{aligned}$$

verifiziert haben. In diesem Beispiel hätten wir aufgrund der Symmetrie der Kreisscheibe die Indizes 1 und 2 vertauschen können.

3. Der x_1 -Normalbereich

$$D = \{(x_1, x_2) : -2 \leq x_1 \leq 1, \quad x_1 + 2 \leq x_2 \leq -x_1^2 + 4\}$$

ist die ebene Fläche, die von oben durch die Parabel $x_2 = -x_1^2 + 4$ und von unten durch die Gerade $x_2 = x_1 + 2$ begrenzt wird, wobei $(-2, 0)$ und $(1, 3)$ die beiden Schnittpunkte dieser Kurven sind. Wir wollen nun den Schwerpunkt der Menge (siehe oben) ermitteln und berechnen dazu die Integrale

$$\begin{aligned}
 \int_D 1 \, d\mathbf{x} &= \int_{-2}^{+1} \left(\int_{x_1+2}^{-x_1^2+4} 1 \, dx_2 \right) dx_1 = \int_{-2}^{+1} (-x_1^2 - x_1 + 2) \, dx_1 \\
 &= \left[-\frac{1}{3}x_1^3 - \frac{1}{2}x_1^2 + 2x_1 \right]_{x_1=-2}^{x_1=1} = \left(-\frac{1}{3} - \frac{1}{2} + 2\right) - \left(+\frac{8}{3} - 2 - 4\right) = \frac{9}{2}
 \end{aligned}$$

sowie

$$\begin{aligned}
 \int_D x_1 \, d\mathbf{x} &= \int_{-2}^{+1} \left(\int_{x_1+2}^{-x_1^2+4} x_1 \, dx_2 \right) dx_1 = \int_{-2}^{+1} x_1 (-x_1^2 - x_1 + 2) \, dx_1 \\
 &= \left[-\frac{1}{4}x_1^4 - \frac{1}{3}x_1^3 + x_1^2 \right]_{x_1=-2}^{x_1=1} = \left(-\frac{1}{4} - \frac{1}{3} + 1\right) - \left(-4 + \frac{8}{3} + 4\right) = -\frac{9}{4}
 \end{aligned}$$

und

$$\begin{aligned} \int_D x_2 \, d\mathbf{x} &= \int_{-2}^{+1} \left(\int_{x_1+2}^{-x_1^2+4} x_2 \, dx_2 \right) dx_1 = \int_{-2}^{+1} \left[\frac{1}{2} x_2^2 \right]_{x_2=x_1+2}^{-x_1^2+4} dx_1 \\ &= \int_{-2}^{+1} \left(\frac{1}{2} x_1^4 - \frac{9}{2} x_1^2 - 2x_1 + 6 \right) dx_1 \\ &= \left[\frac{1}{10} x_1^5 - \frac{3}{2} x_1^3 - x_1^2 + 6x_1 \right]_{x_1=-2}^{x_1=1} \\ &= \left(\frac{1}{10} - \frac{3}{2} - 1 + 6 \right) - \left(-\frac{32}{10} + 12 - 4 - 12 \right) = \frac{54}{5}. \end{aligned}$$

Insgesamt erhalten

$$\text{bar}(D) := \frac{2}{9} \begin{pmatrix} -\frac{9}{4} \\ +\frac{54}{5} \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} \\ +\frac{12}{5} \end{pmatrix}$$

für den gesuchten Schwerpunkt.

Zusammengesetzte Gebiete Nicht jede messbare Menge D ist ein Normalbereich. Oftmals ist es aber möglich, D via

$$D = D_1 \cup \dots \cup D_K$$

in K Normalbereiche zu zerlegen, wobei sich die Teilgebiete nur in Randpunkten überlappen. In diesem Fall gilt

$$\int_D f(\mathbf{x}) \, d\mathbf{x} = \int_{D_1} f(\mathbf{x}) \, d\mathbf{x} + \dots + \int_{D_K} f(\mathbf{x}) \, d\mathbf{x}$$

aufgrund der Gebietsadditivität von Integralen. Die Zerlegung ist aber nicht eindeutig und die Kunst besteht darin, eine Zerlegung zu wählen, für die die Teilintegrale möglichst einfach zu berechnen sind.

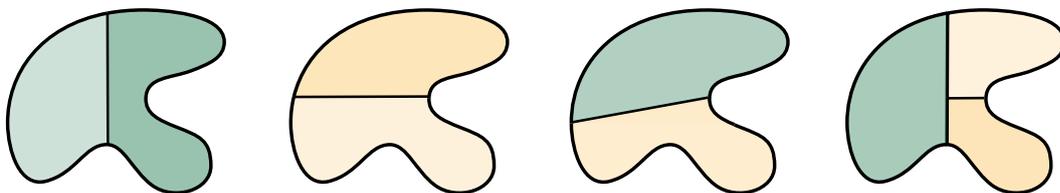


Abbildung Vier Beispiele für die Zerlegung einer Menge D in Normalbereiche, die sich gegenseitig nur in Randpunkten überlappen.

Dreidimensionale Normalbereiche und projizierbare Mengen

Definition Eine Menge $D \subset \mathbb{R}^3$ der Bauart

$$D = \left\{ (x_1, x_2, x_3) : a \leq x_1 \leq b, \quad g(x_1) \leq x_2 \leq h(x_1), \quad u(x_1, x_2) \leq x_3 \leq v(x_1, x_2) \right\}$$

nennen wir x_1 - x_2 -Normalgebiet, wobei a, b zwei Zahlen, g, h zwei stetige skalare Funktionen in einer Variablen und u, v zwei stetige skalare Funktionen in zwei Variablen sind. Analoge Definitionen gelten für jede Permutation der drei Indizes.

Theorem (Fubini für Normalbereiche in 3D) Jeder x_1 - x_2 -Normalbereich D ist messbar und

$$\int_D f(\mathbf{x}) \, d\mathbf{x} = \int_a^b \left(\int_{g(x_1)}^{h(x_1)} \left(\int_{u(x_1, x_2)}^{v(x_1, x_2)} f(x_1, x_2, x_3) \, dx_3 \right) dx_2 \right) dx_1$$

gilt für alle stetigen Funktionen $f : D \rightarrow \mathbb{R}$. Analoge Formeln ergeben sich aus der konsequenten Permutation der Indizes.

Beweisidee Siehe die Kommentare zum entsprechenden 2D-Theorem. \square

Beispiele

- Die dreidimensionale Einheitskugel $D = \{(x_1, x_2, x_3) : x_1^2 + x_2^2 + x_3^2 = 1\}$ ist ein x_i - x_j -Normalgebiet mit

$$a = -1, \quad b = +1, \quad g(x_i) = -\sqrt{1 - x_i^2}, \quad h(x_i) = +\sqrt{1 - x_i^2}$$

und

$$u(x_i, x_j) = -\sqrt{1 - x_i^2 - x_j^2}, \quad v(x_i, x_j) = +\sqrt{1 - x_i^2 - x_j^2},$$

wobei (i, j, k) eine beliebige Permutation von $(1, 2, 3)$ darstellt.

- Quader sind auch immer x_i - x_j -Normalbereiche, wobei g, h sowie u, v jeweils konstante Funktionen sind.

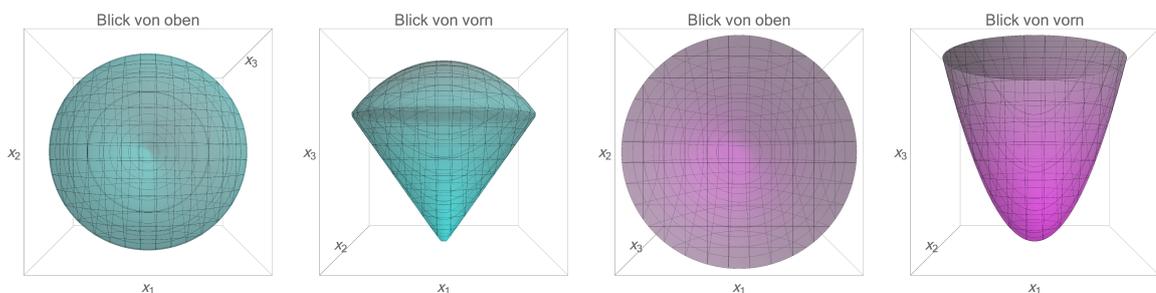


Abbildung Der Kreisel (Türkis) und die parabolische Vase (Lila) sind zwei Beispiele für rotations-symmetrische x_1 - x_2 -Normalbereiche. Sie sind auch x_3 projizierbar, wobei B gerade eine Kreisscheibe ist.

Beispiel Für

$$a = -2, \quad b = +2, \quad g(x_1) = -\sqrt{4 - x_1^2}, \quad h(x_1) = +\sqrt{4 - x_1^2}$$

und

$$u(x_1, x_2) = x_1^2 + x_2^2, \quad v(x_1, x_2) = 4$$

ist der entsprechende x_1 - x_2 -Normalbereich D eine vasenförmige Menge mit Höhe 4 und entarteter Grundfläche, deren Mantelfläche Teil eines Rotationsparaboloiden ist. Das

Fassungsvermögen dieser Vase ergibt sich zu

$$\begin{aligned}
 \text{vol}(D) &= \int_{-2}^{+2} \left(\int_{-\sqrt{4-x_1^2}}^{+\sqrt{4-x_1^2}} \left(\int_{x_1^2+x_2^2}^4 1 \, dx_3 \right) dx_2 \right) dx_1 \\
 &= \int_{-2}^{+2} \left(\int_{-\sqrt{4-x_1^2}}^{+\sqrt{4-x_1^2}} (4 - x_1^2 - x_2^2) dx_2 \right) dx_1 \\
 &= \int_{-2}^{+2} \left[4x_2 - x_1^2 x_2 - \frac{1}{3} x_2^3 \right]_{x_2=-\sqrt{4-x_1^2}}^{x_2=+\sqrt{4-x_1^2}} dx_1 \\
 &= \int_{-2}^{+2} \left(8\sqrt{4-x_1^2} - 2x_1^2 \sqrt{4-x_1^2} - \frac{2}{3}(4-x_1^2) \sqrt{4-x_1^2} \right) dx_1 \\
 &= \frac{4}{3} \int_{-2}^{+2} (4-x_1^2)^{3/2} dx_1 = 8\pi,
 \end{aligned}$$

wobei wir gleich auch den Fall allgemeiner Rotationskörper untersuchen werden. Insbesondere wird sich zeigen, dass wir die Vase aus diesem Beispiel besser als einen x_3 - x_2 -Normalbereich betrachtet hätten.

Bemerkung: Um das letzte Integral auszuwerten, benutzen wir zunächst partielle Integration und berechnen

$$\begin{aligned}
 \int_{-2}^{+2} (4-x_1^2)^{3/2} dx_1 &= \left[x(4-x_1^2)^{3/2} \right]_{x_1=-2}^{x_1=+2} - \int_{-2}^{+2} \frac{3}{2} (4-x_1^2)^{1/2} (-2x_1) dx_1 \\
 &= 3 \int_{-2}^{+2} x_1^2 (4-x_1^2)^{1/2} dx_1.
 \end{aligned}$$

Mit der „nahrhaften Null“ $x_1^2 = -(4-x_1^2) + 4$ und nach Umstellung der Terme erhalten wir

$$\begin{aligned}
 \int_{-2}^{+2} (4-x_1^2)^{3/2} dx_1 &= \int_{-2}^{+2} 3(4-x_1^2)^{1/2} dx_1 \\
 &= \left[\frac{3}{2} x_1 \sqrt{4-x_1^2} + 6 \arcsin\left(\frac{1}{2} x_1\right) \right]_{x_1=-2}^{x_1=+2} = 6\pi,
 \end{aligned}$$

wobei wir eine ähnliche Stammfunktion schon beim Volumen der Kreisscheibe benutzt hatten.

Volumen von Rotationskörpern Die Menge

$$D := \left\{ (x_1, x_2, x_3) : x_1^2 + x_2^2 \leq \varrho^2(x_3), \quad a \leq x_3 \leq b \right\}$$

ist gerade der Rotationskörper, der durch Rotation des Graphen der nichtnegativen Funktion $\varrho : [a, b] \rightarrow [0, \infty)$ um die x_3 -Achse entsteht. Es handelt sich um einen x_3 - x_2 -Normalbereich, sofern wir

$$g(x_3) = -\varrho(x_3), \quad h(x_3) = +\varrho(x_3)$$

und

$$u(x_2, x_3) = -\sqrt{\varrho^2(x_3) - x_2^2}, \quad v(x_2, x_3) = +\sqrt{\varrho^2(x_3) - x_2^2}$$

wählen. Wir erhalten damit

$$\begin{aligned} \text{vol}(G) &= \int_a^b \left(\int_{-\varrho(x_3)}^{+\varrho(x_3)} \left(\int_{-\sqrt{\varrho^2(x_3)-x_2^2}}^{+\sqrt{\varrho^2(x_3)-x_2^2}} 1 \, dx_1 \right) dx_2 \right) dx_3 \\ &= \int_a^b \left(\int_{-\varrho(x_3)}^{+\varrho(x_3)} 2\sqrt{\varrho^2(x_3) - x_2^2} \, dx_2 \right) dx_3 \\ &= \int_a^b \pi \varrho^2(x_3) \, dx_3 \end{aligned}$$

wobei wir wieder die allgemeine Formel $\int_{-r}^{+r} \sqrt{r^2 - s^2} \, ds = \frac{1}{2} \pi r^2$ ausgewertet haben (diesmal mit $s = x_2$ und $r = \varrho(x_3)$). Wir hatten diese Formel schon in *Analysis 1* aus dem Prinzip des Cavalieri abgeleitet (wobei die damalige Argumentation eleganter als unsere heutige Rechnung ist). Beachte auch, dass die Rechnungen im letzten Beispiel komplizierter waren, eben weil wir D als x_1 - x_2 Normalbereich interpretiert hatten.

Definition* Eine Menge $D \subset \mathbb{R}^3$ der Bauart

$$D = \left\{ (x_1, x_2, x_3) : (x_1, x_2) \in B, \quad u(x_1, x_2) \leq x_3 \leq v(x_1, x_2) \right\}$$

heißt x_3 -projizierbar, wobei $B \subset \mathbb{R}^2$ eine zweidimensionale messbare Menge und u, v zwei stetige skalare Funktionen in jeweils zwei Variablen sind. Analog können x_1 -projizierbare und x_2 -projizierbare Mengen eingeführt werden.

Theorem* (**Fubini für projizierbare Mengen**) Mit den obigen Notationen gilt

$$\int_D f(\mathbf{x}) \, d\mathbf{x} = \iint_B \left(\int_{u(x_1, x_2)}^{v(x_1, x_2)} f(x_1, x_2, x_3) \, dx_3 \right) dx_1 \, dx_2$$

für jede stetig Funktion $f : D \rightarrow \mathbb{R}$.

Beweisidee Auch dieses Resultat kann entweder mit Approximationsargumenten oder alternativ aus dem Transformationssatz für Integrale abgeleitet werden. \square

Bemerkung

1. Jeder x_1 - x_2 -Normalbereich ist auch x_3 -projizierbar, sofern

$$B = \{(x_1, x_2) : a \leq x_1 \leq b, g(x_1) \leq x_2 \leq h(x_1)\}$$

gewählt wird. Etwas allgemeiner kann man sagen: Eine x_3 -projizierbare Menge ist genau dann x_1 - x_2 - bzw. x_2 - x_1 -Normalbereich, wenn B eine zweidimensionale x_1 - bzw. x_2 -Normalbereich ist.

2. Der Satz von Fubini ist bei projizierbaren Mengen eher von theoretischem Interesse, denn aus praktischer Sicht kann die Berechnung des Integrals über die zweidimensionale Menge B aufwändig sein. Ist B jedoch selbst Normalbereich, so ist auch D Normalbereich und wir können gleich den Satz von Fubini für dreidimensionale Normalbereiche benutzen.

2.7 Die Transformationsformel für Integrale

Motivation Integrale über komplizierte Gebiete kann man auch dadurch berechnen, dass man eine Transformation anwendet, die das komplizierte Gebiet in ein einfaches überführt. Man muss dann allerdings genaue Regeln der Umrechnung beachten, die man als *Transformationsformel* bezeichnet. Diese spielt auch aus theoretischer Sicht eine wichtige Rolle und wir werden später eine Variante beim Studium von gekrümmten Flächen benutzen.

Vorbereitungen

Setting Wir betrachten in diesem Abschnitt eine Abbildung $\Psi : U \subset \mathbb{R}^n \rightarrow V \subset \mathbb{R}^n$ mit den folgenden Eigenschaften:

1. Ψ bildet U bijektiv auf V ab.
2. Ψ ist stetig differenzierbar

Wir schreiben $\mathbf{x} \in U$ und $\mathbf{y} \in V$ und interpretieren Ψ als den Koordinatenwechsel von den x_j 's zu den y_i 's.

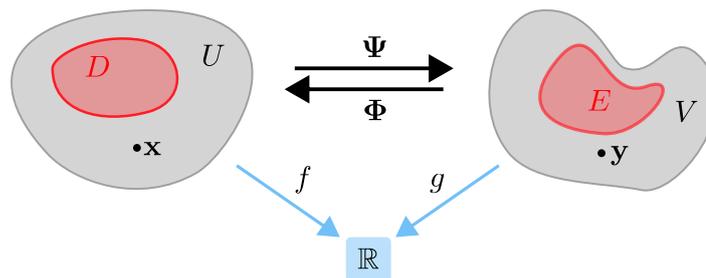


Abbildung Schematische Darstellung der Mengen und Abbildungen in diesem Abschnitt.

Bemerkung

1. Nach Voraussetzung existiert die Umkehrabbildung $\Phi = \Psi^{-1} : V \rightarrow U$ und die Kettenregel impliziert, dass diese auch stetig differenzierbar ist. Insbesondere gilt

$$\mathbf{J}\Phi(\mathbf{y}) \cdot \mathbf{J}\Psi(\mathbf{x}) = \mathbf{I} \quad \mathbf{J}\Psi(\mathbf{x}) \cdot \mathbf{J}\Phi(\mathbf{y}) = \mathbf{I}$$

im Sinne der Matrizen-Multiplikation mit Einheitsmatrix $\mathbf{I} \in \mathbb{R}^{(n,n)}$, sofern \mathbf{x} und \mathbf{y} durch $\mathbf{y} = \Psi(\mathbf{x})$ bzw. $\mathbf{x} = \Phi(\mathbf{y})$ gekoppelt sind.

2. Eine Menge $D \subset U$ wird unter Ψ auf die Bildmenge $E := \Psi(D) \subset V$ abgebildet, wobei dann auch $D = \Phi(E)$ gilt. Außerdem entspricht jede skalare Funktion $f : D \rightarrow \mathbb{R}$ via $g = f \circ \Phi$ bzw. $f = g \circ \Psi$ einer Funktion $g : E \rightarrow \mathbb{R}$ und umgekehrt. Mit anderen Worten: Es gilt

$$\mathbf{x} \in D \Leftrightarrow \mathbf{y} \in E, \quad f(\mathbf{x}) = g(\mathbf{y}),$$

wobei auch bei diesen Formeln immer $\mathbf{y} = \Psi(\mathbf{x})$ bzw. $\mathbf{x} = \Phi(\mathbf{y})$ mitgedacht werden muss.

3. Man kann zeigen: Ist D eine kompakte bzw. messbare Menge, so ist auch E kompakt bzw. messbar und umgekehrt. Außerdem ist $f : D \rightarrow \mathbb{R}$ genau dann stetig, wenn $g : E \rightarrow \mathbb{R}$ stetig ist.
4. Wie immer gilt auch hier: Die in der Literatur verwendeten Notationen sind nicht einheitlich und Sie müssen sich in jedem Kontext klar machen, welches Symbol gerade was bedeutet bzw. welche Relationen zwischen den beteiligten Abbildungen und Mengen bestehen. Schematische Diagramme wie das oben angegebene sind dabei ausgesprochen hilfreich.
5. Die Mengen U und V werden in der Mathematik meist als offen vorausgesetzt, da dann Randpunkte automatisch keine Rolle spielen und die Beweise einfacher werden. In den Anwendungswissenschaften sind die Mengen aber oftmals nicht offen.

Erinnerung: Eine Menge heißt offen, wenn sie keinen ihrer Randpunkte enthält.

6. Es ist sehr wichtig, dass Φ und Ψ stetig differenzierbar und zueinander invers sind. Man kann diese Voraussetzungen abschwächen — vor allem in Randpunkten von U und V — und wir werden solche Entartungspunkte bei konkreten Transformationen (zum Beispiel den Polar- und Kugelkoordinaten) diskutieren.

Transformation von Quadern Eine affine Abbildung

$$\Psi_{\text{aff}}(\mathbf{x}) = \mathbf{c} + \mathbf{A} \cdot (\mathbf{x} - \mathbf{d})$$

mit $\mathbf{A} \in \mathbb{R}^{(n,n)}$ und $\mathbf{c}, \mathbf{d} \in \mathbb{R}^n$ transformiert einen Quader $Q \subset U$ in das n -dimensionale Analogon zu einem Parallelogramm (also in einen Parallelepipid bzw. Spat für $n = 3$). Die Gesetze der linearen Algebra implizieren die Formeln

$$\text{vol}(\Psi_{\text{aff}}(Q)) = |\det \mathbf{A}| \cdot \text{vol}(Q),$$

d.h die Determinante des linearen Anteils \mathbf{A} beschreibt gerade die Volumenverzerrung. Das Betragszeichen ist dabei sehr wichtig, weil wir mit positiven, und nicht mit vorzeichenbehafteten Volumina rechnen.

Für eine nicht-affine Abbildung Ψ besitzt die Bildmenge $\Psi(Q)$ eine kompliziertere Struktur und die Berechnung ihres exakten Volumens kann sehr schwierig sein. Bei kleinen Quadern können wir aber immer Ψ nach dem Satz von Taylor durch eine affine Funktion approximieren und ein Analogon zur oberen Formel benutzen.

Lemma (Volumentransformation kleiner Quader) Es gilt

$$\text{vol}(\Psi(Q)) = |\det(\mathbf{J}\Psi(\boldsymbol{\xi}))| \cdot \text{vol}(Q) + O\left((\text{diam}(Q))^{n+1}\right),$$

wobei $\boldsymbol{\xi} \in Q$ eine beliebig gewählte Stützstelle ist und der Durchmesser

$$\text{diam}(Q) := \max \{ |\mathbf{x} - \tilde{\mathbf{x}}| : \mathbf{x}, \tilde{\mathbf{x}} \in Q \}$$

proportional zur maximalen Kantenlänge von Q ist.

Beweisidee Wir wollen keinen rigorosen Beweis führen, sondern nur die wesentlichen Ideen skizzieren. Der Satz von Taylor garantiert

$$\Psi(\mathbf{x}) = \Psi_{\xi}(\mathbf{x}) + O((\text{diam}(Q))^2), \quad \Psi_{\xi}(x) := \Psi(\xi) + \mathbf{J}\Psi(\xi)(\mathbf{x} - \xi),$$

wobei Ψ_{ξ} gerade das erste Taylor-Polynom von Ψ im Entwicklungspunkt ξ ist. Insbesondere ist Ψ_{ξ} eine affine Abbildung mit linearem Teil $\mathbf{A} = \mathbf{J}\Psi(\xi)$ sowie den Verschiebungsanteilen $\mathbf{c} = \Psi(\xi)$ und $\mathbf{d} = \xi$. Es gilt nun

$$\text{vol}(\Psi_{\xi}(Q)) = |\det(\mathbf{J}\Psi(\xi))| \cdot \text{vol}(Q)$$

und unsere geometrische Intuition sagt, dass für kleine Quader

$$\text{vol}(\Psi_{\xi}(Q)) \approx \text{vol}(\Psi(Q))$$

gilt (siehe Bild). Die mathematische Schwierigkeit liegt nun darin, zum einen die Messbarkeit der deformierten Bildmenge $\Psi(Q)$ zu beweisen und zum anderen alle Fehlerterme rigoros abzuschätzen. \square

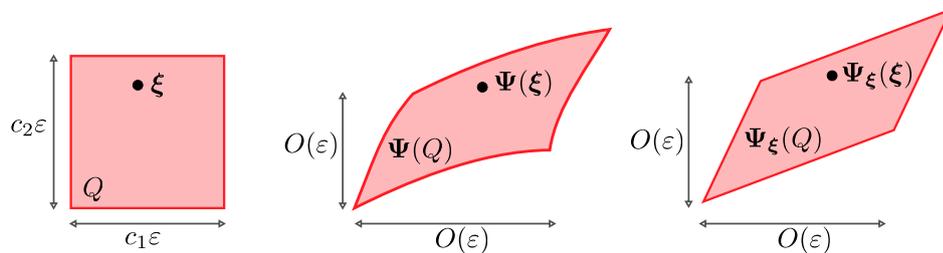


Abbildung Ein zweidimensionaler Quader (links) wird unter einer Abbildung Ψ in die deformierte Menge $\Psi(Q)$ (Mitte) überführt. Sind die Kantenlängen des Quaders klein, so kann man Ψ durch eine Taylor-Approximation Ψ_{ξ} ersetzen, die Q in das Parallelogramm $\Psi_{\xi}(Q)$ (rechts) überführt, dessen Volumen exakt berechnet werden kann. Der Taylor-Entwicklungspunkt ξ wird in der Integrationstheorie auch Stützstelle genannt.

Herleitung des Transformationsatzes

Theorem (allgemeine Transformationsformel) Es gilt

$$\int_E g(\mathbf{y}) \, d\mathbf{y} = \int_D f(\mathbf{x}) |\det(\mathbf{J}\Psi(\mathbf{x}))| \, d\mathbf{x}$$

mit den oben eingeführten Notationen und Voraussetzungen.

Beweisidee Wir approximieren die messbare Menge $D \subset U$ durch eine große, aber endliche Anzahl von Quadern

$$Q_1, \dots, Q_M$$

die ungefähr gleich groß sind und sich nur in Randpunkten gegenseitig überlappen (siehe die roten Quader im Bild). Die Details sowie die genaue Nummerierung der Quader wird dabei keine Rolle spielen. Wichtig ist nur, dass

$$\text{diam}(Q_m) = O(\varepsilon), \quad \text{vol}(Q_m) = O(\varepsilon^n)$$

gelten wird, wobei

$$\varepsilon := 1/\sqrt[n]{M} \ll 1$$

eine positive, aber kleine Zahl ist, die die Feinheit der Quader-Approximation von D quantifiziert. Außerdem wählen wir in jedem Quader eine Stützstelle $\xi_m \in Q_m$ (zum Beispiel den Mittelpunkt oder einen der Eckpunkte). Unter Ψ wird jeder dieser Quader Q_m auf einen deformierten Quader $E_m = \Psi(Q_m)$ abgebildet, dessen Volumen durch

$$\text{vol}(\Psi(Q_m)) = |\det(\mathbf{J}\Psi(\xi_m))| \text{vol}(Q_m) + O(\varepsilon^{n+1})$$

approximiert werden kann, und alle deformierten Quader E_m zusammen approximieren die Bildmenge $E = \Psi(D)$. Es gilt nun

$$\int_E g(\mathbf{y}) \, d\mathbf{y} = \sum_{m=1}^M \int_{\Psi(Q_m)} g(\mathbf{y}) \, d\mathbf{y} + O(\varepsilon),$$

da nur $M^{\frac{n-1}{n}} \ll M$ viele — also relative gesehen nur sehr wenige — dieser Quader den Rand von D überdecken werden, und diese insgesamt ein kleines Volumen der Größenordnung $O(\varepsilon^n \varepsilon^{1-n}) = O(\varepsilon)$ besitzen (hellrote Quader im Bild). Wir können nun jedes Teilintegral nach dem Satz von Taylor durch

$$\begin{aligned} \int_{\Psi(Q_m)} g(\mathbf{y}) \, d\mathbf{y} &= \left(g(\Psi(\xi_m)) + O(\varepsilon) \right) \text{vol}(\Psi(Q_m)) \\ &= g(\Psi(\xi_m)) \text{vol}(\Psi(Q_m)) + O(\varepsilon^{n+1}) \end{aligned}$$

approximieren, wobei wir diesmal g (und nicht die Komponenten von Ψ) durch das entsprechende nullte Taylor-Polynom im Entwicklungspunkt $\Psi(\xi_m)$ ersetzt und

$$\text{diam}(\Psi(Q_m)) = O(\varepsilon), \quad \text{vol}(\Psi(Q_m)) = O(\varepsilon^n)$$

benutzt haben. Insgesamt — und wegen $M = O(\varepsilon^{-n})$ sowie $\sum_{m=1}^M O(\varepsilon^{n+1}) = O(\varepsilon)$ — erhalten wir die Näherungsformel

$$\int_E g(\mathbf{y}) \, d\mathbf{y} = \left(\sum_{m=1}^M g(\Psi(\xi_m)) |\det(\mathbf{J}\Psi(\xi_m))| \text{vol}(Q_m) \right) + O(\varepsilon),$$

wobei der Fehlerterm im Limes $\varepsilon \rightarrow 0$ bzw. $M \rightarrow \infty$ verschwindet. Die rechte Seite ist aber gerade die Riemann-Summen-Approximation des Integrals auf der rechten Seite der Behauptung, wobei wir analog zu unserer bisherigen Diskussion zeigen können, dass der entsprechende Approximationsfehler auch die Ordnung $O(\varepsilon)$ besitzt.

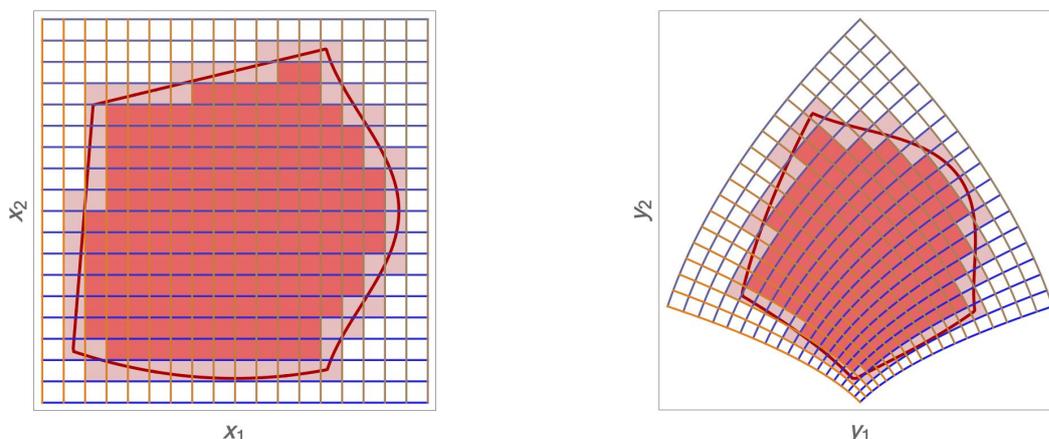


Abbildung Zum Beweis des Transformationssatzes in 2D. *Links:* Die rote Kurve stellt den Rand des kompakten Integrationsgebietes D dar und die roten Quader eine Approximation dieses Gebiets, wobei hier innere Quader und Randquader unterschiedlich schattiert sind. Die Quader könnten mit Multi-Indizes identifiziert werden, aber es ist einfacher, sie irgendwie durchnummerieren (Q_1, \dots, Q_M). *Rechts:* Das entsprechende Bild nach Anwendung von Ψ .

Bemerkungen

1. Der Beweis des Transformationssatzes illustriert sehr schön die folgenden zwei Grundprinzipien der Riemannschen Integrationstheorie:
 - (a) Integrale sind nichts anderes als Grenzwerte von Summen, wobei sehr viele kleine Summanden aufaddiert werden. Im Limes hat man dann — salopp gesprochen — unendlich viele, jeweils unendlich kleine Summanden. Unendlich kleine Größen nennt man außerhalb der Mathematik oftmals auch *infinitesimal klein*.
 - (b) Eine messbare Menge kann durch große Vereinigungen kleiner Mengen approximiert werden, wobei die kleinen Mengen zum Beispiel Quader oder auch deformierte Quader sein können.
2. Die Transformationsformel ist viel einfacher, als sie auf den ersten Blick aussieht, sofern man die symbolische Substitution

$$d\mathbf{y} = \left| \det \left(\frac{\partial \mathbf{y}}{\partial \mathbf{x}} \right) \right| d\mathbf{x}$$

benutzt, wobei $\frac{\partial \mathbf{y}}{\partial \mathbf{x}}$ eine alternative Notation für die Jacobi-Matrix $\mathbf{J}\Psi$ ist. Genauer gesagt: Um das Integral von g bzgl. $\mathbf{y} \in E$ auszurechnen, gehen wir wie folgt vor:

- (a) Wir substituieren \mathbf{y} durch \mathbf{x} im Integranden (das ersetzt $g(\mathbf{y})$ durch $f(\mathbf{x})$).
- (b) Wir passen das Integrationsgebiet in konsistenter Weise an (aus E wird D).
- (c) Wir ersetzen mit obiger Regel $d\mathbf{y}$ durch $d\mathbf{x}$, wobei die Jacobi-Matrix von Ψ ganz natürlich ins Spiel kommt.

Das ist dieselbe Strategie, die wir in *Analysis 1* schon bei der Substitutionsmethode für die Transformation eindimensionaler Integrale (also für $n = 1$) benutzt hatten. Bei $n > 1$ ist jetzt eben alles vektorwertig und wir dürfen nicht vergessen, den Betrag der Determinante einzubauen. Das ist notwendig, weil \mathbf{x} und \mathbf{y} zwar Vektoren sind, aber $d\mathbf{x}$ und $d\mathbf{y}$ infinitesimal kleine *Zahlen* repräsentieren (nämlich das Volumen unendlich kleiner Quader).

3. Aus Symmetriegründen (bzw. nach konsistenter und konsequenter Vertauschung von Symbolen) gilt auch

$$\int_D f(\mathbf{x}) d\mathbf{x} = \int_E g(\mathbf{y}) |\det(\mathbf{J}\Phi(\mathbf{y}))| d\mathbf{y},$$

und in dieser Fassung wird der Transformationssatz meist verwendet: Das Integral einer gegebenen skalaren Funktion f in der vektoriellen Variablen \mathbf{x} über einer geometrisch komplizierten Menge D kann zum Beispiel dadurch berechnet werden, dass wir eine Koordinatentransformation Φ suchen, sodass $D = \Phi(E)$ via $\mathbf{x} = \Phi(\mathbf{y})$ gerade das Bild einer möglichst einfachen Menge E ist. Der zu zahlende Preis ist, dass wir die Transformation erst einmal finden und anschließend den Integranden transformieren müssen, sodass am Ende alles von \mathbf{y} abhängt. Das Auffinden von Φ ist dabei meist der schwierigste Teil.

Allgemeine Riemann-Summen* Wir hatten im Beweis des Transformationsatzes das folgende, allgemeine Prinzip verwendet (mit E und \mathbf{y} statt D und \mathbf{x}). Ist D eine messbare Menge und

$$D = D_1 \cup \dots \cup D_M$$

eine Zerlegung $Z = \{D_1, \dots, D_M\}$ von D in M messbare Mengen (zum Beispiel in deformierte Quader), die sich nur in Randpunkten überlappen, so nennt man

$$\|Z\| = \max_{m=1 \dots M} \text{diam}(D_m) \quad \text{diam}(D_m) := \max \{ |\mathbf{x} - \tilde{\mathbf{x}}| : \mathbf{x}, \tilde{\mathbf{x}} \in D_m \}$$

die Feinheit der Zerlegung. Ist nun eine integrierbare Funktion f auf D gegeben und wird nun in jedem D_m eine Stützstelle $\xi_m \in D_m$ gewählt, so kann die Riemann-Summe

$$R(Z) = \sum_{m=1}^M f(\xi_m) \text{vol}(D_m)$$

gebildet werden. Es gilt dann immer $\int_D f(\mathbf{x}) \, d\mathbf{x} = \lim_{\|Z\| \rightarrow 0} R(Z)$, d.h. bei immer feiner werdender Zerlegung konvergiert die Riemann-Summe gegen das Integral von f über Q . Für den Fall, dass D und D_m selbst alle Quader sind, ergibt sich gerade wieder das am Anfang dieses Kapitels eingeführte Konzept. Wird außerdem ξ_m als Minimierer bzw. Maximierer von f in D_m gewählt, so spricht man wieder von Unter- bzw. Obersumme.

Anwendung des Transformationsatzes

Beispiel Wir betrachten die Transformation

$$y_1 = \Psi_1(x_1, x_2) = \frac{-2x_2}{(1-x_1)^2 + x_2^2}, \quad y_2 = \Psi_2(x_1, x_2) = \frac{1-x_1^2 - x_2^2}{(1-x_1)^2 + x_2^2}$$

sowie

$$x_1 = \Phi_1(y_1, y_2) = \frac{y_1^2 + y_2^2 - 1}{y_1^2 + (1+y_2)^2}, \quad x_2 = \Phi_2(y_1, y_2) = \frac{-2y_1}{y_1^2 + (1+y_2)^2},$$

wobei sowohl Φ als auch Ψ zu Klasse der *Möbius-Transformationen* (siehe dazu die Vorlesung *Funktionentheorie*) gehören und man üblicherweise die entsprechenden Definitionsbereiche U bzw. V via

$$U := \{(x_1, x_2) : x_1^2 + x_2^2 < 1\}, \quad V := \{(y_1, y_2) : y_2 > 0\}$$

als offene Einheitskreisscheibe bzw. als obere offene Halbebene wählt. Wir berechnen

$$\mathbf{J}\Psi(x_1, x_2) = \frac{2}{((1-x_1)^2 + x_2^2)^2} \begin{pmatrix} -2(1-x_1)x_2 & -(1-x_1^2) + x_2^2 \\ +(1-x_1^2) - x_2^2 & -2(1-x_1)x_2 \end{pmatrix}$$

sowie

$$\mathbf{J}\Phi(y_1, y_2) = \frac{2}{(y_1^2 + (1+y_2)^2)^2} \begin{pmatrix} +2y_1(1+y_2) & -y_1^2 + (1+y_2)^2 \\ +y_1^2 - (1+y_2)^2 & +2y_1(1+y_2) \end{pmatrix}$$

und erhalten

$$\det(\mathbf{J}\Psi(x_1, x_2)) = \frac{4}{((1-x_1)^2 + x_2^2)^2}, \quad \det(\mathbf{J}\Phi(y_1, y_2)) = \frac{4}{(y_1^2 + (1+y_2)^2)^2}.$$

Insbesondere ergeben sich aus den obigen Notationen und Ausdrücken mit $f \equiv 1$ und $g \equiv 1$ die Volumenformeln

$$\text{vol}(D) = \iint_E \frac{4 \, dy_1 \, dy_2}{(y_1^2 + (1 + y_2)^2)^2}, \quad \text{vol}(E) = \iint_D \frac{4 \, dx_1 \, dx_2}{((1 - x_1)^2 + x_2^2)^2},$$

d.h. das Volumen der Menge $D \subset U$ im \mathbf{x} -Bereich kann durch ein Integral über die entsprechende Menge $E \subset V$ im \mathbf{y} -Bereich dargestellt werden und umgekehrt. Für den Quader $D_1 = [-\frac{1}{2}, +\frac{1}{2}] \times [-\frac{1}{2}, +\frac{1}{2}]$ erhalten wir durch Auswertung der iterierten Integrale

$$\text{vol}(E_1) = 4\pi + \frac{8}{9} \left(+6 - \text{arctan}(3) - 9 \text{arctan}(3) \right) \approx 7.261$$

für das Volumen seines Bildes E_1 unter Ψ , aber die exakte Berechnung der Integrale erfordert mehrere, nicht triviale Zwischenschritte, die wir hier nicht darstellen können. Analog folgt

$$\text{vol}(D_2) = \frac{1}{8} \left(4 + \pi - \text{arctan} \left(\frac{44}{117} \right) \right) \approx 0.848$$

mit $D_2 = \Phi(E_2)$ für den Quader $E_2 = [0, 1] \times [0, 1]$.

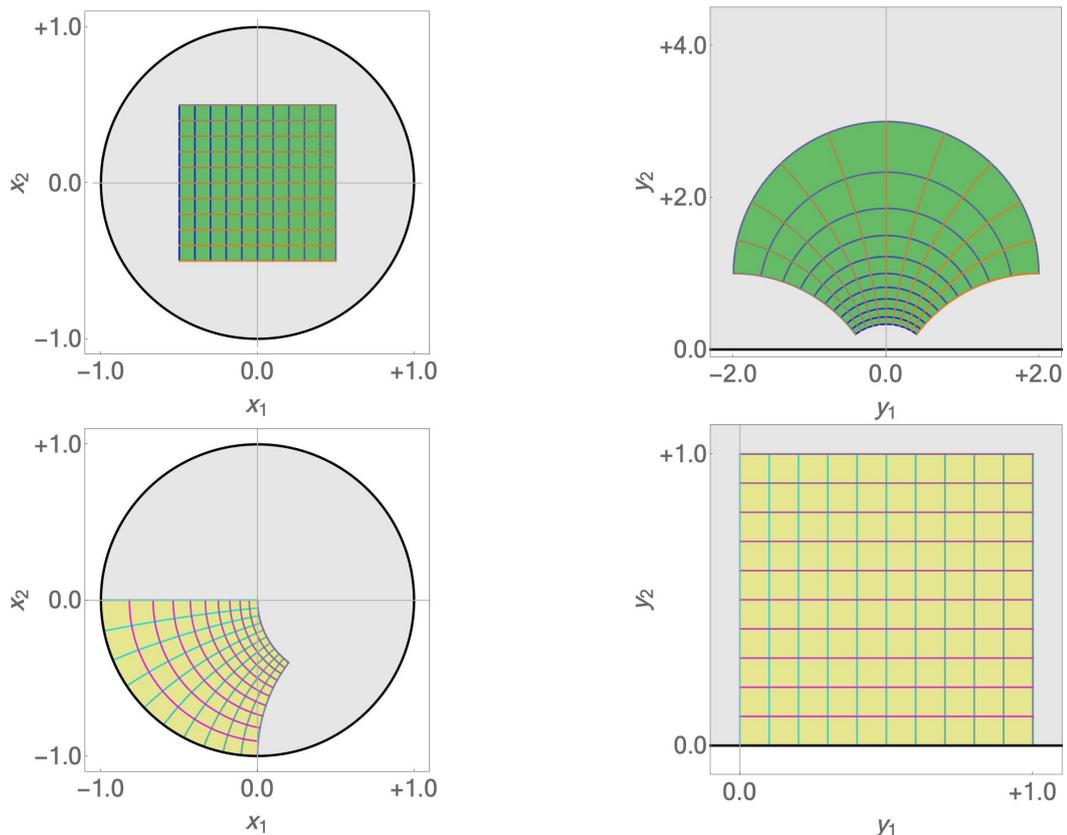


Abbildung Die Bilder zum eben gerechneten Beispiel, wobei die Kreisscheibe U sowie die Halbebene V in Grau dargestellt sind und Φ sowie Ψ als Möbius-Transformationen winkeltreu sind. Eine weitere bemerkenswerte Eigenschaft ist, dass Kreissegmente auf Kreissegmente abgebildet werden, wobei Geraden auch Kreise, nur mit unendlich großem Radius, sind. *Oben:* Die Menge D_1 (grün links) ist ein Quader und das Volumen der deformierten Menge $E_1 = \Psi(D_1)$ (grün rechts) kann durch ein \mathbf{x} -Integral über D_1 berechnet werden. *Unten:* Diesmal ist E_2 (gelb, rechts) ein Quader, so dass das Volumen der deformierten Menge $D_2 = \Phi(E_2)$ (gelb, links) als Integral bzgl. $\mathbf{y} \in E_2$ berechnet werden kann.

Beispiel Wir rechnen diesmal in Physiker-Notation. Die Formeln

$$x = r \cos(\theta), \quad y = r \sin(\theta)$$

beschreiben den Zusammenhang zwischen den kartesischen Koordinaten der Ebene (hier x und y genannt) und den Polarkoordinaten (Radius r und Winkel θ). Einfache Rechnungen zeigen

$$\frac{\partial(x, y)}{\partial(r, \theta)} = \begin{pmatrix} \cos(\theta) & -r \sin(\theta) \\ \sin(\theta) & r \cos(\theta) \end{pmatrix}, \quad \left| \det \left(\frac{\partial(x, y)}{\partial(r, \theta)} \right) \right| = r$$

und liefern

$$dx \, dy = r \, dr \, d\theta$$

als symbolische Transformationsregel für infinitesimale Volumina. Die Transformationsformel kann nun als

$$\iint_D f(x, y) \, dx \, dy = \iint_E g(r, \theta) \, r \, dr \, d\theta$$

geschrieben werden, wobei die Konsistenz und damit Richtigkeit der Formel durch

$$(x, y) \in D \Leftrightarrow (r, \theta) \in E, \quad f(x, y) = g(r, \theta)$$

sichergestellt wird. Mit der speziellen Wahl

$$f(x, y) = 1, \quad g(r, \theta) = 1$$

und

$$D = \{(x, y) : x^2 + y^2 \leq \varrho^2\}, \quad E = \{(r, \theta) : 0 \leq r \leq \varrho, 0 \leq \theta \leq 2\pi\}$$

erhalten wir

$$\text{vol}(D) = \iint_E r \, dr \, d\theta = \int_0^{2\pi} \left(\int_0^{\varrho} r \, dr \right) d\theta = \int_0^{2\pi} \left[\frac{1}{2} r^2 \right]_{r=0}^{r=\varrho} d\theta = \int_0^{2\pi} \frac{1}{2} \varrho^2 \, d\theta = \frac{1}{2} \pi \varrho^2$$

und haben damit das Volumen von D , also der kartesischen Kreisscheibe mit Radius ϱ , durch ein Integral über den polaren Quader E berechnet. Mit der alternativen Wahl

$$f(x, y) = x + y^2, \quad g(r, \theta) = r \cos(\theta) + r^2 \sin^2(\theta)$$

ergibt sich

$$\begin{aligned} \iint_D (x + y^2) \, dx \, dy &= \iint_E (r \cos(\theta) + r^2 \sin^2(\theta)) \, r \, dr \, d\theta \\ &= \int_0^{\varrho} \left(\int_0^{2\pi} (r^2 \cos(\theta) + r^3 \sin^2(\theta)) \, d\theta \right) dr \\ &= \int_0^{\varrho} (\pi r^3) \, dr = \frac{1}{4} \pi \varrho^4 \end{aligned}$$

durch direkte Rechnungen.

über Entartungspunkte Bei den Polarkoordinaten treten gewisse Entartungen und Singularitäten auf. Zum einen entspricht der kartesische Koordinatenursprung — also der nulldimensionale Punkt mit $x = 0$ und $y = 0$ — der polaren Koordinatenachse $r = 0$, also einer eindimensionalen Kurve. Diese Entartung spiegelt sich auch darin wider, dass die Jacobi-Matrix

$$\frac{\partial(r, \theta)}{\partial(x, y)} = \left(\frac{\partial(x, y)}{\partial(r, \theta)} \right)^{-1} = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -r^{-1} \sin(\theta) & r^{-1} \cos(\theta) \end{pmatrix}$$

für $r = 0$ gar nicht definiert ist. Außerdem ist die Zuordnung $(r, \theta) \rightarrow (x, y)$ nicht eindeutig, sondern 2π -periodisch in θ . Man kann natürlich Winkelrestriktionen (zum Beispiel $0 \leq \theta \leq 2\pi$ oder $-\pi \leq \theta \leq +\pi$) stellen, aber auch dann gibt es gewisse Subtilitäten. Bei Kugelkoordinaten gibt es ähnliche Probleme mit der Mehrdeutigkeit der Winkel und der Nicht-Differenzierbarkeit für $r = 0$.

Merkregel: Die Transformationsformel bleibt bei gewissen (aber nicht bei allen) Entartungen richtig, besonders wenn die Probleme nur am Rand der betrachteten Mengen auftreten. Man muss dann aber immer besonders aufpassen und sollte alle Rechenergebnisse zumindest auf Plausibilität prüfen.

Beispiel Eine klassische Anwendung der zweidimensionalen Transformationsformel mit Polarkoordinaten auf unbeschränkten Gebieten ist die Berechnung des Integrals unter einer Gaußschen Glockenkurve. Mit den Notationen aus dem letzten Beispiel gilt zum einen

$$\begin{aligned} \iint_{\mathbb{R}^2} \exp(-x^2 - y^2) \, dx \, dy &= \int_0^\infty \left(\int_0^{2\pi} \exp(-r^2) r \, d\theta \right) dr \\ &= 2\pi \int_0^\infty r \exp(-r^2) \, dr = 2\pi \left[-\frac{1}{2} \exp(-r^2) \right]_{r=0}^{r=\infty} \\ &= \pi, \end{aligned}$$

wobei wir $\exp(-\infty) = \lim_{r \rightarrow \infty} \exp(-r^2) = 0$ verwendet haben. Zum anderen impliziert der Satz von Fubini

$$\iint_{\mathbb{R}^2} \exp(-x^2 - y^2) \, dx \, dy = \left(\int_{-\infty}^{+\infty} \exp(-x^2) \, dx \right) \left(\int_{-\infty}^{+\infty} \exp(-y^2) \, dy \right),$$

und weil die beiden Integrale auf der rechten Seite gleich sind, erhalten wir die bekannte Formel

$$\int_{-\infty}^{+\infty} \exp(-x^2) \, dx = \sqrt{\pi}$$

durch Kombination der beiden Teilergebnisse. Beachte, dass wir hier zwei Theoreme über zweidimensionale Integrale verwendet haben, um ein eindimensionales Integral zu berechnen.

Affine Transformation Im Fall einer affinen Transformation

$$\mathbf{y} = \Psi(\mathbf{x}) = \mathbf{c} + \mathbf{A} \cdot (\mathbf{x} - \mathbf{d}), \quad \mathbf{x} = \Phi(\mathbf{y}) = \mathbf{d} + \mathbf{A}^{-1} \cdot (\mathbf{y} - \mathbf{c})$$

mit invertierbarer Matrix $\mathbf{A} \in \mathbb{R}^{(n,n)}$ und Verschiebungsanteilen $\mathbf{c}, \mathbf{d} \in \mathbb{R}^n$ sind beide Jacobi-Matrizen konstant, d.h. es gilt

$$\mathbf{J}\Psi(\mathbf{x}) = \mathbf{A}, \quad \mathbf{J}\Phi(\mathbf{y}) = \mathbf{A}^{-1}.$$

Die Transformationsformel kann daher vereinfacht werden und impliziert via

$$\text{vol}(D) = |\det(\mathbf{A})| \text{vol}(E), \quad \text{vol}(E) = |\det(\mathbf{A}^{-1})| \text{vol}(D),$$

dass sich das Volumen bei der Komposition von Verschiebungen und Rotationen oder Spiegelungen wegen $\det(\mathbf{A}) = \pm 1$ nicht ändert.

Skalierung von Volumina Affine Transformationen der Bauart

$$\mathbf{y} = \Psi(\mathbf{x}) = \lambda^{+1} \mathbf{x}, \quad \mathbf{x} = \Phi(\mathbf{y}) = \lambda^{-1} \mathbf{y}$$

mit Parameter $\lambda > 0$ beschreiben uniforme Streckungen bzw. Stauchungen und für Quader gilt zum Beispiel

$$\mathbf{x} \in D = [a_1, b_1] \times \dots \times [a_n, b_n] \quad \Leftrightarrow \quad \mathbf{y} \in E = [\lambda a_1, \lambda b_1] \times \dots \times [\lambda a_n, \lambda b_n],$$

In Physikernotation berechnen wir

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \lambda^{+1} \mathbf{I}, \quad \frac{\partial \mathbf{x}}{\partial \mathbf{y}} = \lambda^{-1} \mathbf{I}$$

und damit

$$d\mathbf{x} = \lambda^{-n} d\mathbf{y}, \quad d\mathbf{y} = \lambda^{+n} d\mathbf{x},$$

wobei sich die Potenzen $-n$ bzw. $+n$ aus der Multi-Linearität der Determinante — d.h. aus der Formel $\det(\lambda^{\pm 1} \mathbf{A}) = \lambda^{\pm n} \det(\mathbf{A})$ — ergeben (und gerne mal vergessen werden). Insbesondere erhalten wir

$$\text{vol}(D) = \lambda^{-n} \text{vol}(E), \quad \text{vol}(E) = \lambda^{+n} \text{vol}(D)$$

als mathematische Beschreibung einer intuitiven Tatsache: Der Übergang von \mathbf{x} zu \mathbf{y} bzw. von D zu E multipliziert alle Längen mit λ und daher das n -dimensionale Volumen mit λ^n .

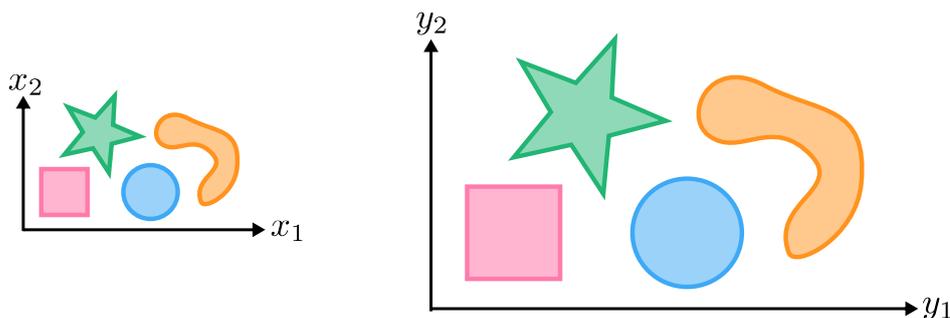


Abbildung Skalierungstransformation in 2D mit $\lambda = 2$. Alle Längen werden verdoppelt, alle Flächeninhalte jedoch vervierfacht. Alle Winkel bleiben erhalten.

2.8 Integralsätze von Gauß und Stokes in 2D

Ziel Wir wollen die Integralsätze von Gauß und Stokes in zwei Dimensionen in ihrer allgemeinen Form angeben (d.h. nicht nur auf Quadern). Damit wird es uns im nächsten Kapitel leichter fallen, die dreidimensionalen Versionen zu verstehen.

Geometrische Vorüberlegungen

Setting In diesem Abschnitt betrachten wir eine messbare (also auch kompakte Menge) $D \subset \mathbb{R}^2$, deren Rand

$$\partial D := \{\mathbf{x} \in \mathbb{R}^2 : \mathbf{x} \text{ ist Randpunkt von } D\} \subset \mathbb{R}^2,$$

durch endlich viele (meist sehr wenige) parametrisierte Kurven beschrieben werden kann. Genauer gesagt, im Folgenden nehmen wir an, dass K parametrisierte Kurven $\mathbf{c}_k : [t_{k,0}, t_{k,1}] \rightarrow \mathbb{R}^2$ mit den folgenden Eigenschaften existieren:

1. Es gilt

$$\partial D := \{\mathbf{c}_1(t) : t \in [t_{1,0}, t_{1,1}]\} \cup \dots \cup \{\mathbf{c}_K(t) : t \in [t_{K,0}, t_{K,1}]\},$$

d.h. ∂D ist die Vereinigung der Bilder der parametrisierten Kurven \mathbf{c}_k .

2. Jede Kurve \mathbf{c}_k ist zumindest stückweise stetig differenzierbar.
3. Die Kurven liefern eine minimale Parametrisierung in dem Sinne, dass es nur endlich viele Punkte $\mathbf{x} \in \partial D$ gibt, für die es via $\mathbf{x} = \mathbf{c}_k(t_k) = \mathbf{c}_l(t_l)$ mehr als eine Darstellung gibt.

Hinweis: Die nachfolgendem Beispiele werden verdeutlichen, was diese Bedingungen im Einzelnen meinen.

Bemerkungen

1. Es mag auf den ersten Blick verwirrend sein, dass man das Symbol ∂ sowohl für Ableitungen von Funktionen als auch für Ränder von Mengen benutzt. Es hat sich aber so eingebürgert. Außerdem gibt es eine innere Verbindung der Konzepte *Ableitung* und *Rand*, die sich (zum Beispiel) in den Integralsätzen manifestiert.
2. Es ist zugelassen, dass ∂D endlich viele *Ecken* oder *Spitzen* besitzt, in denen die \mathbf{c}_k nicht differenzierbar sind. Insbesondere darf D auch ein Dreieck, ein Viereck oder eine andere *polygonal berandete* Menge sein.
3. Wir wollen auch geometrische *Doppelpunkte* nicht ausschließen. Man kann sie zwar durch Aufteilung von D vermeiden, aber manchmal tauchen Sie in sehr natürlicher Weise auf (siehe einige der Beispiele unten).
4. Bei uns ist D immer abgeschlossen und deshalb gilt $\partial D \subset D$. Man könnte aber auch offene Mengen D zulassen (für diese gilt $\partial D \subset \mathbb{R}^2 \setminus D$). In den Anwendungen ist es jedoch meist nicht wichtig, ob D als offen oder abgeschlossen (oder weder noch) angesetzt wird.

5. Die Komponenten von \mathbf{c}_k bezeichnen wir wie üblich mit $c_{k,j}$, wobei $k = 1 \dots K$ und $j = 1$ oder $j = 2$ gilt. Im Fall von $K = 1$ schreiben wir wieder \mathbf{c} statt \mathbf{c}_1 bzw. c_j statt $c_{1,j}$.

Beispiele

1. Ist $D = [a_1, b_1] \times [a_2, b_2]$ ein Quader, so hatten wir bereits explizit beschrieben, wie ∂D durch eine einzige Funktion \mathbf{c} ($K = 1$) parametrisiert werden kann, wobei diese Funktion dann stückweise stetig differenzierbar ist. Alternativ können wir die vier Kantensegmente des Quader separat durch vier stetig differenzierbare Kurven beschreiben ($K = 4$), die sich gerade in den vier Eckpunkten überlappen. Bei konkreten Rechnungen hatten wir das eigentlich auch schon getan und die vier Teilparametrisierungen mit \mathbf{c}_u , \mathbf{c}_r , \mathbf{c}_o und \mathbf{c}_l bezeichnet.

Bemerkung: Wir haben sehr viele Freiheiten, die Parametrisierung von ∂D zu wählen. In der Praxis besteht die Kunst darin, eine „optimale“ zu finden, für die die Rechnungen möglichst einfach werden.

2. Die Formeln

$$\overline{B}_\varrho(\mathbf{m}) = \{\mathbf{x} \in \mathbb{R}^2 : |\mathbf{x} - \mathbf{m}| \leq \varrho\}, \quad \partial \overline{B}_\varrho(\mathbf{m}) = \{\mathbf{x} \in \mathbb{R}^2 : |\mathbf{x} - \mathbf{m}| = \varrho\}$$

beschreiben, dass der Rand der abgeschlossenen Kreisscheibe mit Radius ϱ und Mittelpunkt $\mathbf{m} \in \mathbb{R}^2$ die entsprechende Kreislinie ist. (Analoges gilt für die offene Kreisscheibe und das Wort *Kreis* kann sowohl die Scheibe als auch die Linie bezeichnen). Wir können nun den Rand in naheliegender Weise durch

$$\mathbf{c}(t) = \begin{pmatrix} m_1 \\ m_2 \end{pmatrix} + \varrho \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}, \quad t \in [0, 2\pi]$$

parametrisieren, aber es gibt natürlich auch andere Parametrisierungen.

3. Für zwei gegebene Radien $0 < \varrho_1 < \varrho_2 < \infty$ besteht der Rand des *abgeschlossenen Kreisringes*

$$\overline{A}_{\varrho_1, \varrho_2}(\mathbf{m}) := \overline{B}_{\varrho_2}(\mathbf{m}) \setminus B_{\varrho_1}(\mathbf{m}) = \{\mathbf{x} \in \mathbb{R}^2 : \varrho_1 \leq |\mathbf{x} - \mathbf{m}| \leq \varrho_2\}$$

aus den zwei Kreislinien

$$\begin{aligned} \partial \overline{A}_{\varrho_1, \varrho_2}(\mathbf{m}) &:= \partial \overline{B}_{\varrho_1}(\mathbf{m}) \cup \partial \overline{B}_{\varrho_2}(\mathbf{m}) \\ &= \{\mathbf{x} \in \mathbb{R}^2 : |\mathbf{x} - \mathbf{m}| = \varrho_1 \text{ oder } |\mathbf{x} - \mathbf{m}| = \varrho_2\}. \end{aligned}$$

Dieser Rand, der offensichtlich aus zwei Komponenten besteht, kann mit $K = 2$ durch

$$\mathbf{c}_1(t) = \begin{pmatrix} c_{1,1}(t) \\ c_{1,2}(t) \end{pmatrix} = \begin{pmatrix} m_1 \\ m_2 \end{pmatrix} + \varrho_1 \begin{pmatrix} +\cos(t) \\ -\sin(t) \end{pmatrix} = \begin{pmatrix} m_1 \\ m_2 \end{pmatrix} + \varrho_1 \begin{pmatrix} \cos(-t) \\ \sin(-t) \end{pmatrix}$$

und

$$\mathbf{c}_2(t) = \begin{pmatrix} c_{2,1}(t) \\ c_{2,2}(t) \end{pmatrix} = \begin{pmatrix} m_1 \\ m_2 \end{pmatrix} + \varrho_2 \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}$$

parametrisiert werden, wobei jeweils $t \in [0, 2\pi]$ gilt. Beachte, dass die äußere bzw. innere Randkreislinie entgegen bzw. mit dem Uhrzeigersinn durchlaufen wird, sodass der Umlaufsinn mit der jeweils positiven Orientierung am Rand (siehe unten) übereinstimmt.

4. Die *Doppelhorn-Kurve*

$$\mathbf{c}(t) = \begin{pmatrix} \sin(t) \\ \frac{\cos^2(t)(2 + \cos(t))}{3 + \sin^2(t)} \end{pmatrix}, \quad t \in [0, 2\pi]$$

parametrisiert den Rand einer messbaren Menge D , wobei der Geschwindigkeitsvektor $\dot{\mathbf{c}}(t)$ in den beiden Eckpunkten ($t = \frac{1}{2}\pi$, $x_1 = +1$, $x_2 = 0$ bzw. $t = \frac{3}{2}\pi$, $x_1 = -1$, $x_2 = 0$) jeweils verschwindet.

5. Das *Zweiblatt*

$$\mathbf{c}(t) = \sin^2(t) \cos(t) \begin{pmatrix} \sin(t) \\ \cos(t) \end{pmatrix}, \quad t \in [0, \pi]$$

parametrisiert auch den Rand einer messbaren Menge D . Diesmal gibt es einen entarteten Doppelpunkt, nämlich $\mathbf{x} = \mathbf{0}$, der von der angegebenen Parametrisierung mehrmals realisiert wird. Insbesondere gilt

$$\mathbf{c}(0) = \mathbf{c}\left(\frac{1}{2}\pi\right) = \mathbf{c}(\pi) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \dot{\mathbf{c}}(0) = \dot{\mathbf{c}}(\pi) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \dot{\mathbf{c}}\left(\frac{1}{2}\pi\right) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}.$$

Wir könnten D auch in zwei Teile aufspalten und diese separat betrachten, wobei dann $t \in [0, \frac{1}{2}\pi]$ oder $t \in [\frac{1}{2}\pi, \pi]$ für die Randparametrisierung zu wählen ist.

6. Die *Torpedokurve*

$$\mathbf{c}(t) = \sin(t) \cos(2t) \begin{pmatrix} \sin(t) \\ \cos(t) \end{pmatrix}, \quad t \in [0, \pi]$$

besitzt den echten Dreifachpunkt

$$\mathbf{0} = \mathbf{c}(0) = \mathbf{c}\left(\frac{1}{2}\pi\right) = \mathbf{c}\left(\frac{3}{2}\pi\right) = \mathbf{c}(\pi)$$

und die von ihr berandete Menge könnte in drei Teile aufgespalten werden.

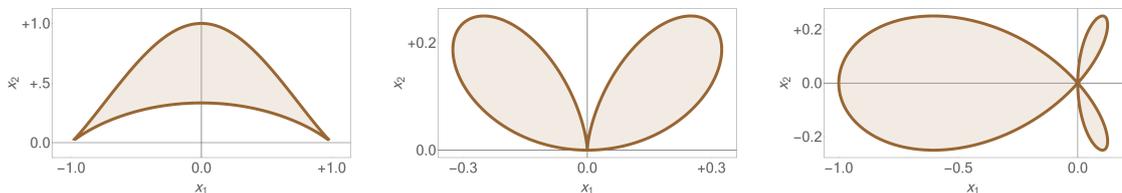


Abbildung Die drei nicht-trivialen parametrisierten Randkurven aus den Beispielen.

Normalen- und Tangentialvektor, Teil 1 In *jedem* Randpunkt $\mathbf{x} \in \partial D$ (mit Ausnahme von Eck- und Doppelpunkten) existiert ein normierter Tangentialvektor $\boldsymbol{\tau}(\mathbf{x})$ sowie ein normierter Normalenvektor $\boldsymbol{\nu}(\mathbf{x})$, so dass die folgenden Eigenschaften erfüllt sind:

1. $\boldsymbol{\tau}(\mathbf{x})$ und $\boldsymbol{\nu}(\mathbf{x})$ bilden immer eine ON-Basis des \mathbb{R}^2 , d.h. es gilt

$$\langle \boldsymbol{\tau}(\mathbf{x}), \boldsymbol{\tau}(\mathbf{x}) \rangle = 1, \quad \langle \boldsymbol{\nu}(\mathbf{x}), \boldsymbol{\nu}(\mathbf{x}) \rangle = 1, \quad \langle \boldsymbol{\tau}(\mathbf{x}), \boldsymbol{\nu}(\mathbf{x}) \rangle = 0.$$

- $\boldsymbol{\tau}(\mathbf{x})$ ist parallel zur Tangentialgeraden an ∂D im Punkt \mathbf{x} , wohingegen $\boldsymbol{\nu}(\mathbf{x})$ immer senkrecht auf dieser Geraden steht.
- $\boldsymbol{\nu}(\mathbf{x})$ zeigt nach außen und $\boldsymbol{\tau}(\mathbf{x})$ ist positiv orientiert, d.h. D liegt immer links von $\boldsymbol{\tau}(\mathbf{x})$. Siehe dazu die Bilder.

Hinweis: Das Konzept *Orientierung* ist am Anfang etwas vertrackt und es Bedarf einiger Übung, die positive (also die richtige) von der negativen (der falschen) auf der Ebene von Formeln zu unterscheiden. Malen Sie daher *immer* ein Bild, da man dort alles sehen kann. Wie (fast) immer gilt außerdem: Vertrauen Sie Ihrer geometrischen Intuition und dem gesunden Menschenverstand!

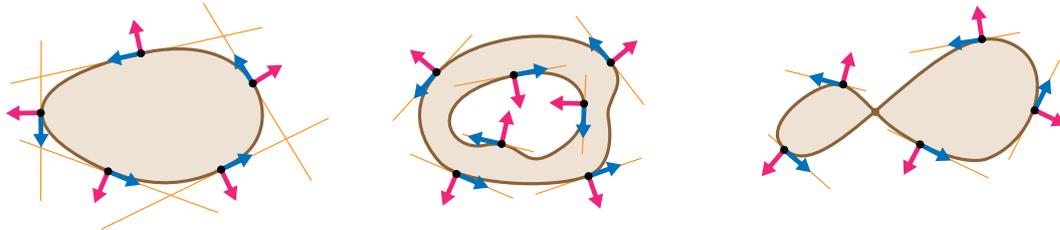


Abbildung Drei Beispiele für eine zulässige Menge D , wobei diese im zweiten Beispiel ein Loch und im dritten einen Doppelpunkt besitzt. Die roten bzw. blauen Pfeile repräsentieren den nach außen zeigenden Normalenvektor $\boldsymbol{\nu}(\mathbf{x})$ bzw. den positiv orientierten Tangentialvektor $\boldsymbol{\tau}(\mathbf{x})$ in ausgewählten Randpunkten $\mathbf{x} \in \partial D$ (schwarz), wobei die jeweilige Tangentialgerade durch die orangen Linien angedeutet wird. *Bemerkung:* Werden die roten Pfeile umgeklappt, so entsteht gerade der nach innen zeigenden Normalenvektor. Das Umklappen der blauen ändert die Orientierung.

Bemerkungen

- $\boldsymbol{\tau}(\mathbf{x})$ und $\boldsymbol{\nu}(\mathbf{x})$ werden nur in Randpunkten von D betrachtet. Es ist nicht wichtig, ob oder wie man diese auch in inneren oder äußeren Punkten von D einführen kann.
- In Eck- oder Doppelpunkten von ∂D sind $\boldsymbol{\tau}(\mathbf{x})$ und $\boldsymbol{\nu}(\mathbf{x})$ im Allgemeinen nicht (oder nicht eindeutig) definiert.
- Es gilt immer

$$\boldsymbol{\tau}(\mathbf{x}) = \begin{pmatrix} -\nu_2(\mathbf{x}) \\ +\nu_1(\mathbf{x}) \end{pmatrix} \quad \text{und damit auch} \quad \boldsymbol{\nu}(\mathbf{x}) = \begin{pmatrix} +\tau_2(\mathbf{x}) \\ -\tau_1(\mathbf{x}) \end{pmatrix}.$$

Oder anders gesagt: Wenn $\boldsymbol{\nu}(\mathbf{x})$ in Richtung „Norden“ zeigt, so zeigt $\boldsymbol{\tau}(\mathbf{x})$ nach „Westen“. Man muss allerdings für jeden Punkt $\mathbf{x} \in \partial D$ entscheiden, wo gerade „Norden“ liegt.

- Manchmal schreibt man auch $\boldsymbol{\tau}_{\partial D}(\mathbf{x})$ und $\boldsymbol{\nu}_{\partial D}(\mathbf{x})$ um deutlich zu machen, um welchen Rand es geht. Dies ist besonders dann sinnvoll, wenn man mehrere Gebiete gleichzeitig untersucht.

Normalen- und Tangentialvektor, Teil 2 Die Vektoren $\boldsymbol{\tau}(\mathbf{x})$ und $\boldsymbol{\nu}(\mathbf{x})$ spiegeln geometrische Eigenschaften der Punktmenge ∂D wider und können unabhängig von einer Parametrisierung eingeführt werden. Mit Parametrisierungen kann man aber sehr einfach explizite Formeln angeben. Insbesondere gilt

$$\boldsymbol{\tau}(\mathbf{c}_k(t)) = \frac{\pm \dot{\mathbf{c}}_k(t)}{|\dot{\mathbf{c}}_k(t)|} = \frac{\pm 1}{\sqrt{(\dot{c}_{k,1}(t))^2 + (\dot{c}_{k,2}(t))^2}} \begin{pmatrix} \dot{c}_{k,1}(t) \\ \dot{c}_{k,2}(t) \end{pmatrix},$$

sowie

$$\boldsymbol{\nu}(\mathbf{c}_k(t)) = \frac{\pm 1}{\sqrt{(\dot{c}_{k,1}(t))^2 + (\dot{c}_{k,2}(t))^2}} \begin{pmatrix} +\dot{c}_{k,2}(t) \\ -\dot{c}_{k,1}(t) \end{pmatrix},$$

wobei das Vorzeichen für jeden Zeitparameter t so zu wählen ist, dass $\boldsymbol{\nu}(\mathbf{c}_k(t))$ nach außen zeigt und $\boldsymbol{\tau}(\mathbf{c}_k(t))$ dazu positiv orientiert ist. Für jedes feste t legt aber das Vorzeichen in der einen Formel schon das Vorzeichen in der anderen Formel fest und umgekehrt.

Bemerkungen

1. Bei „einfachen“ Mengen D kann man eine globale Wahl des Vorzeichens für alle t treffen, aber manchmal ist das nicht möglich, insbesondere wenn es Doppelpunkte gibt.
2. In konkreten Fällen kann die richtige Vorzeichenwahl meist problemlos aus einer Skizze abgelesen werden.
3. Die Vektoren $\boldsymbol{\tau}(\mathbf{c}_k(t))$ und $\boldsymbol{\nu}(\mathbf{c}_k(t))$ liefern — ggf. bis auf das Vorzeichen — gerade das Frenetsche Zweibein der parametrisierten Randkurve.

Beispiele

1. Auf dem Rand von zweidimensionalen Quadern hatten wir oben schon die entsprechenden Normalen- und Tangentialvektoren identifiziert (siehe auch die Übungsaufgaben zu Stokes und Gauß auf 2D Quadern), wobei auf jeder der vier Kanten des Randes die Werte von $\boldsymbol{\tau}(\mathbf{x})$ und $\boldsymbol{\nu}(\mathbf{x})$ konstant bzgl. \mathbf{x} sind.

2. Es gilt

$$\boldsymbol{\nu}(\mathbf{x}) = + \frac{\mathbf{x} - \mathbf{m}}{|\mathbf{x} - \mathbf{m}|} = \varrho^{-1} \begin{pmatrix} +(x_1 - m_1) \\ +(x_2 - m_2) \end{pmatrix}, \quad \boldsymbol{\tau}(\mathbf{x}) = \varrho^{-1} \begin{pmatrix} -(x_2 - m_2) \\ +(x_1 - m_1) \end{pmatrix}$$

auf dem Rand der Kreisscheibe $B_\varrho(\mathbf{m})$.

3. Für den oben eingeführten Kreisring $\bar{A}_{\varrho_1, \varrho_2}(\mathbf{0})$ mit Mittelpunkt $\mathbf{m} = \mathbf{0}$ gilt

$$\boldsymbol{\nu}(\mathbf{x}) = + \frac{\mathbf{x}}{|\mathbf{x}|} = \varrho_2^{-1} \begin{pmatrix} +x_1 \\ +x_2 \end{pmatrix}, \quad \boldsymbol{\tau}(\mathbf{x}) = \varrho_2^{-2} \begin{pmatrix} -x_2 \\ +x_1 \end{pmatrix}$$

auf dem Außenrand $|\mathbf{x}| = \varrho_2$ sowie

$$\boldsymbol{\nu}(\mathbf{x}) = - \frac{\mathbf{x}}{|\mathbf{x}|} = \varrho_1^{-1} \begin{pmatrix} -x_1 \\ -x_2 \end{pmatrix}, \quad \boldsymbol{\tau}(\mathbf{x}) = \varrho_1^{-1} \begin{pmatrix} +x_2 \\ -x_1 \end{pmatrix}$$

auf dem Innenrand $|\mathbf{x}| = \varrho_1$. Analoge Formeln können für $\mathbf{m} \neq \mathbf{0}$ angegeben werden.

4. Der Rand von D ist auch immer der Rand der Komplementärmenge $\tilde{D} := \mathbb{R}^2 \setminus D$, aber es gilt

$$\boldsymbol{\nu}_{\partial D}(\mathbf{x}) = -\boldsymbol{\nu}_{\partial \tilde{D}}(\mathbf{x}), \quad \boldsymbol{\tau}_{\partial D}(\mathbf{x}) = -\boldsymbol{\tau}_{\partial \tilde{D}}(\mathbf{x}),$$

da außen aus Sicht von D gerade innen aus Sicht von \tilde{D} ist und umgekehrt (siehe das Bild). Insbesondere wird aus einer positiv orientierten Parametrisierung $\mathbf{c} : [t_0, t_1] \rightarrow \mathbb{R}^2$ von ∂D durch

$$\tilde{\mathbf{c}}(t) = \mathbf{c}(t_1 - t_0 - t)$$

eine positiv orientierte Parametrisierung $\mathbf{c} : [t_0, t_1] \rightarrow \mathbb{R}^2$ von $\partial \tilde{D}$.

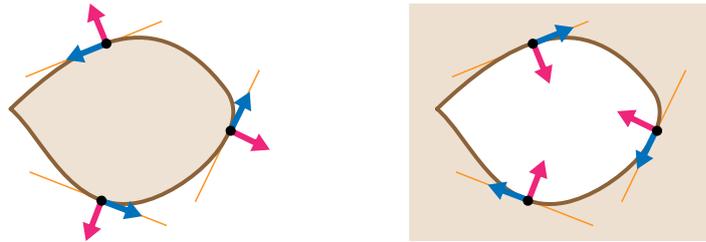


Abbildung Beim Übergang von einer Menge D (links) zur Komplementärmenge $\tilde{D} = \mathbb{R}^2 \setminus D$ (rechts) müssen Vorzeichen getauscht werden. In diesem Beispiel D ist beschränkt und \tilde{D} unbeschränkt, wobei letzteres graphisch nicht wirklich dargestellt werden kann.

Ränder als Niveaumenge Es kommt häufig vor, dass ∂D via

$$\partial D = \{(x_1, x_2) : g(x_1, x_2) = c\}$$

als Niveaumenge einer stetig differenzierbaren Funktion $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ definiert wird oder betrachtet werden kann, wobei dann D entweder die entsprechende Subniveau-Menge

$$D = \{(x_1, x_2) : g(x_1, x_2) \leq c\}$$

oder die Superniveau-Menge

$$D = \{(x_1, x_2) : g(x_1, x_2) \geq c\}$$

ist. In diesem Fall gilt

$$\nu(\mathbf{x}) = \pm \mathbf{grad} g(\mathbf{x}) = \pm \lambda(\mathbf{x}) \begin{pmatrix} \partial_{x_1} g(\mathbf{x}) \\ \partial_{x_2} g(\mathbf{x}) \end{pmatrix}$$

sowie

$$\tau(\mathbf{x}) = \pm \lambda(\mathbf{x}) \begin{pmatrix} -\partial_{x_2} g(\mathbf{x}) \\ +\partial_{x_1} g(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} -\nu_2(\mathbf{x}) \\ +\nu_1(\mathbf{x}) \end{pmatrix},$$

wobei der skalare Faktor

$$\lambda(\mathbf{x}) = \frac{1}{|\mathbf{grad} g(\mathbf{x})|} = \frac{1}{\sqrt{(\partial_{x_1} g(\mathbf{x}))^2 + (\partial_{x_2} g(\mathbf{x}))^2}}$$

wohldefiniert ist, sofern \mathbf{x} ein regulärer Punkt von g ist, d.h. falls $\mathbf{grad} g(\mathbf{x}) \neq \mathbf{0}$.

Bemerkung

1. Diese Formeln gelten, weil *Gradienten senkrecht auf Niveaukurven stehen*.
2. Bei einer Niveaukurve können wir $\boldsymbol{\tau}(\mathbf{x})$ und $\boldsymbol{\nu}(\mathbf{x})$ bis auf eine Vorzeichenwahl direkt aus den partiellen Ableitungen von g gewinnen und müssen keine Parametrisierung von ∂D explizit angeben (was im Einzelfall schwierig sein kann). Die Berechnung von Kurvenintegralen gelingt aber in aller Regel nur mit Hilfe einer Parametrisierung.
3. Die richtige und konsistente Vorzeichenwahl ergibt sich aus der Tatsache, dass der *Gradient in Richtung wachsender Funktionswerte* zeigt, und stellt sicher, dass $\boldsymbol{\nu}(\mathbf{x})$ für jedes \mathbf{x} nach außen zeigt und dass $\boldsymbol{\tau}(\mathbf{x})$ positiv orientiert ist. In konkreten Fällen hilft meist auch eine Skizze.
4. Bei *stetigen* Funktionen g sind Subniveau- und Superniveaumengen wegen der nicht-strikten Ungleichungen („ \leq “ bzw. „ \geq “) *immer abgeschlossen*, wohingegen die analogen Formeln mit strikten Ungleichungen („ $<$ “ bzw. „ $>$ “) *immer offene* Mengen liefern. Bei un stetigen Funktionen g gilt diese Aussage nicht mehr.
5. Analoge Formeln können für Zwischenniveaumengen der Bauart

$$D = \{(x_1, x_2) : c_1 \leq g(x_1, x_2) \leq c_2\}$$

abgeleitet werden, wobei der Rand ∂D dann aus zwei Niveaumengen besteht.

Beispiele

1. Für $g(x_1, x_2) = x_1^2 + x_2^2$ und $c = \varrho^2$ ist die entsprechende Subniveaumenge gerade $\overline{B}_\varrho(0)$ und die obigen Formeln mit den Ableitungen von g liefern die schon bekannten Ausdrücke für den Tangential- und den Normalenvektor, sofern wir \pm immer durch $+$ ersetzen.
2. Die Nullstellenmenge der Funktion

$$g(x_1, x_2) = (x_1^2 + x_2^2 - 3x_1)^2 - x_1^2 - x_2^2$$

wird als *Pascalsche Schnecke* bezeichnet und direkte Rechnungen sowie eine Auswertung der obigen Formeln liefern

$$\begin{pmatrix} \nu_1(\mathbf{x}) \\ \nu_2(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} +\tau_2(\mathbf{x}) \\ -\tau_1(\mathbf{x}) \end{pmatrix} = \pm \lambda(\mathbf{x}) \begin{pmatrix} 4x_1^3 - 18x_1^2 + 4x_1x_2^2 + 16x_1 - 6x_2^2 \\ 4x_1^2x_2 - 12x_1x_2 + 4x_2^3 - 2x_2 \end{pmatrix},$$

wobei wir die explizite, aber längliche Formel für $\lambda(\mathbf{x})$ nicht hingeschrieben haben. Wählen wir D als die entsprechende Subniveaumenge, so müssen wir in den obigen Formeln für jeden Punkt $\mathbf{x} \in \partial D$ den +-Ast wählen, da dann $\boldsymbol{\nu}$ in Richtung wachsender Werte von g und damit nach außen zeigt. Man kann die Pascalsche Schnecke übrigens durch

$$\mathbf{c}(t) = (1 + 3 \cos(t)) \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}, \quad t \in [0, 2\pi]$$

parametrisieren, aber das folgt nicht so einfach aus der Formel für g . Wir können aber (mit Additionstheoremen) zeigen, dass $g(\mathbf{c}(t)) = 0$ für alle Zeiten t gilt bzw. dass die Gleichung $g(\mathbf{x}) = 0$ in Polarkoordinaten als $r^2 = (1 + 3 \cos(\theta))^2$ geschrieben werden kann, wobei r der Radius und θ der Winkel ist und $t = \theta$ für die angegebene Parametrisierung gilt.

3. Die *Bernoullische Lemniskate* ist der Rand der Subniveaumenge

$$D = \{(x_1, x_2) : (x_1^2 + x_2^2)^2 - 2x_1^2 + 2x_2^2 \leq 0\}$$

und mit der richtigen Vorzeichenwahl ergibt sich

$$\begin{pmatrix} \nu_1(\mathbf{x}) \\ \nu_2(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} +\tau_2(\mathbf{x}) \\ -\tau_1(\mathbf{x}) \end{pmatrix} = 4\lambda(\mathbf{x}) \begin{pmatrix} x_1^3 + x_1 x_2^2 - x_1 \\ x_1^2 x_2 + x_2^3 + x_2 \end{pmatrix}$$

für jeden Punkt $\mathbf{x} \in \partial D$, wobei wir den Faktor $\lambda(\mathbf{x})$ wieder nicht angegeben haben. Auch für diese Kurve gibt es bekannte Parametrisierungen, zum Beispiel

$$\mathbf{c}(t) = \frac{\sqrt{2}}{1 + \sin^2(t)} \begin{pmatrix} \cos(t) \\ \cos(t) \sin(t) \end{pmatrix}, \quad t \in [0, 2\pi],$$

die zwar nicht leicht zu finden, aber einfach zu überprüfen ist. In Polarkoordinaten lautet die entsprechende Gleichung übrigens $r^2 = 2 \cos(2\theta)$.

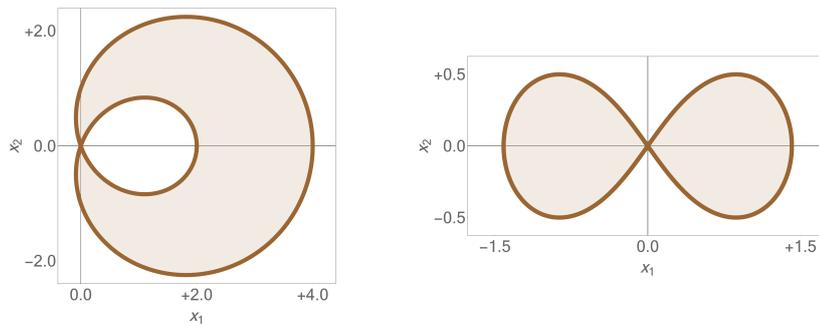


Abbildung Die Pascalsche Schnecke (links) und die Benoullische Lemniskate (rechts) sowie die entsprechenden Subniveaumengen.

Vorlesung 08-3 : 21. Dezember

Formulierung und Herleitung

Theorem (Sätze von Gauß und Stokes in 2D) Mit den Notationen und Voraussetzungen von oben gilt die Gauß-Formel

$$\int_D \operatorname{div} \mathbf{f}(\mathbf{x}) \, d\mathbf{x} = \sum_{k=1}^K \int_{\mathbf{c}_k} \langle \mathbf{f}(\mathbf{x}), \boldsymbol{\nu}(\mathbf{x}) \rangle \, ds$$

sowie die Stokes-Formel

$$\int_D \operatorname{rot} \mathbf{f}(\mathbf{x}) \, d\mathbf{x} = \sum_{k=1}^K \int_{\mathbf{c}_k} \langle \mathbf{f}(\mathbf{x}), \boldsymbol{\tau}(\mathbf{x}) \rangle \, ds$$

für jedes stetig differenzierbare Vektorfeld $\mathbf{f} : D \rightarrow \mathbb{R}^2$. Dabei steht auf der linken Seite jeweils ein Gebietsintegral und rechts immer Kurvenintegrale der ersten Art.

Beweisidee Es gibt einen wichtigen, und erstaunlich einfachen, ersten Teil sowie einen sehr technischen, und für uns nicht ganz so wichtigen, zweiten Teil.

Teil 1: Ist D die Vereinigung endlich vieler Quader, so folgt die Behauptung mit einem eleganten Auslöschungsargument direkt und ohne große Rechnungen aus dem bereits bekannten Resultat für Einzelquader. Das wird gleich im Detail erklärt.

Teil 2: Im Allgemeinen approximiert man D durch viele kleine Quader (so wie bei den Riemann-Summen oder bei der Herleitung der Transformationsformel beschrieben) und zeigt, dass der Approximationsfehler durch die Feinheit der Quaderapproximation kontrolliert wird. Die Kombination beider Argumente liefert dann die Behauptung. \square

Bemerkungen

1. Man schreibt die rechten Seiten oftmals als

$$\int_{\partial D} \langle \mathbf{f}(\mathbf{x}), \boldsymbol{\nu}(\mathbf{x}) \rangle ds \quad \text{bzw.} \quad \int_{\partial D} \langle \mathbf{f}(\mathbf{x}), \boldsymbol{\tau}(\mathbf{x}) \rangle ds,$$

da Kurvenintegrale invariant unter Reparametrisierung sind und daher das Ergebnis nur von ∂D und \mathbf{f} , aber nicht von den Details der Parametrisierung abhängt. Ohne Parametrisierung kann man die Integrale aber in der Regel nicht berechnen.

2. Aus dem Integralsatz von Gauß kann der von Stokes hergeleitet werden und umgekehrt. Wir hatten das in den Hausaufgaben schon für Quader gesehen, aber das Argument mit den um 90° gedrehten Vektorfeldern kann in 2D immer angewendet werden.
3. Die geometrisch-physikalische Interpretation der Sätze von Gauß und Stokes auf allgemeinen Mengen D ist dieselbe wie für Quader (siehe oben). Die 2D-Variante des Satzes von Stokes wird auch wieder Satz von Green genannt und oftmals anders geschrieben (siehe nochmal oben).
4. Wir haben das Stokessche Kurvenintegral immer als Kurvenintegral der 1. Art geschrieben, wobei explizit der positive orientierte Tangentialvektor $\boldsymbol{\tau}$ auftaucht. In der Literatur wird manchmal ein äquivalentes Kurvenintegral der 2. Art angegeben. Wir werden dies weiter unten genauer diskutieren.

Vereinigung endlich vieler Quader Um die wesentlichen Argumenten in Beweis der Sätze von Gauß und Stokes herauszuarbeiten, betrachten wir zunächst den Fall, dass D selbst ein Quader ist, der aus zwei kleinen Quadern $D^{(1)}$ und $D^{(2)}$ besteht, die sich in einer vertikalen Kante berühren. Ein typisches Beispiel ist

$$D = [0, 2] \times [0, 1], \quad D^{(1)} = [0, 1] \times [0, 1], \quad D^{(2)} = [1, 2] \times [0, 1].$$

Wir wissen zwar schon, dass der Satz von Gauß auf D erfüllt ist, wollen ihn hier aber noch einmal anders ableiten. Der Ausgangspunkt ist die Formel

$$\int_D \operatorname{div} \mathbf{f}(\mathbf{x}) \, d\mathbf{x} = \sum_{j=1}^2 \int_{D^{(j)}} \operatorname{div} \mathbf{f}(\mathbf{x}) \, d\mathbf{x} = \sum_{j=1}^2 \sum_{p \in \{\text{u,r,o,l}\}} \int_{\mathbf{c}_p^{(j)}} \langle \mathbf{f}(\mathbf{x}), \boldsymbol{\nu}_p^{(j)}(\mathbf{x}) \rangle ds,$$

wobei wir die Gebietsadditivität von Integralen sowie den Satz von Gauß auf beiden Teilquadern benutzt haben (und der Index p immer als „unten“, „rechts“, „oben“, „links“ zu lesen ist). Insgesamt haben wir das Gebietsintegral von $\operatorname{div} \mathbf{f}$ über D als Summe von 8 Kurvenintegralen dargestellt.

Die wesentliche Beobachtung ist nun, dass der rechte Rand vom Quader $D^{(1)}$ gerade der linke Rand vom Quader $D^{(2)}$ ist und dass

$$\boldsymbol{\nu}_r^{(1)}(\mathbf{x}) = -\boldsymbol{\nu}_l^{(2)}(\mathbf{x})$$

in jedem Punkt $\mathbf{x} \in \partial D^{(1)} \cap \partial D^{(2)}$ erfüllt ist, der zum Rand von $D^{(1)}$ und zum Rand von $D^{(2)}$ gehört. Oder anders gesagt: Auf der gemeinsamen Randkante zeigen die äußeren Normalenvektoren der beiden Teilquader in entgegengesetzte Richtungen. Hieraus folgt

$$\int_{\mathbf{c}_r^{(1)}} \langle \mathbf{f}(\mathbf{x}), \boldsymbol{\nu}_r^{(1)}(\mathbf{x}) \rangle ds = - \int_{\mathbf{c}_l^{(2)}} \langle \mathbf{f}(\mathbf{x}), \boldsymbol{\nu}_l^{(2)}(\mathbf{x}) \rangle ds,$$

d.h. von den insgesamt 8 Kurvenintegralen heben sich die zwei gegenseitig auf, die zur gemeinsamen Kante von $D^{(1)}$ und $D^{(2)}$ gehören. Die restlichen 6 Kurvenintegrale bilden aber zusammen das Gaußsche Kurvenintegral auf dem Gesamtgebiet D , wobei auf diesen 6 Randstücken der jeweilige Normalenvektor auf ∂D mit den entsprechenden Vektoren auf $\partial D^{(1)}$ bzw. $\partial D^{(2)}$ übereinstimmt. Es gilt also

$$\sum_{j=1}^2 \sum_{p \in \{\text{u,r,o,l}\}} \int_{\mathbf{c}_p^{(j)}} \langle \mathbf{f}(\mathbf{x}), \boldsymbol{\nu}_p^{(j)}(\mathbf{x}) \rangle ds = \int_{\partial D} \langle \mathbf{f}(\mathbf{x}), \boldsymbol{\nu}(\mathbf{x}) \rangle ds$$

und wir haben damit den Satz von Gauß auf D aus dem Satz von Gauß auf den Teilquadern $D^{(1)}$ und $D^{(2)}$ hergeleitet.

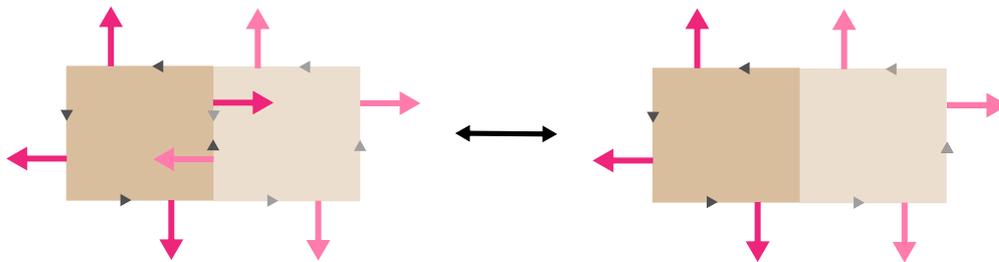


Abbildung Wird der Satz von Gauß auf zwei sich in einer Kante berührenden Quader (hell- und dunkelbraun) angewandt, so löschen sich die Einzelbeiträge zum Kurvenintegral auf der gemeinsamen Kante gegenseitig aus.

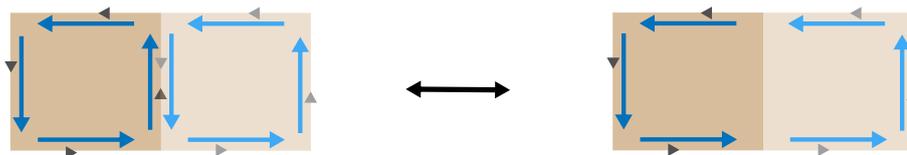


Abbildung Auch beim Stokesschen Kurvenintegral gibt es Auslöschungseffekte auf gemeinsamen Quaderkanten.

Beim Satz von Stokes kann man analog argumentieren, da auch

$$\boldsymbol{\tau}_r^{(1)}(\mathbf{x}) = -\boldsymbol{\tau}_l^{(2)}(\mathbf{x})$$

auf der gemeinsamen Quaderkante von $D^{(1)}$ und $D^{(2)}$ gilt. Alle Schritte (Aufspaltung in Teilquader, Auswertung der Integralsätze in jedem Teilquader, Auslöschung entlang von gemeinsamen Kanten, Zusammensetzen von Kurvenintegralen) können auch auf die Vereinigung von endlich vielen Quadern angewendet werden, und liefern dann einen Beweis der Sätze von Gauß und Stokes.

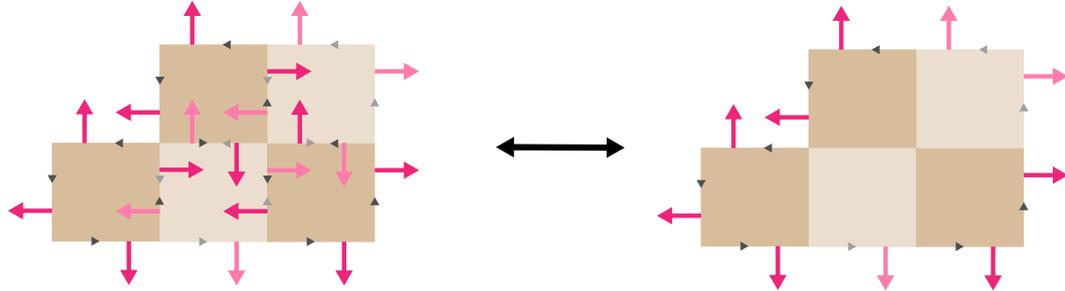


Abbildung Auslöschungseffekte implizieren die Gültigkeit des Satzes von Gauß auf beliebigen Vereinigungen von zweidimensionalen Quadern, sofern diese sich nur in gemeinsamen Randkanten überlappen. Allgemeinere Gebiete können durch Approximationsargumente behandelt werden.

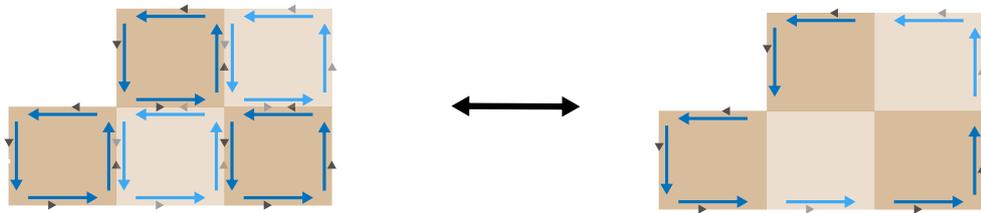


Abbildung Der Satz von Stokes kann analog bewiesen werden.

Bemerkung* Wir können das Stokessche Kurvenintegral via

$$\int_{\mathbf{c}} \langle \mathbf{f}(\mathbf{x}), \boldsymbol{\tau}(\mathbf{x}) \rangle ds = \int_{t_0}^{t_1} \langle \mathbf{f}(\mathbf{c}(t)), \dot{\mathbf{c}}(t) \rangle dt = \int_{\mathbf{c}} \mathbf{f}(\mathbf{x}) \cdot d\mathbf{x}$$

auch als Kurvenintegral 2. Art schreiben, sofern (1.) die Menge ∂D durch eine einzige Randkurve $\mathbf{c} : [t_0, t_1] \rightarrow \mathbb{R}^2$ parametrisiert wird und (2.) die Vektoren $\dot{\mathbf{c}}(t)$ und $\boldsymbol{\tau}(\mathbf{c}(t))$ für jedes $t \in [t_0, t_1]$ in die gleiche (und nicht in die entgegengesetzte) Richtung zeigen. Unter diesen Voraussetzungen gilt nämlich

$$\dot{\mathbf{c}}(t) = |\dot{\mathbf{c}}(t)| \boldsymbol{\tau}(\mathbf{c}(t)) \quad \text{und damit} \quad \dot{\mathbf{c}}(t) dt = \boldsymbol{\tau}(\mathbf{c}(t)) ds.$$

Bei der Schreibweise als Kurvenintegral 1. Art gibt es diese Restriktionen nicht, da dort $\boldsymbol{\tau}(\mathbf{x})$ explizit auftaucht.

Flächenberechnung durch Randintegrale In vielen Fällen kann man den Rand von D durch eine einzelne geschlossene Kurve $\mathbf{c} : [t_0, t_1] \rightarrow \mathbb{R}^2$ ohne Doppelpunkte parametrisieren. Dann gilt

$$\boldsymbol{\tau}(\mathbf{c}(t)) = \pm \frac{\dot{\mathbf{c}}(t)}{|\dot{\mathbf{c}}(t)|},$$

wobei das Vorzeichen für alle $t \in [t_0, t_1]$ gleich ist und davon abhängt, ob der Umlaufsinn der Parametrisierung mit $\boldsymbol{\tau}$ übereinstimmt oder nicht. Wir können nun

den Satz von Stokes für sehr spezielle Vektorfelder mit konstanter Rotation auswerten. Zum Beispiel gilt

$$\begin{aligned} \text{vol}(D) &= \int_D \text{rot} \begin{pmatrix} 0 \\ +x_1 \end{pmatrix} d\mathbf{x} = \int_{\mathbf{c}} \left\langle \begin{pmatrix} 0 \\ +x_1 \end{pmatrix}, \boldsymbol{\tau}(\mathbf{x}) \right\rangle ds = + \int_{\mathbf{c}} x_1 \tau_2(\mathbf{x}) ds \\ &= \pm \int_{t_0}^{t_1} c_1(t) \frac{\dot{c}_2(t)}{|\dot{\mathbf{c}}(t)|} |\dot{\mathbf{c}}(t)| dt = \pm \int_{t_0}^{t_1} c_1(t) \dot{c}_2(t) dt, \end{aligned}$$

wobei wir $ds = |\dot{\mathbf{c}}(t)| dt$ verwendet haben, und analog ergibt sich

$$\text{vol}(D) = \int_D \text{rot} \begin{pmatrix} -x_2 \\ 0 \end{pmatrix} d\mathbf{x} = - \int_{\mathbf{c}} x_2 \tau_1(\mathbf{x}) ds = \mp \int_{t_0}^{t_1} c_2(t) \dot{c}_1(t) dt.$$

Bei Kurven mit Doppelpunkt können diese Formeln nicht verwendet werden, da dann das Vorzeichen des Proportionalitätsfaktors zwischen $\boldsymbol{\tau}(\mathbf{c}(t))$ und $\dot{\mathbf{c}}(t)$ von t abhängt und es ungewollte Auslöschungen zwischen positiven und negativen Beiträgen geben kann.

Beispiel Der Flächeninhalt der Kreisscheibe von Radius ϱ kann auch mit der Parametrisierung

$$\mathbf{c}(t) = \varrho \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}, \quad t \in [0, 2\pi]$$

und durch das Kurvenintegral

$$\text{vol}(\overline{B}_\varrho(\mathbf{0})) = \int_0^{2\pi} c_1(t) \dot{c}_2(t) dt = \varrho^2 \int_0^{2\pi} \cos^2(t) dt = \pi \varrho^2$$

berechnet werden.

rotationssymmetrische Vektorfelder Für eine gegebene stetig differenzierbare Funktion $\phi : (0, \infty) \rightarrow \mathbb{R}$ werden durch

$$(RZ) \quad \mathbf{f}(\mathbf{x}) = \phi\left(\sqrt{x_1^2 + x_2^2}\right) \begin{pmatrix} +x_1 \\ +x_2 \end{pmatrix}, \quad (RD) \quad \mathbf{f}(\mathbf{x}) = \phi\left(\sqrt{x_1^2 + x_2^2}\right) \begin{pmatrix} -x_2 \\ +x_1 \end{pmatrix},$$

zwei rotationssymmetrische Vektorfelder auf $\mathbb{R}^2 \setminus \{\mathbf{0}\}$ definiert. Das Feld RZ ist dabei ein sogenanntes *Zentralfeld*, wohingegen beim *Drehfeld* RD sich die Vektoren entlang von Kreisen um den Koordinatenursprung drehen. Für jeden Kreisring mit Mittelpunkt $\mathbf{x} = \mathbf{0}$ können wir sowohl die Gebiets- als auch die Kurvenintegrale in den Integralsätzen von Gauß und Stokes explizit berechnen und damit die Bedeutung der zweidimensionalen Differentialoperatoren div und rot sowie der entsprechenden Gebiets- und Kurvenintegrale an einfachen Beispielen studieren (siehe die Übungen).

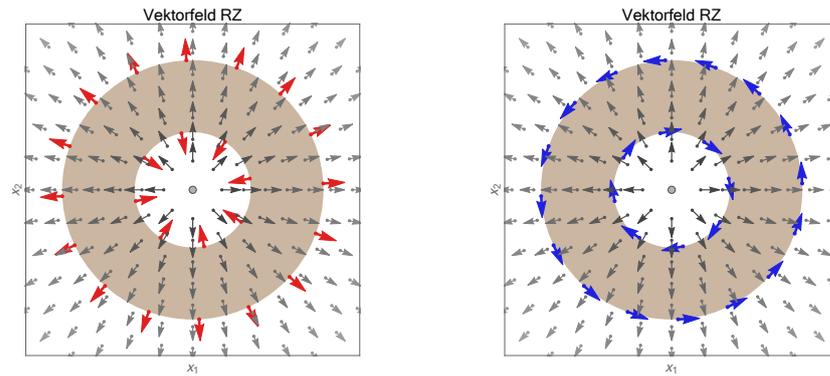


Abbildung Das Vektorfeld RZ (grau) ist zwar immer rotationsfrei, aber der Wert von $\operatorname{div} \mathbf{f}(\mathbf{x})$ unterliegt keiner solchen universellen Einschränkung (selbst dann nicht, wenn ϕ positiv und monoton fallend ist). Insbesondere kann dieses Vektorfeld nur für $\phi(r) = cr^{-2}$ (hier *nicht* dargestellt) als Geschwindigkeitsfeld einer idealen Flüssigkeit interpretiert werden, die aus einer Quelle im Ursprung sprudelt und in einer Ebene ins Unendliche abfließt. Die roten bzw. blauen Pfeile stellen den äußeren Normalenvektor bzw. den positiv orientierten Tangentialvektor in ausgewählten Randpunkten des braunen Kreisringes dar.

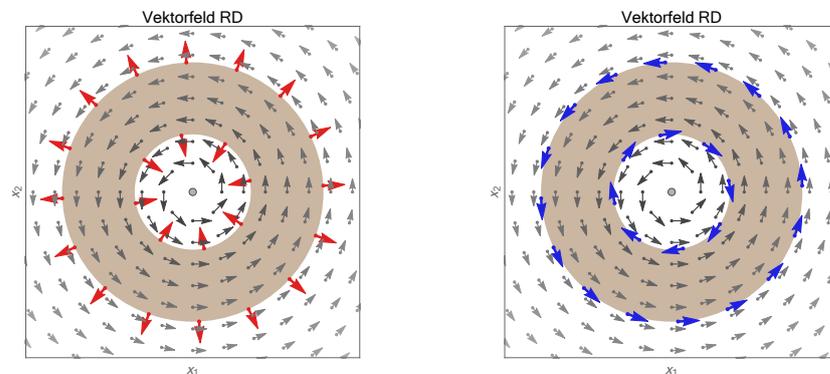


Abbildung Links: Das Vektorfeld RD ist immer divergenzfrei, aber diesmal kann $\operatorname{rot} \mathbf{f}(\mathbf{x})$ im Prinzip jeden Wert annehmen. Dies gilt selbst dann, ϕ positiv und monoton fallend ist. In diesem Fall dreht sich das Vektorfeld zwar entgegen dem Uhrzeigersinn um den Ursprung, aber $\operatorname{rot} \mathbf{f}(\mathbf{x})$ misst eben nicht diese *globale Rotation*, sondern quantifiziert die *lokale Rotation*, d.h. die *Verwirbelungseffekte* des Vektorfeldes in der Nähe von \mathbf{x} . Stellen wir uns eine *kleine Kugel* mit Radius $0 < \varepsilon \ll 1$ und Mittelpunkt \mathbf{x} vor, so beschreibt $\operatorname{rot} \mathbf{f}(\mathbf{x})$ wie stark und in welcher Orientierung das Vektorfeld diese kleine Kugel um ihre eigene Achse drehen möchte (*Spin*). Vollkommen unabhängig davon will das Vektorfeld den Mittelpunkt der Kugel auch noch verschieben, nämlich hier entlang einer Kreisbahn um den Ursprung.

Bemerkung Ein rotationssymmetrisches Vektorfeld kann in zwei Dimensionen als

$$\mathbf{f}(\mathbf{x}) = \phi_1 \left(\sqrt{x_1^2 + x_2^2} \right) \begin{pmatrix} +x_1 \\ +x_2 \end{pmatrix} + \phi_2 \left(\sqrt{x_1^2 + x_2^2} \right) \begin{pmatrix} -x_2 \\ +x_1 \end{pmatrix},$$

geschrieben werden, wobei ϕ_1 und ϕ_2 nur vom Polarradius abhängen. Der erste bzw. zweite Summand ist dabei rotationsfrei bzw. divergenzfrei.

Zwei wichtige singuläre Vektorfelder Die Vektorfelder

$$\mathbf{f}_Q(\mathbf{x}) = \frac{c}{x_1^2 + x_2^2} \begin{pmatrix} +x_1 \\ +x_2 \end{pmatrix} \quad \text{bzw.} \quad \mathbf{f}_W(\mathbf{x}) = \frac{c}{x_1^2 + x_2^2} \begin{pmatrix} -x_2 \\ +x_1 \end{pmatrix}$$

werden elementares Quellenfeld bzw. elementares Wirbelfeld genannt (jeweils in 2D), wobei $c \neq 0$ eine beliebige reelle Konstante ist. Sie sind auf ganz $\mathbb{R}^2 \setminus \{\mathbf{0}\}$ sowohl divergenzfrei als auch rotationsfrei (siehe die Hausaufgaben), aber besitzen jeweils eine *Singulärität* im Koordinatenursprung $\mathbf{x} = \mathbf{0}$.

Bemerkung* Für eine Kugel $\overline{B}_\rho(\mathbf{0})$ können wir aufgrund der Singularitäten die Gebietsintegrale in den Sätzen von Gauß und Stokes nicht auswerten. Wir können aber problemlos die entsprechenden Kurvenintegrale bestimmen und erhalten

$$\int_{\partial\overline{B}_\rho(\mathbf{0})} \langle \mathbf{f}_Q(\mathbf{x}), \boldsymbol{\nu}(\mathbf{x}) \rangle ds = 2\pi c, \quad \int_{\partial\overline{B}_\rho(\mathbf{0})} \langle \mathbf{f}_Q(\mathbf{x}), \boldsymbol{\tau}(\mathbf{x}) \rangle ds = 0$$

bzw.

$$\int_{\partial\overline{B}_\rho(\mathbf{0})} \langle \mathbf{f}_W(\mathbf{x}), \boldsymbol{\nu}(\mathbf{x}) \rangle ds = 0, \quad \int_{\partial\overline{B}_\rho(\mathbf{0})} \langle \mathbf{f}_W(\mathbf{x}), \boldsymbol{\tau}(\mathbf{x}) \rangle ds = 2\pi c,$$

wobei das Ergebnis *unabhängig vom Kugelradius* ρ ist. Beachte, dass bei jedem Vektorfeld, das stetig differenzierbar auf ganz \mathbb{R}^2 ist, hingegen

$$\int_{\partial\overline{B}_\rho(\mathbf{0})} \langle \mathbf{f}(\mathbf{x}), \boldsymbol{\nu}(\mathbf{x}) \rangle ds = \pi \rho^2 \operatorname{div} \mathbf{f}(\mathbf{0}) + O(\rho^3)$$

sowie

$$\int_{\partial\overline{B}_\rho(\mathbf{0})} \langle \mathbf{f}(\mathbf{x}), \boldsymbol{\tau}(\mathbf{x}) \rangle ds = \pi \rho^2 \operatorname{rot} \mathbf{f}(\mathbf{0}) + O(\rho^3)$$

gelten wird, wobei diese Formeln mittels einer Taylor-Approximation der Komponenten von \mathbf{f} aus den Integralsätzen abgeleitet werden können. Diese Beispiele illustrieren eine sehr wichtige Strategie in der Mathematik und den Anwendungswissenschaften: Wir können nämlich \mathbf{f}_Q eine *distributionelle Divergenz* und \mathbf{f}_W eine *distributionelle Rotation* zuweisen, wobei der Wert jeweils $2\pi c$ ist. Diese sind im Ursprung konzentriert (in etwas so, wie Masse in einem *Punkt* konzentriert sein kann) und können nicht mehr durch Differentiation, sondern nur durch ein *duales* Konzept (im konkreten Fall durch ein Kurvenintegral) berechnet werden. Wir können diese beiden Vektorfelder sogar als „*Elementarbausteine*“ betrachten, aus denen jedes andere Vektorfeld durch eine „unendliche Überlagerung“ zusammengesetzt werden kann (analog sind Punktmassen die „Bausteine“ in der Kontinuumsmechanik).

Ausblick: Man kann sehr gut mit Singularitäten rechnen, aber im Moment übersteigt das noch unseren Erkenntnisstand.

Kapitel 3

Vektoranalysis

Vorlesung 09-1 : 11. Januar

3.1 mehr über Kurvenintegrale

Ziel In diesem Abschnitt wiederholen wir einige Konzepte über parametrisierte Kurven im \mathbb{R}^n sowie die entsprechenden Integrale. Wir können eine Kurve oftmals als den *Weg* interpretieren, den ein (reales oder gedachtes) Teilchen im \mathbb{R}^n zurücklegt.

Erinnerung

1. Eine parametrisierte Kurve (bzw. ein Weg) in einer gegebenen Menge $D \subset \mathbb{R}^n$ ist eine stetige Abbildung $\mathbf{c} : I \rightarrow D$, wobei I ein Intervall ist. Die Elemente von I bezeichnen wir mit t und werden diese meist als *Zeit* interpretieren.
2. Das Bild von \mathbf{c} , also die Menge

$$\text{im}(\mathbf{c}) = \{\mathbf{c}(t) : t \in I\} \subset D$$

ist das geometrische Objekt (Kreislinie, Spirale usw.), das von \mathbf{c} parametrisiert wird.

3. Ist I kompakt, so schreiben wir $I = [t_a, t_e]$ und nennen $\mathbf{c}(t_a)$ den Anfangspunkt und $\mathbf{c}(t_e)$ den Endpunkt. Man sagt auch, die Kurve \mathbf{c} verbindet diese beiden Punkte.
4. Die Kurve \mathbf{c} heißt geschlossen, falls $\mathbf{c}(t_a) = \mathbf{c}(t_e)$ gilt, d.h. falls Anfangs- und Endpunkt zusammenfallen.
5. Gilt $\mathbf{c}(t_1) = \mathbf{c}(t_2)$ für zwei verschiedene Zeiten t_1, t_2 mit $t_a \leq t_1 < t_2 \leq t_e$, so sprechen wir von einem Doppelpunkt der Kurve, es sein denn, es gilt $t_1 = t_a$ und $t_2 = t_e$. Diese Ausnahme ist sinnvoll, da andernfalls jede geschlossene Kurve einen Doppelpunkt besäße.
6. Ist \mathbf{c} stetig differenzierbar in t , so heißt

$$\dot{\mathbf{c}}(t) = \frac{d\mathbf{c}}{dt}(t)$$

der (momentane) Tangentialvektor oder Geschwindigkeitsvektor zur Zeit t . Die zweite Ableitung

$$\ddot{\mathbf{c}}(t) = \frac{d\dot{\mathbf{c}}}{dt}(t)$$

wird, falls sie existiert, der (momentane) Beschleunigungsvektor genannt.

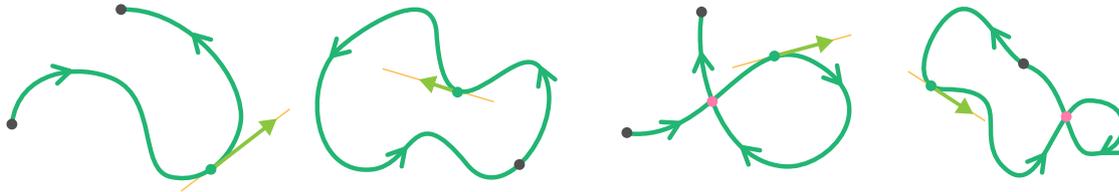


Abbildung Vier Beispiele für Kurven (bzw. Wege) im \mathbb{R}^2 , wobei die schwarzen Kreise die jeweiligen Anfangs- und Endpunkte darstellen und die Durchlaufrichtung durch stilisierte Pfeile angedeutet wird. In einem ausgewählten Punkt jeder Kurve sind außerdem die Tangentialgerade und der Tangentialvektor in gelb und hellgrün gezeichnet. Die rosa Kreise markieren Doppelpunkte.

Beispiel

1. Die parametrisierte Kurve

$$\mathbf{c}(t) = (1-t)\mathbf{p} + t\mathbf{q} = (1-t) \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} + t \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix}, \quad t \in [0, 1]$$

beschreibt ein dreidimensionales Geradenstück, wobei die reellen Zahlen p_j bzw. q_j die Koordinaten des Anfangspunktes \mathbf{p} bzw. des Endpunkts \mathbf{q} sind.

2. Durch

$$\mathbf{c}(t) = \begin{pmatrix} \varrho \cos(\omega t) \\ \varrho \sin(\omega t) \\ \sigma t \end{pmatrix}, \quad t \in \mathbb{R}$$

wird eine unendliche Spirale parametrisiert, wobei ϱ , ω und σ reelle Parameter sind und die Mittellinie der Spirale gerade x_3 -Achse ist.

Voraussetzung Wir werden im Folgenden immer annehmen, dass \mathbf{c} eine stückweise stetig differenzierbare Kurve ist, wobei wir akzeptieren, dass der Tangentialvektor $\dot{\mathbf{c}}$ zu endlich vielen Zeiten (die *Knicken* oder *Ecken* entsprechen) doppelt, d.h. als $\dot{\mathbf{c}}(t-0)$ und $\dot{\mathbf{c}}(t+0)$ im Sinne einseitiger Ableitungen zu berechnen ist. Wir setzen außerdem voraus, dass die Kurve regulär ist, d.h. dass

$$\dot{\mathbf{c}}(t) \neq \mathbf{0}$$

für alle $t \in I$ gilt. Diese Bedingung schließt einige Entartungsfälle aus.

Erinnerung

1. Für eine skalare Funktion $f : D \rightarrow \mathbb{R}$ wird durch

$$\int_{\mathbf{c}} f(\mathbf{x}) \, ds := \int_{t_a}^{t_e} f(\mathbf{c}(t)) |\dot{\mathbf{c}}(t)| \, dt$$

das Kurvenintegral der 1. Art definiert, wobei die formale Substitution

$$\mathbf{x} = \mathbf{c}(t), \quad ds = |\dot{\mathbf{c}}(t)| \, dt$$

zu Grunde liegt. Im Spezialfall $f \equiv 1$ erhalten wir mit

$$L(\mathbf{c}) := \int_{\mathbf{c}} 1 \, ds = \int_{t_a}^{t_e} |\dot{\mathbf{c}}(t)| \, dt$$

die Integralformel für die Länge einer Kurve.

2. Das Kurvenintegral der 2. Art bezieht sich auf ein Vektorfeld $\mathbf{f} : D \rightarrow \mathbb{R}^n$ und wird durch

$$\int_{\mathbf{c}} \mathbf{f}(\mathbf{x}) \cdot d\mathbf{x} := \int_{t_a}^{t_e} \langle \mathbf{f}(\mathbf{c}(t)), \dot{\mathbf{c}}(t) \rangle \, dt$$

eingeführt, wobei das Skalarprodukt im \mathbb{R}^n auf der linken Seite durch \cdot angedeutet und auf der rechten Seite als $\langle \cdot, \cdot \rangle$ geschrieben ist. Formal gilt diesmal

$$\mathbf{x} = \mathbf{c}(t), \quad d\mathbf{x} = \dot{\mathbf{c}}(t) \, dt$$

und der Integrand auf der rechten Seite kann alternativ als $\mathbf{f}(\mathbf{c}(t)) \cdot \dot{\mathbf{c}}(t)$ geschrieben werden.

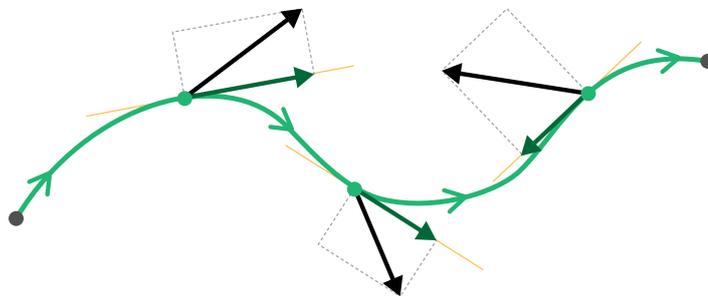


Abbildung Eine Kurve (grün) mit Anfangs- und Endpunkt (grau). Ein Vektorfeld (schwarz) besitzt in jedem Punkt einen entsprechenden tangentialen Anteil (dunkelgrün), der entweder in dieselbe oder die entgegengesetzte Richtung wie der Tangentialvektor an die Kurve (nicht dargestellt) zeigt. Das Kurvenintegral der 2. Art integriert diese tangentialen Komponente, wobei das Ergebnis invariant unter Reparametrisierungen der Kurve ist, aber sein Vorzeichen beim Wechsel der Durchlaufrichtung ändert. Beachte, dass es in 2D in jedem Punkt nur eine Normalenrichtung an die Kurve gibt, wohingegen es in nD für jeden Punkt der Kurve einen $n - 1$ -dimensionalen Unterraum des \mathbb{R}^n gibt, der senkrecht auf dem momentanen Tangentialvektor steht.

Bemerkung

1. Es gilt

$$\dot{\mathbf{c}}(t) = |\dot{\mathbf{c}}(t)| \frac{\dot{\mathbf{c}}(t)}{|\dot{\mathbf{c}}(t)|}$$

und damit symbolisch auch

$$\langle \mathbf{f}(\mathbf{c}(t)), \dot{\mathbf{c}}(t) \rangle dt = \left\langle \mathbf{f}(\mathbf{c}(t)), \frac{\dot{\mathbf{c}}(t)}{|\dot{\mathbf{c}}(t)|} \right\rangle ds,$$

d.h. jedes Kurvenintegral der 2. Art ist eigentlich ein spezielles Kurvenintegral der 1. Art. Beachte auch, dass $\dot{\mathbf{c}}(t)/|\dot{\mathbf{c}}(t)|$ gerade der *normierte Tangentialvektor* und damit der erste Frenet-Vektor der Kurve ist.

2. Bei Kurvenintegralen der 2. Art ist $d\mathbf{x}$ als infinitesimaler Vektor zu interpretieren, wohingegen bei Gebietsintegralen $d\mathbf{x} = dx_1 \cdot \dots \cdot dx_n$ als infinitesimaler Skalar anzusehen ist.

Interpretation 1* Vektorfelder \mathbf{f} können oftmals als Kraftfelder interpretiert werden, wobei $\mathbf{f}(\mathbf{x})$ die Kraft beschreibt, die im Punkt \mathbf{x} auf ein (reales oder gedachtes) Teilchen wirkt. In diesem Kontext modelliert eine parametrisierte Kurve \mathbf{c} die Bahn oder den Weg eines Teilchens und das Kurvenintegral 2. Art beschreibt gerade die vom Kraftfeld entlang der Teilchenbahn verrichtete Arbeit.

Beispiele*

1. In der Mechanik entspricht $\mathbf{f}(\mathbf{x})$ zum Beispiel einem Gravitationsfeld (dessen Details von der betrachteten Masseverteilung abhängen) und das Kurvenintegral von \mathbf{f} bzgl. \mathbf{c} ist gerade die mechanische Arbeit die man leisten muss, um ein gedachtes Probeteilchen mit Einheitsmasse entlang der Kurve \mathbf{c} zu verschieben. Da Gravitationsfelder in der Regel Gradientenfelder sind (siehe unten), kann die verrichtete Arbeit mit der Differenz der potentiellen Energie zwischen End- und Anfangspunkt in Beziehung gebracht werden.
2. Analog können wir mit Hilfe eines Vektorfeldes \mathbf{f} die elektrischen Coulomb-Kräfte beschreiben, die von einer statischen Ladungsverteilung im Raum erzeugt werden und auf ein gedachtes Probeteilchen mit Einheitsladung wirken. In der Elektrodynamik sind die Kraftfelder aber dynamisch, d.h. \mathbf{f} wird nicht nur von der Ortsvariablen \mathbf{x} , sondern auch noch von der Zeit t abhängen. Beachte auch, das magnetische Lorentz-Kräfte anders sind, da sie nicht nur von der Position, sondern auch von der Geschwindigkeit des Teilchens abhängen.

Bemerkung* Die Mathematik ignoriert in der Regel alle physikalischen Einheiten, wobei diese Strategie viele Vorzüge, aber auch einige Nachteile mit sich bringt. Wir wollen zur Vollständigkeit und für den Standardfall die physikalischen Dimensionen

aller Größen sowie die entsprechenden SI-Einheiten auflisten:

t, dt	:	Zeit	bzw.	Sekunde
c_j	:	Länge	bzw.	Meter
\dot{c}_j	:	Geschwindigkeit = $\frac{\text{Länge}}{\text{Zeit}}$	bzw.	$\frac{\text{Meter}}{\text{Sekunde}}$
f_j	:	Kraft = $\frac{\text{Masse} \cdot \text{Länge}}{\text{Zeit}^2}$	bzw. Newton =	$\frac{\text{Kilogramm} \cdot \text{Meter}}{\text{Sekunde}^2}$
$\langle \mathbf{f}, \dot{\mathbf{c}} \rangle$:	Leistung = $\frac{\text{Arbeit}}{\text{Zeit}}$	bzw. Watt =	$\frac{\text{Kilogramm} \cdot \text{Meter}^2}{\text{Sekunde}^3}$

Das Kurvenintegral ist eine Arbeit = Kraft · Länge mit SI-Einheit Joule = Newton · Meter.

Interpretation 2* Beschreibt \mathbf{f} das Geschwindigkeitsfeld einer Stömung (etwa von Wasser), so quantifiziert das Kurvenintegral wie viel Strömungsmittel entlang der Kurve fließt. Insbesondere können wir die Zirkulation entlang einer geschlossenen Kurve auswerten. Bei Kraftfeldern liefert die Zirkulation gerade die Arbeit, die entlang einer geschlossenen Kurve verrichtet wird.

Bemerkung: Es besteht ein innerer Zusammenhang zwischen den beiden Konzepten Zirkulation und Rotation, den wir in 2D schon kennengelernt haben: Ist \mathbf{c} eine geschlossene planare Kurve ohne Doppelpunkt, die den Rand einer Menge B im richtigen Umlaufsinn parametrisiert, so garantiert der Satz von Stokes die Formel

$$\int_{\mathbf{c}} \mathbf{f}(\mathbf{x}) \cdot d\mathbf{x} = \int_{\partial B} \langle \mathbf{f}(\mathbf{x}), \boldsymbol{\tau}(\mathbf{x}) \rangle ds = \int_B \text{rot } \mathbf{f}(\mathbf{x}) \, d\mathbf{x}.$$

Wir werden unten sehen, dass es einen ähnlicher Zusammenhang auch in drei Dimensionen gibt.

Beispiel Für das elementare Wirbelfeld \mathbf{f}_W vom Ende des letzten Kapitels sowie die natürliche Parametrisierung der Kreislinie $\partial B_\varrho(\mathbf{0})$ erhalten wir

$$\mathbf{c}(t) = \varrho \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}, \quad \dot{\mathbf{c}}(t) = \varrho \begin{pmatrix} -\sin(t) \\ +\cos(t) \end{pmatrix}, \quad \mathbf{f}_W(\mathbf{c}(t)) = \frac{c}{\varrho} \begin{pmatrix} -\sin(t) \\ +\cos(t) \end{pmatrix}$$

und damit

$$\langle \mathbf{f}_W(\mathbf{c}(t)), \dot{\mathbf{c}}(t) \rangle = c, \quad \int_{\mathbf{c}} \mathbf{f}_W(\mathbf{x}) \cdot d\mathbf{x} = \int_0^{2\pi} c \, dt = 2\pi c,$$

wobei c eine reelle Konstante ist. Oder anders gesagt: Der Wert des Parameters c gibt gerade an, wie viel Stömungsmittel entlang eines konzentrischen Kreises um den Ursprung fließt. Mit mehr Aufwand könnte man denselben Wert, also wieder $2\pi c$, für jede Kurve berechnen, die *genau einmal* und *entgegen dem Uhrzeigersinn* um den Koordinatenursprung (dort lebt die Singularität des Vektorfeldes) läuft. Wir erhalten aber $-2\pi c$, sofern der Umlaufsinn der Kurve umgekehrt wird. Für jede geschlossene Kurve, die nicht um $\mathbf{0}$ läuft, verschwindet jedoch das Kurvenintegral nach dem Stokesschen Integralssatz, denn es gilt $\text{rot } \mathbf{f}_W(\mathbf{x}) = 0$ in jedem Punkt $\mathbf{x} \neq \mathbf{0}$.

Alternative Notationen*

1. Manchmal wird auch der Punkt im Kurvenintegral weggelassen, d.h. man schreibt

$$\int_{\mathbf{c}} \mathbf{f}(\mathbf{x}) \, d\mathbf{x} \quad \text{statt} \quad \int_{\mathbf{c}} \mathbf{f}(\mathbf{x}) \cdot d\mathbf{x},$$

aber $\int_{\mathbf{c}} \langle \mathbf{f}(\mathbf{x}), d\mathbf{x} \rangle$ wird nur sehr selten benutzt. Für *geschlossene* Kurven findet sich in der Literatur häufig die Schreibweise

$$\oint_{\mathbf{c}} \mathbf{f}(\mathbf{x}) \, d\mathbf{x} \quad \text{oder} \quad \oint_{\mathbf{c}} \mathbf{f}(\mathbf{x}) \cdot d\mathbf{x}.$$

Bemerkung: Die Notationen sind wieder mal nicht einheitlich. Daher ist es um so wichtiger, dass Sie die Konzepte und Ideen verinnerlichen: Ein Kurvenintegral der 2. Art integriert zum Beispiel ein Vektorfeld entlang einer Kurve und liefert eine Zahl (und nicht etwa einen Vektor). Man erkennt ein Kurvenintegral in der Regel daran, dass das Integrationsgebiet eine eindimensionale Kurve und nicht (wie beim Gebietsintegral) eine n -dimensionale Menge ist.

2. Manchmal wird das Kurvenintegral der 2. Art auch als

$$\int_{\mathbf{c}} \mathbf{f}(\mathbf{x}) \cdot d\mathbf{x} = \sum_{j=1}^n \int_{\mathbf{c}} f_j(\mathbf{x}) \, dx_j$$

geschrieben, wobei dies im Sinne der Substitutionen $x_j = c_j(t)$ und $dx_j = \dot{c}_j(t) \, dt$ zu verstehen ist. Eine ähnliche Notation finden wir auch bei der 2D-Variante des Satzes von Stokes bzw. dem Satz von Green (siehe oben).

Bemerkung: Ein Integral $\int_{\mathbf{c}} f_j(\mathbf{x}) \, dx_j$ muss entlang der parametrisierten Kurve \mathbf{c} ausgewertet werden und stellt eigentlich eine Zwischen- oder Mischform von Integralen dar. Alternativ kann man es als das Integral einer 1-Form betrachten, aber das zu Grunde liegende Konzept der *Differentialformen* spielt in dieser Vorlesung keine Rolle.

Beispiele

1. Für ein affines Vektorfeld $\mathbf{f}(\mathbf{x}) = \mathbf{A} \cdot \mathbf{x} + \mathbf{v}$ mit $\mathbf{v} \in \mathbb{R}^n$ und symmetrischer Matrix $\mathbf{A} = \mathbf{A}^T \in \mathbb{R}^{(n,n)}$ erhalten wir

$$\int_{\mathbf{c}} \mathbf{f}(\mathbf{x}) \cdot d\mathbf{x} = \int_{t_a}^{t_b} \langle \mathbf{A} \cdot \mathbf{c}(t) + \mathbf{v}, \dot{\mathbf{c}}(t) \rangle \, dt = \left[\frac{1}{2} \langle \mathbf{A} \cdot \mathbf{c}(t), \mathbf{c}(t) \rangle + \langle \mathbf{v}, \mathbf{c}(t) \rangle \right]_{t=t_a}^{t=t_b},$$

wobei wir den Hauptsatz der Differential- und Integralrechnung bzgl. t sowie die Formeln

$$\frac{d}{dt} \langle \mathbf{v}, \mathbf{c}(t) \rangle = \langle \mathbf{v}, \dot{\mathbf{c}}(t) \rangle$$

und

$$\frac{d}{dt} \langle \mathbf{A} \cdot \mathbf{c}(t), \mathbf{c}(t) \rangle = \langle \mathbf{A} \cdot \dot{\mathbf{c}}(t), \mathbf{c}(t) \rangle + \langle \mathbf{A} \cdot \mathbf{c}(t), \dot{\mathbf{c}}(t) \rangle = \langle (\mathbf{A} + \mathbf{A}^T) \cdot \mathbf{c}(t), \dot{\mathbf{c}}(t) \rangle$$

benutzt haben, die man einfach nachrechnen kann.

2. Für Gradientenfelder $\mathbf{f}(\mathbf{x}) = \mathbf{grad} \psi(\mathbf{x})$ ergibt sich analog

$$\begin{aligned} \int_{\mathbf{c}} \mathbf{grad} \psi(\mathbf{x}) \cdot d\mathbf{x} &= \int_{t_a}^{t_b} \langle \mathbf{grad} \psi(\mathbf{c}(t)), \dot{\mathbf{c}}(t) \rangle dt \\ &= \int_{t_a}^{t_b} \frac{d}{dt} \psi(\mathbf{c}(t)) dt = \psi(\mathbf{c}(t_b)) - \psi(\mathbf{c}(t_a)). \end{aligned}$$

Insbesondere hängt in diesem Fall der Wert des Kurvenintegrals nur vom Anfangs- und vom Endpunkt der Kurve \mathbf{c} sowie von ψ ab.

Approximation durch Polygonzüge* Jede Kurve $\mathbf{c} : I \rightarrow \mathbb{R}^n$ mit kompakten Parametrisierungsintervall I kann durch einen Polygonzug approximiert werden kann. Dazu zerlegen wir I via

$$t_a = t_0 < t_1 < \dots < t_{K-1} < t_K = t_e$$

in K Teilintervalle und definieren durch

$$\mathbf{c}_{\text{app}}(t) := \frac{t_k - t}{t_k - t_{k-1}} \mathbf{c}(t_{k-1}) + \frac{t - t_{k-1}}{t_k - t_{k-1}} \mathbf{c}(t_k) \quad \text{für } t \in [t_{k-1}, t_k]$$

eine stückweise affine Kurve $\mathbf{c}_{\text{app}} : I \rightarrow \mathbb{R}^n$, für die

$$\dot{\mathbf{c}}_{\text{app}}(t) = \frac{\mathbf{c}(t_k) - \mathbf{c}(t_{k-1})}{t_k - t_{k-1}}, \quad t \in (t_{k-1}, t_k)$$

gilt (siehe Bild). Diese Kurve \mathbf{c}_{app} approximiert \mathbf{c} , sofern K hinreichend groß und die Feinheit der Zerlegung — also $\max_{k=1 \dots K} t_k - t_{k-1}$ — hinreichend klein ist. Sie ist nach Konstruktion immer stetig, da

$$\mathbf{c}_{\text{app}}(t_{k-1} + 0) = \mathbf{c}_{\text{app}}(t_k - 0) = \mathbf{c}(t_k)$$

für jedes $k = 2 \dots K-1$ im Sinne einseitiger Grenzwerte gilt, aber im Allgemeinen nur stückweise stetig differenzierbar wegen

$$\dot{\mathbf{c}}_{\text{app}}(t_{k-1} + 0) \neq \dot{\mathbf{c}}_{\text{app}}(t_k - 0).$$

Ein Kurvenintegral bzgl. \mathbf{c}_{app} approximiert das entsprechende Kurvenintegral bzgl. \mathbf{c} und kann segmentweise ausgewertet werden. Zum Beispiel gilt

$$\int_{\mathbf{c}_{\text{app}}} \mathbf{f}(\mathbf{x}) \cdot d\mathbf{x} = \sum_{k=1}^K \int_{t_{k-1}}^{t_k} \langle \mathbf{f}(\mathbf{c}_{\text{app}}(t)), \dot{\mathbf{c}}_{\text{app}}(t) \rangle dt$$

und wenn wir $\mathbf{f}(\mathbf{c}_{\text{app}}(t))$ durch $\mathbf{f}(\mathbf{c}_{\text{app}}(t_k)) = \mathbf{f}(\mathbf{c}(t_k))$ ersetzen und die so vereinfachten Integrale (in denen der Integrand nicht mehr von t abhängt) direkt berechnen, erhalten wir schließlich die Näherungsformel

$$\int_{\mathbf{c}} \mathbf{f}(\mathbf{x}) \cdot d\mathbf{x} \approx \sum_{k=1}^K \langle \mathbf{f}(\mathbf{c}(t_k)), \mathbf{c}(t_k) - \mathbf{c}(t_{k-1}) \rangle.$$

Analog kann

$$\int_{\mathbf{c}} f(\mathbf{x}) \, ds \approx \sum_{k=1}^K f(\mathbf{c}(t_k)) |\mathbf{c}(t_k) - \mathbf{c}(t_{k-1})|$$

hergeleitet werden. Diese zwei Formeln sind nicht nur aus numerischer Sicht nützlich, sondern sie verdeutlichen einmal mehr das Konzept eines Kurvenintegrals. Die Definitionen der Kurvenintegrale sind nämlich nichts anderes als die Grenzwerte dieser Summenformeln im Limes $K \rightarrow \infty$ (so wie Gebietsintegrale die Grenzwerte von Riemann-Summen sind).

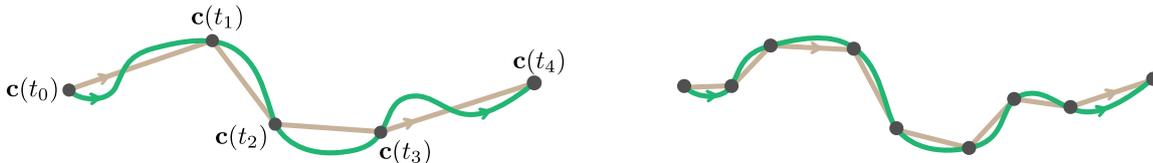


Abbildung Approximation einer Kurve \mathbf{c} (grün) durch einen Polygonzug \mathbf{c}_{app} (braun), wobei links $K = 4$ und rechts $K = 8$ gewählt wurde. Zu jedem Kurvenintegral bzgl. \mathbf{c} gibt es eine entsprechende Näherungsformeln mit Summen.

Addition und Subtraktion mit Kurven Sind $\mathbf{c}^{(1)}$ und $\mathbf{c}^{(2)}$ zwei Kurven mit

$$\mathbf{c}^{(1)}(t_e^{(1)}) := \mathbf{c}^{(2)}(t_a^{(2)}), \quad t_e^{(1)} = t_a^{(2)}$$

so können wir diese „zusammenkleben“ (siehe Bild). In Formeln beschreiben wir dies durch die Kurve

$$(\mathbf{c}^{(1)} \oplus \mathbf{c}^{(2)})(t) = \begin{cases} \mathbf{c}^{(1)}(t) & \text{für } t_a^{(1)} \leq t \leq t_e^{(1)}, \\ \mathbf{c}^{(2)}(t) & \text{für } t_a^{(2)} \leq t \leq t_e^{(2)}, \end{cases}$$

wobei wir das Parametrisierungsintervall $[t_a^{(1)}, t_e^{(2)}]$ zu Grunde gelegt haben. Beachte, dass die Bedingung $\mathbf{c}^{(1)}(t_e^{(1)}) = \mathbf{c}^{(2)}(t_a^{(2)})$ wirklich eine geometrische Restriktion ist, wohingegen $t_e^{(1)} = t_a^{(2)}$ immer durch eine sehr einfache Reparametrisierung (bzw. durch eine Zeitverschiebung) von $\mathbf{c}^{(1)}$ oder $\mathbf{c}^{(2)}$ sichergestellt werden kann. Oder anders gesagt: Wir können $\mathbf{c}^{(1)} \oplus \mathbf{c}^{(2)}$ immer dann sinnvoll definieren, wenn der Endpunkt von $\mathbf{c}^{(1)}$ der Anfangspunkt von $\mathbf{c}^{(2)}$ ist, wobei man ein beliebiges Parametrisierungsintervall der Länge $(t_e^{(1)} - t_a^{(1)}) + (t_e^{(2)} - t_a^{(2)})$ wählen kann.

Zu einer gegebenen Kurve \mathbf{c} können wir durch

$$(\ominus \mathbf{c})(t) := \mathbf{c}(2t_e - t) \quad t \in [t_e, 2t_e - t_a]$$

ihre Umkehrung $\ominus \mathbf{c}$ einführen, deren Anfangs- bzw. Endpunkt gerade der End- bzw. Anfangspunkt von \mathbf{c} ist. Beide Kurven besitzen dieselbe Bildmenge, aber einen anderen Durchlaufsin.

Sind $\mathbf{c}^{(1)}$ und $\mathbf{c}^{(2)}$ zwei Kurven mit identischen Endpunkten $\mathbf{c}^{(1)}(t_e^{(1)}) = \mathbf{c}^{(2)}(t_e^{(2)})$, so startet die Kurve

$$\mathbf{c}^{(1)} \ominus \mathbf{c}^{(2)} := \mathbf{c}^{(1)} \oplus (\ominus \mathbf{c}^{(2)})$$

im Punkt $\mathbf{c}^{(1)}(t_a^{(1)})$, läuft so wie $\mathbf{c}^{(1)}$ zum gemeinsamen Endpunkt, um anschließend sich rückwärts entlang von $\mathbf{c}^{(2)}$ zum Punkt $\mathbf{c}^{(2)}(t_a^{(2)})$ zu bewegen (siehe Bild). Insbesondere ist $\mathbf{c} \ominus \mathbf{c}$ immer ein geschlossener Weg, wobei jeder Punkt aus im (\mathbf{c}) genau zweimal, nämlich einmal auf dem Hinweg und einmal auf dem Rückweg, passiert wird.

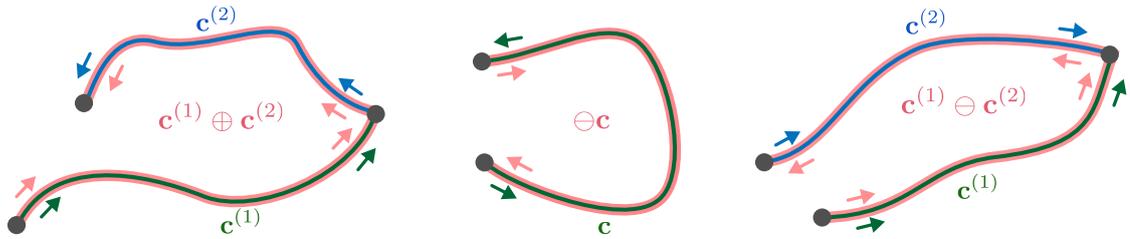


Abbildung Die Addition (links) und die Subtraktion (rechts) von jeweils zwei Kurven sowie die „Umkehrung“ einer Kurve (Mitte). Der jeweilige Durchlaufsinne wird durch die farbigen Pfeile angegeben.

Beispiel Die planare Kurve

$$\mathbf{c} = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}, \quad t \in [0, 2\pi]$$

durchläuft die Einheitskreislinie genau einmal in mathematisch positiver Richtung, also entgegen dem Uhrzeigersinn, wohingegen $\ominus \mathbf{c}$ dasselbe, aber im Uhrzeigersinn tut. Die Kurven $\mathbf{c} \oplus \mathbf{c}$ bzw. $\mathbf{c} \oplus \mathbf{c} \oplus \mathbf{c}$ laufen auch im Uhrzeigersinn, aber passieren dabei jeden Punkt der Kreislinie zwei- bzw. dreimal.

Rechenregeln für Kurvenintegrale der 2. Art Unsere Definitionen implizieren die folgenden Aussagen:

1. (Linearität bzgl. \mathbf{f})
$$\int_{\mathbf{c}} (\alpha \mathbf{f}(\mathbf{x}) + \tilde{\alpha} \tilde{\mathbf{f}}(\mathbf{x})) \cdot d\mathbf{x} = \alpha \int_{\mathbf{c}} \mathbf{f}(\mathbf{x}) \cdot d\mathbf{x} + \tilde{\alpha} \int_{\mathbf{c}} \tilde{\mathbf{f}}(\mathbf{x}) \cdot d\mathbf{x}$$
2. (Additivität bzgl. der Kurve)
$$\int_{\mathbf{c}^{(1)} \oplus \mathbf{c}^{(2)}} \mathbf{f}(\mathbf{x}) \cdot d\mathbf{x} = \int_{\mathbf{c}^{(1)}} \mathbf{f}(\mathbf{x}) \cdot d\mathbf{x} + \int_{\mathbf{c}^{(2)}} \mathbf{f}(\mathbf{x}) \cdot d\mathbf{x}$$
3. (Subtraktivität bzgl. der Kurve)
$$\int_{\mathbf{c}^{(1)} \ominus \mathbf{c}^{(2)}} \mathbf{f}(\mathbf{x}) \cdot d\mathbf{x} = \int_{\mathbf{c}^{(1)}} \mathbf{f}(\mathbf{x}) \cdot d\mathbf{x} - \int_{\mathbf{c}^{(2)}} \mathbf{f}(\mathbf{x}) \cdot d\mathbf{x}$$

Bemerkung

1. Ein Spezialfall der dritten Formel ist

$$\int_{\ominus \mathbf{c}} \mathbf{f}(\mathbf{x}) \cdot d\mathbf{x} = - \int_{\mathbf{c}} \mathbf{f}(\mathbf{x}) \cdot d\mathbf{x}$$

und physikalisch ist das sehr einleuchtend: Ist W die mechanische Arbeit, um ein Teilchen entlang einer Kurve von ihrem Anfangspunkt zu ihrem Endpunkt zu bewegen, so muss die Arbeit $-W$ verrichtet werden, um das Teilchen auf der Kurve zurückzubewegen.

2. Die ersten beiden Eigenschaften gelten analog für Kurvenintegrale der 1. Art. Bei der \ominus -Operation ändert sich aber nicht das Vorzeichen, d.h. es gilt

$$\int_{\ominus \mathbf{c}} \mathbf{f}(\mathbf{x}) \, ds = \int_{\mathbf{c}} \mathbf{f}(\mathbf{x}) \, ds.$$

Dies liegt daran, dass sich bei einer Umkehrung des Durchlaufsinns zwar die Richtung der Tangentialvektoren $\dot{\mathbf{c}}$, aber nicht ihre Norm ändert. Etwas genauer: Mit $\tilde{\mathbf{c}} = \ominus \mathbf{c}$ ergibt sich $\dot{\tilde{\mathbf{c}}}(t) = -\dot{\mathbf{c}}(2t_e - t)$ und $|\dot{\tilde{\mathbf{c}}}(t)| = |\dot{\mathbf{c}}(2t_e - t)|$ aus der obigen Definition von $\ominus \mathbf{c}$ nach Differentiation bzgl. t .

Vorlesung 09-2 : 13. Januar

Theorem (Invarianz bei Reparametrisierung) Seien $\mathbf{c} : I \rightarrow D$ und $\tilde{\mathbf{c}} : \tilde{I} \rightarrow D$ zwei Kurven, so dass die Gleichungen

$$\tilde{\mathbf{c}}(\tilde{t}) = \mathbf{c}(h(\tilde{t})) \quad \text{und} \quad \mathbf{c}(t) = \tilde{\mathbf{c}}(\tilde{h}(t))$$

für alle $t \in I$ und $\tilde{t} \in \tilde{I}$ sowie zwei *bijektive, monoton wachsende* und *zueinander inverse* Funktionen $h : \tilde{I} \rightarrow I$ und $\tilde{h} : I \rightarrow \tilde{I}$ erfüllt sind. Dann gilt

$$\int_{\tilde{\mathbf{c}}} f(\mathbf{x}) \, ds = \int_{\mathbf{c}} f(\mathbf{x}) \, ds \quad \text{bzw.} \quad \int_{\tilde{\mathbf{c}}} \mathbf{f}(\mathbf{x}) \cdot d\mathbf{x} = \int_{\mathbf{c}} \mathbf{f}(\mathbf{x}) \cdot d\mathbf{x}$$

für jedes $f : D \rightarrow \mathbb{R}$ bzw. $\mathbf{f} : D \rightarrow \mathbb{R}^n$.

Beweis Wir beweisen nur die zweite Formel; die Herleitung der ersten erfolgt analog. Nach Voraussetzung und Kettenregel gilt

$$\mathbf{f}(\tilde{\mathbf{c}}(\tilde{t})) = \mathbf{f}(\mathbf{c}(h(\tilde{t}))), \quad \frac{d\tilde{\mathbf{c}}}{d\tilde{t}}(\tilde{t}) = \frac{d\mathbf{c}}{dt}(h(\tilde{t})) \frac{dh}{d\tilde{t}}(\tilde{t})$$

und die Transformationsformel für 1D-Integrale liefert

$$\begin{aligned} \int_{\tilde{\mathbf{c}}} \mathbf{f}(\mathbf{x}) \cdot d\mathbf{x} &= \int_{\tilde{t}_a}^{\tilde{t}_e} \left\langle \mathbf{f}(\tilde{\mathbf{c}}(\tilde{t})), \frac{d\tilde{\mathbf{c}}}{d\tilde{t}}(\tilde{t}) \right\rangle d\tilde{t} = \int_{\tilde{t}_a}^{\tilde{t}_e} \left\langle \mathbf{f}(\mathbf{c}(h(\tilde{t}))), \frac{d\mathbf{c}}{dt}(h(\tilde{t})) \right\rangle \frac{dh}{d\tilde{t}}(\tilde{t}) d\tilde{t} \\ &= \int_{t_a}^{t_e} \left\langle \mathbf{f}(\mathbf{c}(t)), \frac{d\mathbf{c}}{dt}(t) \right\rangle dt = \int_{\mathbf{c}} \mathbf{f}(\mathbf{x}) \cdot d\mathbf{x} \end{aligned}$$

wobei die Details der Substitution durch

$$t = h(\tilde{t}), \quad t_a = h(\tilde{t}_a), \quad t_e = h(\tilde{t}_e), \quad dt = \frac{dh}{d\tilde{t}}(\tilde{t}) d\tilde{t}$$

beschrieben sind.

Bemerkung

1. Die Kurve $\tilde{\mathbf{c}}$ ist eine Reparametrisierung von \mathbf{c} und umgekehrt, wobei die jeweiligen Parameterwechsel durch die Funktionen h bzw. \tilde{h} vermittelt werden. Da diese invers zueinander sind, gilt stets $\tilde{t} = \tilde{h}(h(\tilde{t}))$ bzw. $t = h(\tilde{h}(t))$.
2. Das monotone Wachstum von h und \tilde{h} ist wichtig, da aus $t_a < t_e$ bei monoton fallenden Reparametrisierungen $\tilde{t}_a = \tilde{h}(t_a) > \tilde{t}_e = \tilde{h}(t_e)$ wird, d.h. ein Wechsel im Durchlaufsinns stattfindet. Bei Kurvenintegralen der 2. Art ändert sich dadurch das Vorzeichen, wohingegen Kurvenintegrale der 1. Art (zum Beispiel die Länge der Kurve) unverändert bleiben.

3. Reparametrisierungen der Bauart $\tilde{t} = t - t_*$ mit festem t_* entsprechen einer Zeitverschiebung, wohingegen $\tilde{t} = \sigma t$ mit Konstante $\sigma > 0$ eine uniforme Beschleunigung (oder Entschleunigung) der Zeit beschreibt. Alternativ kann man diese Parameterwechsel als Wahl einer Referenzzeit bzw. einer anderen Zeiteinheit interpretieren.
4. Es gibt viele sinnvolle und wichtige Reparametrisierungen, bei denen t , aber nicht \tilde{t} als Zeit interpretiert werden kann. In diesem Fall benutzt man an Stelle von \tilde{t} und $\tilde{\mathbf{c}}$ gerne andere Notationen und schreibt meist $'$ statt \cdot für die Ableitung nach dem Kurvenparameter. Ein prominentes Beispiel ist die Reparametrisierung einer Kurve nach ihrer Bogenlänge (siehe Kapitel 1), bei der t durch den Parameter s ersetzt wird, für den $ds = |\dot{\mathbf{c}}(t)| dt$ gilt.

3.2 Grundlagen der Potentialtheorie

Ziel Gradientenfelder spielen in vielen Anwendungsbereichen eine wichtige Rolle. In diesem Abschnitt untersuchen wir, unter welchen Umständen ein gegebenes Vektorfeld der Gradient einer skalaren Funktion ist oder nicht. Wir setzen dabei immer voraus, dass $D \subset \mathbb{R}^n$ zusammenhängend ist und dass $\mathbf{f} : D \rightarrow \mathbb{R}^n$ stetig ist.

Erinnerung Eine Menge $D \subset \mathbb{R}^n$ heißt zusammenhängend, falls es zu je zwei Punkten eine verbindende Kurve gibt. Salopp kann man auch sagen: Eine Menge ist genau dann zusammenhängend, wenn sie nicht in zwei oder mehrere, voneinander getrennte, Teile zerfällt.

Gradientenfelder und Potentiale

Theorem (Hauptsatz über Kurvenintegrale) Unter den Voraussetzungen an D und \mathbf{f} sind die folgenden drei Aussagen paarweise äquivalent:

1. Das Vektorfeld \mathbf{f} ist zirkulationsfrei, d.h. es gilt

$$\int_{\mathbf{c}} \mathbf{f}(\mathbf{x}) \cdot d\mathbf{x} = 0$$

für jede *geschlossene* Kurve \mathbf{c} in D .

2. Das Kurvenintegral bzgl. \mathbf{f} ist wegunabhängig, d.h. es gilt

$$\int_{\mathbf{c}^{(1)}} \mathbf{f}(\mathbf{x}) \cdot d\mathbf{x} = \int_{\mathbf{c}^{(2)}} \mathbf{f}(\mathbf{x}) \cdot d\mathbf{x}$$

für je zwei Wege $\mathbf{c}^{(1)}$ und $\mathbf{c}^{(2)}$ in D , die denselben Anfangspunkt mit demselben Endpunkt verbinden.

3. Das Vektorfeld \mathbf{f} ist ein Gradientenfeld, d.h. es existiert eine stetig differenzierbare Funktion $\psi : D \rightarrow \mathbb{R}$, so dass

$$\mathbf{f}(\mathbf{x}) = \mathbf{grad} \psi(\mathbf{x})$$

für alle $\mathbf{x} \in D$ gilt. Dabei wird ψ das Potential von \mathbf{f} genannt.

Beweis Wir weisen die Äquivalenzen (1.) \Leftrightarrow (2.) und (2.) \Leftrightarrow (3.) nach, wobei jede in Form von zwei Implikationen etabliert wird. Die Gültigkeit von (1.) \Leftrightarrow (3.) ist dann automatisch sichergestellt.

(1.) \Rightarrow (2.) und (2.) \Rightarrow (1.): Zwei gegebene Kurven $\mathbf{c}^{(1)}$ und $\mathbf{c}^{(2)}$ mit identischem Anfangs- und Endpunkt können via $\mathbf{c} := \mathbf{c}^{(1)} \ominus \mathbf{c}^{(2)}$ zu einer geschlossenen Kurve vereint werden (siehe Bild). Unter Verwendung von (1.) erhalten wir dann

$$0 = \int_{\mathbf{c}} \mathbf{f}(\mathbf{x}) \cdot d\mathbf{x} = \int_{\mathbf{c}^{(1)}} \mathbf{f}(\mathbf{x}) \cdot d\mathbf{x} - \int_{\mathbf{c}^{(2)}} \mathbf{f}(\mathbf{x}) \cdot d\mathbf{x}$$

und damit (2.) nach Umstellung der Terme. Umgekehrt kann jede geschlossene Kurve \mathbf{c} als $\mathbf{c} = \mathbf{c}^{(1)} \ominus \mathbf{c}^{(2)}$ geschrieben werden und unsere Formel zeigt deshalb, dass (2.) aus (1.) folgt.

(3.) \Rightarrow (2.): Wir hatten schon oben gesehen, dass die Existenz von ψ die Formel

$$\int_{\mathbf{c}^{(k)}} \mathbf{grad} \psi(\mathbf{x}) \cdot d\mathbf{x} = \psi(\mathbf{c}^{(k)}(t_e^{(k)})) - \psi(\mathbf{c}^{(k)}(t_a^{(k)}))$$

für $k = 1$ und $k = 2$ und damit auch die Wegunabhängigkeit impliziert.

(2.) \Rightarrow (3.): Dies ist der anspruchsvollste Teil des gesamten Beweises, aber die zu Grunde liegenden Ideen sind eigentlich sehr naheliegend (siehe wieder das Bild) und spiegeln wichtige physikalische und geometrische Prinzipien wider. Wir fixieren einen beliebigen Punkt $\xi_* \in D$, wählen für jeden anderen Punkt $\xi \in D$ eine Kurve $\mathbf{c}_{\xi_*, \xi}$, die ξ_* mit ξ verbindet, und setzen

$$\psi(\xi) := \int_{\mathbf{c}_{\xi_*, \xi}} \mathbf{f}(\mathbf{x}) \cdot d\mathbf{x}.$$

Es wird sehr viele solcher Kurven geben, die von ξ_* nach ξ führen, aber die Bedingung (2.) garantiert, dass $\psi(\xi)$ trotzdem wohldefiniert ist, eben weil jede Kurve denselben Wert für das Integral liefert. Ist $\zeta \in D$ ein weiterer Punkt und sind $\mathbf{c}_{\xi_*, \zeta}$ bzw. $\mathbf{c}_{\zeta, \xi}$ eine beliebige Verbindungskurve von ξ_* nach ζ bzw. von ζ nach ξ (siehe das Bild), so implizieren (2.) sowie die Rechenregeln für Kurvenintegrale die Formel

$$\psi(\xi) = \int_{\mathbf{c}_{\xi_*, \zeta} \oplus \mathbf{c}_{\zeta, \xi}} \mathbf{f}(\mathbf{x}) \cdot d\mathbf{x} = \int_{\mathbf{c}_{\xi_*, \zeta}} \mathbf{f}(\mathbf{x}) \cdot d\mathbf{x} + \int_{\mathbf{c}_{\zeta, \xi}} \mathbf{f}(\mathbf{x}) \cdot d\mathbf{x} = \psi(\zeta) + \int_{\mathbf{c}_{\zeta, \xi}} \mathbf{f}(\mathbf{x}) \cdot d\mathbf{x},$$

da auch die Kurve $\mathbf{c}_{\xi_*, \zeta} \oplus \mathbf{c}_{\zeta, \xi}$ die Punkte ξ_* mit ξ miteinander verbindet. Liegen die Punkte ξ und ζ nahe beieinander, so können wir annehmen, dass \mathbf{f} sich auf $\mathbf{c}_{\zeta, \xi}$ nur wenig ändert. Insbesondere gilt dann $\mathbf{f}(\mathbf{x}) \approx \mathbf{f}(\xi)$ entlang dieser Kurve und wir erhalten die Näherungsformel

$$\int_{\mathbf{c}_{\zeta, \xi}} \mathbf{f}(\mathbf{x}) \cdot d\mathbf{x} = \int_{\mathbf{c}_{\zeta, \xi}} \mathbf{f}(\xi) \cdot d\mathbf{x} + O(|\xi - \zeta|^2) = \langle \mathbf{f}(\xi), \xi - \zeta \rangle + O(|\xi - \zeta|^2),$$

wobei wir das approximative Kurvenintegral mit dem (aus Sicht von \mathbf{x}) konstanten Vektorfeld explizit berechnet haben. Insgesamt ergibt sich

$$\psi(\xi) - \psi(\zeta) = \langle \mathbf{f}(\xi), \xi - \zeta \rangle + O(|\xi - \zeta|^2)$$

und wir schließen, dass ψ im Punkt ξ in der Tat differenzierbar mit $\mathbf{grad} \psi(\xi) = \mathbf{f}(\xi)$ ist. Da ξ beliebig war, können wir am Ende ξ durch \mathbf{x} ersetzen und haben damit (3.) aus (2.) abgeleitet. \square

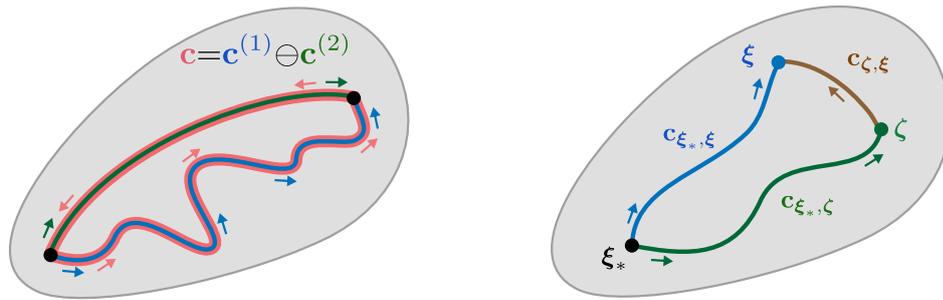


Abbildung Der Beweis des Hauptsatzes in Bildern, wobei links bzw. rechts die Schritte (1) \Leftrightarrow (2) bzw. (2) \Rightarrow (3) illustriert sind.

Bemerkung

1. Das Potential ψ ist nicht eindeutig, sondern nur bis auf eine Konstante bestimmt. Oder anders gesagt: Sind ψ und $\tilde{\psi}$ zwei Potentiale von \mathbf{x} , so gilt $\tilde{\psi}(\mathbf{x}) = \psi(\mathbf{x}) + d$ für alle $x \in D$ mit einer Konstanten $d \in \mathbb{R}$, die nicht von \mathbf{x} abhängt. Physikalisch meint dies, dass man das Potential in einem frei wählbaren Punkt Null setzen darf (die sogenannte *Erdung* in der Elektrostatik). Die (harmlose) Mehrdeutigkeit erkennen wir auch im letzten Beweisschritt, da dort der Punkt ξ_* beliebig gewählt wurde und $\psi(\xi_*) = 0$ für das konstruierte Potential gilt.
2. Der Hauptsatz kann als eine von vielen möglichen Verallgemeinerung des eindimensionalen Hauptsatzes der Differential- und Integralrechnung verstanden werden. Andere Verallgemeinerungen sind die Integralsätze von Gauß und Stokes.
3. Ist D nicht zusammenhängend, so kann der Hauptsatz auf jeder Zusammenhangskomponente, d.h. auf jedem der separierten Teile, von D angewendet werden. Zwei Potentiale zu \mathbf{f} werden sich dann aber nicht mehr nur durch eine globale Konstante unterscheiden, sondern auf jeder Komponente kann es eine andere Konstante geben.
4. Achtung: In der Physik wird üblicherweise eine andere Vorzeichenkonvention getroffen. Die Gleichung

$$\mathbf{f}(\mathbf{x}) = + \mathbf{grad} \psi(\mathbf{x})$$

wird dort meist als

$$\mathbf{f}(\mathbf{x}) = - \mathbf{grad} \phi(\mathbf{x})$$

geschrieben, das heißt das physikalische Potential

$$\phi(\mathbf{x}) = -\psi(\mathbf{x})$$

ist aus mathematischer Sicht eigentlich das negative Potential und umgekehrt.

Bemerkung: Ähnliche Abweichungen in der Vorzeichenwahl gibt es bei der mathematischen und der physikalischen Entropie, der physikalischen und der technischen Stromrichtung sowie bei der Fourier-Transformation.

5. Der Hauptsatz über Kurvenintegrale ist von fundamentaler Bedeutung (auch in den Anwendungswissenschaften), da er unter allen Vektorfeldern die Klasse der Gradientenfelder identifiziert. Mit ihm kann man zum Beispiel sehr leicht zeigen,

dass für ein gegebenes Vektorfeld *kein* Potential existiert, denn man muss ja nur *eine* geschlossene Kurve angeben, für die das entsprechende Kurvenintegral nicht verschwindet. Will man allerdings zeigen, dass ein gegebenes Vektorfeld tatsächlich Gradientenfeld ist, so ist der Hauptsatz eher ungeeignet, da man zum Beispiel *alle* denkbaren geschlossenen Kurven betrachten muss. Wir werden unten auf diesen Aspekt zurückkommen.

Beispiele

1. Für jedes $\mathbf{v} \in \mathbb{R}^n$ besitzt das entsprechende konstante Vektorfeld ein lineares Potential, wobei dann

$$\mathbf{f}(\mathbf{x}) = \mathbf{v}, \quad \psi(\mathbf{x}) = \langle \mathbf{x}, \mathbf{v} \rangle$$

für alle $\mathbf{x} \in \mathbb{R}^n$ gilt.

2. Analog gilt

$$\mathbf{f}(\mathbf{x}) = \mathbf{A} \cdot \mathbf{x}, \quad \psi(\mathbf{x}) = \frac{1}{2} \langle \mathbf{A} \cdot \mathbf{x}, \mathbf{x} \rangle$$

für jedes lineare Vektorfeld mit symmetrischer Matrix $\mathbf{A} = \mathbf{A}^T \in \mathbb{R}^{(n,n)}$.

3. Das dreidimensionale Vektorfeld

$$\mathbf{f}(\mathbf{x}) = \begin{pmatrix} 2x_1 + x_2 x_3 \\ x_1 x_3 \\ x_1 x_2 \end{pmatrix}$$

ist ein Gradientenfeld auf $D = \mathbb{R}^3$, denn wir können durch scharfes Hinsehen das Potential

$$\psi(\mathbf{x}) = x_1^2 + x_1 x_2 x_3$$

ablesen. Alternativ können wir es auch durch Kurvenintegrale berechnen. Mit den Notationen des Beweises wählen wir zum Beispiel

$$\boldsymbol{\xi}_* = \mathbf{0} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{c}_{\boldsymbol{\xi}_*, \boldsymbol{\xi}}(t) = \begin{pmatrix} t \xi_1 \\ t \xi_2 \\ t \xi_3 \end{pmatrix} \quad \text{mit } t \in [0, 1],$$

d.h. wir verbinden den Ursprung $\boldsymbol{\xi}_*$ mit jedem Punkt $\boldsymbol{\xi}$ durch ein Geradenstück. Durch diese Wahl sind die Integrale relativ einfach zu berechnen und wir erhalten via

$$\begin{aligned} \psi(\boldsymbol{\xi}) &= \int_{\mathbf{c}_{\boldsymbol{\xi}_*, \boldsymbol{\xi}}} \mathbf{f}(\mathbf{x}) \cdot d\mathbf{x} = \int_0^1 \left\langle \begin{pmatrix} 2t \xi_1 + t^2 \xi_2 \xi_3 \\ t^2 \xi_1 \xi_3 \\ t^2 \xi_1 \xi_2 \end{pmatrix}, \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix} \right\rangle dt \\ &= \int_0^1 (2 \xi_1^2 t + 3 \xi_1 \xi_2 \xi_3 t^2) dt \\ &= \xi_1 [t^2]_{t=0}^{t=1} + \xi_1 \xi_2 \xi_3 [t^3]_{t=0}^{t=1} = \xi_1^2 + \xi_1 \xi_2 \xi_3 \end{aligned}$$

das gleiche Ergebnis für ψ , sofern wir am Ende \mathbf{x} statt $\boldsymbol{\xi}$ schreiben.

4. Wir können für jedes $n \geq 2$ und das Zentralfeld

$$\mathbf{f}(\mathbf{x}) = c |\mathbf{x}|^p \mathbf{x}, \quad \mathbf{x} \neq \mathbf{0}$$

leicht nachrechnen (mit $\partial_{x_j} |\mathbf{x}| = x_j / |\mathbf{x}|$), dass durch

$$\psi(\mathbf{x}) = \frac{c}{p+2} |\mathbf{x}|^{p+2}$$

ein entsprechendes Potential auf $\mathbb{R}^n \setminus \{\mathbf{0}\}$ definiert wird, wobei c eine reelle Konstante und $p \neq -2$ ein gegebener Exponent ist. Insbesondere gilt

$$\int_{\mathbf{c}} \mathbf{f}(\mathbf{x}) \cdot d\mathbf{x} = \frac{c}{p+2} (|\mathbf{c}(t_e)|^{p+2} - |\mathbf{c}(t_a)|^{p+2})$$

für jede Kurve \mathbf{c} , die nicht durch $\mathbf{0}$ läuft (dort besitzt das Vektorfeld für $p < -2$ eine Singularität). Für $p = -2$ lautet die entsprechende Formel jedoch

$$\psi(\mathbf{x}) = c \ln(|\mathbf{x}|),$$

und mit $n = 2$ erhalten wir gerade das elementare Quellenfeld f_Q in 2D, das wir am Ende des letzten Kapitels eingeführt hatten.

Wichtiger Spezialfall: Mit $n = 3$, $p = -3$, $c = -m c_g$ ergibt sich für

$$\mathbf{f}(\mathbf{x}) = -m c_g \frac{\mathbf{x}}{|\mathbf{x}|^3}$$

via

$$\phi(\mathbf{x}) = -\psi(\mathbf{x}) = -\frac{m c_g}{|\mathbf{x}|}$$

gerade das physikalische Gravitationspotential einer im Ursprung konzentrierten Punktmasse, wobei m die Masse und c_g die universelle Gravitationskonstante ist. Mit einer anderen Wahl der Konstanten erhalten wir das dreidimensionale elektrische Feld, dass von einer Punktladung erzeugt wird.

Gegenbeispiel Das elementare Wirbelfeld \mathbf{f}_W (siehe oben) besitzt kein Potential auf $\mathbb{R}^2 \setminus \{\mathbf{0}\}$, eben weil es geschlossene Kurven mit nicht verschwindendem Kurvenintegral gibt. Zum Beispiel Kreislinien mit Mittelpunkt $\mathbf{0}$.

Bemerkung: Auf jeder *einfach zusammenhängenden* Teilmenge $D \subset \mathbb{R}^2 \setminus \{\mathbf{0}\}$ können wir aber immer ein lokales Potential einführen, eben weil wir in einer solchen Menge nicht um den Punkt $\mathbf{x} = \mathbf{0}$ herumlaufen können. Wir werden diesen Aspekt in der Vorlesung *Funktionentheorie* genauer studieren.

Integrabilitätsbedingungen

Ziel Wir untersuchen nun, ob und wie man einem Vektorfeld direkt ansehen kann, ob es ein Gradientenfeld ist. Wir setzen dabei immer voraus, dass $\mathbf{f} : D \rightarrow \mathbb{R}^n$ nicht nur stetig, sondern sogar stetig differenzierbar ist.

Notwendige Bedingungen Gilt $\mathbf{f} = \mathbf{grad} \psi$, d.h. $f_j = \partial_{x_j} \psi$, so impliziert der Satz von Schwarz die Bedingung

$$\partial_{x_j} f_i(\mathbf{x}) = \partial_{x_i} f_j(\mathbf{x})$$

für alle $\mathbf{x} \in D$ und alle $i, j = 1 \dots n$ mit $i \neq j$. Oder anders gesagt: \mathbf{f} kann nur dann ein Gradientenfeld sein, wenn die Jacobi-Matrix $\mathbf{Jf}(\mathbf{x}) \in \mathbb{R}^{(n,n)}$ in jedem Punkt symmetrisch ist. Die Frage ist nun, ob auch die Umkehrung gilt.

Bemerkung Für $n = 2$ gilt offensichtlich

$$\mathbf{Jf}(\mathbf{x}) = (\mathbf{Jf}(\mathbf{x}))^T \quad \Leftrightarrow \quad \mathbf{rot} \mathbf{f}(\mathbf{x}) = 0,$$

und für $n = 3$ verifizieren wir analog

$$\mathbf{Jf}(\mathbf{x}) = (\mathbf{Jf}(\mathbf{x}))^T \quad \Leftrightarrow \quad \mathbf{rot} \mathbf{f}(\mathbf{x}) = 0.$$

Die Rotation von \mathbf{f} misst also gerade die Asymmetrie der $(2, 2)$ -Matrix $\mathbf{Jf}(\mathbf{x})$ und wir sehen, dass nur rotationsfreie Vektorfelder auch Gradientenfelder sein können.

Beispiel und Gegenbeispiel Wir betrachten für $n = 2$ nochmal das elementare Quellenfeld \mathbf{f}_Q sowie das elementare Wirbelfeld \mathbf{f}_W vom Ende des letzten Kapitels auf der Menge $D = \mathbb{R}^2 \setminus \{\mathbf{0}\}$ und wählen die Konstante $c = 1$. Die entsprechenden Jacobi-Matrizen berechnen sich zu

$$\mathbf{Jf}_Q(\mathbf{x}) = \frac{1}{(x_1^2 + x_2^2)^2} \begin{pmatrix} x_2^2 - x_1^2 & x_1 x_2 \\ x_1 x_2 & x_1^2 - x_2^2 \end{pmatrix}$$

bzw.

$$\mathbf{Jf}_W(\mathbf{x}) = \frac{1}{(x_1^2 + x_2^2)^2} \begin{pmatrix} -x_1 x_2 & x_2^2 - x_1^2 \\ x_2^2 - x_1^2 & +x_1 x_2 \end{pmatrix}$$

und sind offensichtlich symmetrisch. Beide Vektorfelder sind also rotationsfrei (sie sind außerdem auch divergenzfrei, aber das interessiert uns im Moment nicht) und erfüllen damit die notwendigen Bedingungen in jedem Punkt $\mathbf{x} \neq \mathbf{0}$ (für $\mathbf{x} = \mathbf{0}$ sind die Terme nicht definiert). Wir hatten aber schon oben gesehen, dass das Quellenfeld ein Potential besitzt, nämlich

$$\psi_Q(\mathbf{x}) = \ln \left(\sqrt{x_1^2 + x_2^2} \right),$$

aber dass Wirbelfeld \mathbf{f}_W kein Gradientenfeld sein kann, da das Kurvenintegral entlang einer beliebigen Kreislinie mit Mittelpunkt $\mathbf{0}$ nicht verschwindet. Die notwendigen Bedingungen können also nicht hinreichend sein.

Bemerkung: Es existieren zwei Probleme bei \mathbf{f}_W : Zum einen gibt es die Singularität bei $\mathbf{0}$, aber dieses *Loch* im Definitionsbereich D existiert es auch bei \mathbf{f}_Q . Das zweite Problem ist, dass das Drehfeld \mathbf{f}_W um dieses Loch rotiert, wohingegen \mathbf{f}_Q dies nicht tut, sondern als Zentralfeld von diesem Loch weg zeigt.

Einfach zusammenhängende Mengen Zusammenhängende Mengen können sehr unterschiedliche Eigenschaften haben und man unterteilt bzw. klassifiziert sie nach der Anzahl ihrer *wesentlichen Löcher*, wobei eine Menge ohne wesentliches Loch einfach zusammenhängend genannt wird. Ob ein Loch als wesentlich oder unwesentlich anzusehen ist, hängt allerdings ganz entscheidend von der Struktur des Loches sowie der Dimension n ab.

In *zwei Dimensionen* sind alle Löcher wesentlich und die Mathematik kennt die folgenden drei Charakterisierungen, die paarweise äquivalent sind und jeweils die Existenz von Löchern ausschließen:

1. (Komplementärmenge) Die Menge $\mathbb{R}^n \setminus D$ ist zusammenhängend.
2. (Nullhomologie) Die Windungszahl jeder geschlossenen Kurve \mathbf{c} in D bzgl. eines äußeren Punktes $\mathbf{x}_* \notin D$ ist Null. Siehe dazu die Vorlesung *Funktionentheorie*.
3. (Nullhomotopie) Jede geschlossene Kurve \mathbf{c} in D kann innerhalb von D solange stetig deformiert werden, bis sie auf einen Punkt zusammengeschrunpft ist.

In *drei oder mehr Dimensionen* wird es unübersichtlicher, da es nun auch unwesentliche Löcher geben kann und die obigen Bedingungen nicht mehr äquivalent sind. Für $n \geq 3$ ist eine Menge $D \subset \mathbb{R}^n$ genau dann einfach zusammenhängend, wenn sie die dritte Eigenschaft besitzt. Die ersten beiden Kriterien dürfen nicht mehr verwendet werden.



Abbildung Links: Zwei Beispiele für eine einfach zusammenhängende Menge $D \subset \mathbb{R}^2$ (grün), wobei die vier farbigen Kurven jeweils eine Nullhomotopie der orangenen Kurve illustrieren. Rechts: Zwei Beispiele für eine nicht einfach zusammenhängende Menge $D \subset \mathbb{R}^2$ (rot, mit einem bzw. zwei Löchern), wobei jede der blauen Kurven innerhalb von D nicht auf einen Punkt zusammengezogen werden kann. Außerdem verschwindet ihre Windungszahl nicht, sofern diese bzgl. des blau markierten Punktes ausgewertet wird.

Bemerkungen

1. Bei der Nullhomotopie sind nur Deformationen innerhalb von D erlaubt und Windungszahlen können nur für $n = 2$ sinnvoll definiert werden.
2. Konvexe Mengen sind immer einfach zusammenhängend. Dabei wird $D \subset \mathbb{R}^n$ konvex genannt, wenn die Implikation

$$\mathbf{x} \in D, \tilde{\mathbf{x}} \in D, t \in [0, 1] \quad \Rightarrow \quad (1-t)\mathbf{x} + t\tilde{\mathbf{x}} \in D$$

gilt, d.h. wenn zu je zwei Punkten aus D auch die entsprechende Verbindungsstrecke ganz zu D gehört. Der Ganzraum \mathbb{R}^n ist immer konvex und damit auch einfach zusammenhängend (für jedes n).

3. In 2D sind insbesondere auch *Einpunktlöcher* wesentlich, d.h. die *punktierte Ebene* $\mathbb{R}^2 \setminus \{\mathbf{0}\}$ ist zum Beispiel *nicht* einfach zusammenhängend. Solche Einpunktlöcher sind oftmals mit Singularitäten von Vektorfeldern verbunden.

4. Singularitäten in 3D sind zwar immer noch mit Einpunktlöcher verbunden, aber diese sind nicht mehr wesentlich und können nicht den einfachen Zusammenhang einer Menge zerstören. Insbesondere ist der *punktierte Raum* $\mathbb{R}^3 \setminus \{\mathbf{0}\}$ einfach zusammenhängend. Dasselbe gilt für $\mathbb{R}^3 \setminus \overline{B}_\rho(\mathbf{0})$, d.h. wenn das Loch nicht nur aus einem Punkt, sondern aus einer ganzen Kugel besteht. Die Mengen

$$\mathbb{R}^3 \setminus \{(x_1, x_2, x_3) : x_1^2 + x_2^2 = 1, x_3 = 0\}, \quad \mathbb{R}^3 \setminus \{(x_1, x_2, x_3) : x_1 = x_2 = 0\}$$

sind jedoch beide *nicht* einfach zusammenhängend, d.h. Löcher in Form einer Kreislinie oder einer Geraden sind in 3D wesentlich.

Hinweis: Im Rahmen dieser Vorlesung müssen Sie nur für $n = 2$ (aber nicht für $n \geq 3$) entscheiden können, ob eine vorgelegte Menge $D \subset \mathbb{R}^n$ einfach zusammenhängend ist oder nicht.

Theorem (Integrabilitätsbedingungen für Vektorfelder in nD) Ist $D \subset \mathbb{R}^n$ einfach zusammenhängend und $\mathbf{f} : D \rightarrow \mathbb{R}^n$ stetig differenzierbar, so ist \mathbf{f} genau dann Gradientenfeld, wenn $\mathbf{Jf}(\mathbf{x})$ für jedes $\mathbf{x} \in D$ eine symmetrische Matrix ist.

Beweis Den allgemeinen Beweis wollen wir an dieser Stelle nicht führen, aber für $n = 2$ können wir die Behauptung aus dem Satz von Stokes ableiten. Dies wird gleich genauer erklärt. \square

Bemerkung

1. Auf einfach zusammenhängenden Mengen sind also die notwendigen Bedingungen von oben auch hinreichend für die Existenz eines Potentials zu einem gegebenen Vektorfeld \mathbf{f} .
2. Die notwendigen Bedingungen, d.h die Symmetrieforderung an die Jacobi-Matrix, nennt man aus historischen Gründen die *Integrabilitätsbedingungen* an das Vektorfeld \mathbf{f} . Sie sind in der Regel leicht auszuwerten, da nur Ableitungen berechnet werden müssen.
3. Will man das Potential ψ zu einem gegebenen Vektorfeld \mathbf{f} berechnen, so kann man – wie im Beweis des Hauptsatzes beschrieben – Kurvenintegrale benutzen. In der Praxis versucht man jedoch meist, eine entsprechende partielle Differentialgleichung zu lösen (siehe den Ausblick unten). In konkreten Fällen können wir ψ auch direkt durch mehrere eindimensionale Integrationsschritte ableiten (siehe das nächste Beispiel).

Spezialfall $n = 2$ Sei $D \subset \mathbb{R}^2$ einfach zusammenhängend und sei $\mathbf{f} : D \rightarrow \mathbb{R}^2$ rotationsfrei. Sei außerdem \mathbf{c} eine beliebige geschlossene Kurve in D , wobei wir der Einfachheit halber annehmen, dass \mathbf{c} den Rand einer Menge $B \subset D$ parametrisiert und keine Doppelpunkte aufweist. Dann gilt

$$\int_{\mathbf{c}} \mathbf{f}(\mathbf{x}) \cdot d\mathbf{x} = \pm \int_{\partial B} \langle \mathbf{f}(\mathbf{x}), \boldsymbol{\tau}(\mathbf{x}) \rangle ds = \pm \int_B \operatorname{rot} \mathbf{f}(\mathbf{x}) d\mathbf{x} = 0$$

nach Auswertung des Integralsatzes von Stokes für B , wobei $\boldsymbol{\tau}$ der positiv orientierte Tangentialvektor auf ∂B ist. Die Existenz eines Potentials kann nun mit dem Hauptsatz über Kurvenintegrale begründet werden.

Beispiel Wir wollen untersuchen, für welche Werte des Parameters $p \in \mathbb{R}$ das dreidimensionale Vektorfeld

$$\mathbf{f}(\mathbf{x}) = \begin{pmatrix} \sin(x_3) + p \exp(x_2) \\ x_3 \exp(x_2) \\ \exp(x_2) + x_1 \cos(x_3) \end{pmatrix}$$

ein Gradientenfeld auf der einfach zusammenhängenden Menge $D = \mathbb{R}^3$ ist. Dazu prüfen wir die Integrabilitätsbedingungen, indem wir die Rotation von \mathbf{f} auswerten (siehe oben). Einfache Rechnungen liefern

$$\begin{aligned} \mathbf{rot} \mathbf{f}(\mathbf{x}) &= \begin{pmatrix} \partial_{x_2}(\exp(x_2) + x_1 \cos(x_3)) - \partial_{x_3}(x_3 \exp(x_2)) \\ \partial_{x_3}(\sin(x_3) + p \exp(x_2)) - \partial_{x_1}(\exp(x_2) + x_1 \cos(x_3)) \\ \partial_{x_1}(x_3 \exp(x_2)) - \partial_{x_2}(\sin(x_3) + p \exp(x_2)) \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ 0 \\ -p \exp(x_2) \end{pmatrix} \end{aligned}$$

und wir folgern, dass nur im Fall $p = 0$ ein Potential ψ für \mathbf{f} existiert. Für dieses Beispiel wollen wir mit $p = 0$ eine entsprechende Formel direkt, d.h. ohne Benutzung von Kurvenintegralen, ableiten. Durch Auswertung der ersten Gleichung $\partial_{x_1}\psi(\mathbf{x}) = f_1(\mathbf{x})$ und Anwendung des Hauptsatzes der Differential- und Integralrechnung — bzgl. der Variablen x_1 und bei festgehaltenen x_2, x_3 — schließen wir, dass

$$\psi(\mathbf{x}) = x_1 \sin(x_3) + \alpha(x_2, x_3)$$

gelten muss, wobei $\alpha(x_2, x_3)$ eine Integrationskonstante ist, die von x_2 und x_3 abhängen darf. Die zweite Gleichung $\partial_{x_2}\psi(\mathbf{x}) = f_2(\mathbf{x})$ kann nun als

$$\partial_{x_2}\alpha(x_2, x_3) = x_3 \exp(x_2)$$

geschrieben werden und liefert

$$\alpha(x_2, x_3) = x_3 \exp(x_2) + \beta(x_3), \quad \psi(\mathbf{x}) = x_1 \sin(x_3) + x_3 \exp(x_2) + \beta(x_3)$$

mit einer Integrationskonstanten $\beta(x_3)$. Aufgrund der verbleibenden dritten Gleichung $\partial_{x_3}\psi(\mathbf{x}) = f_3(\mathbf{x})$ erhalten wir

$$\partial_{x_2}\beta(x_3) = 0$$

und insgesamt

$$\psi(\mathbf{x}) = x_1 \sin(x_3) + x_3 \exp(x_2) + \gamma,$$

wobei die finale Integrationskonstante γ nicht von x_1, x_2 oder x_3 abhängt.

Ausblick: Poisson-Gleichung für Potentiale* Die Gleichung $\mathbf{grad} \psi(\mathbf{x}) = \mathbf{f}(\mathbf{x})$ impliziert

$$\Delta\psi(\mathbf{x}) = \operatorname{div} \mathbf{grad} \psi(\mathbf{x}) = \operatorname{div} \mathbf{f}(\mathbf{x})$$

wobei $\Delta = \partial_{x_1}^2 + \dots + \partial_{x_n}^2$ der *Laplace-Operator* ist und die rechte Seite aus \mathbf{f} berechnet werden kann. Diese partielle Differentialgleichung kodiert das sogenannte Poisson-Problem und kann benutzt werden, um das Potential ψ aus \mathbf{f} zu berechnen, wobei dann noch geeignete Randbedingungen an ψ auf ∂D gestellt werden müssen. Diese Umformulierung ist vor allem aus numerischer Sicht nützlich, da die Differentialgleichung sehr effektiv auf dem Computer gelöst werden kann.

Ausblick: Helmholtz-Zerlegung von Vektorfelder* Man kann — zumindest in 2D und 3D — jedes Vektorfeld $\mathbf{f} : D \rightarrow \mathbb{R}^n$ via

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}_D(\mathbf{x}) + \mathbf{f}_R(\mathbf{x})$$

in einen *divergenzfreien* Anteil \mathbf{f}_D und einen *rotationsfreien* Anteil \mathbf{f}_R zerlegen, wobei auf einer einfach zusammenhängenden Menge D der rotationsfreie Anteil auch als Gradient eines Potentials geschrieben werden kann. Diese Zerlegung ist im Allgemeinen nicht eindeutig, aber sehr wichtig in der Hydro- und Elektrodynamik.

3.3 Frenet-Theorie und Geometrie von Kurven

Ziel Wir wiederholen das Konzept *Reparametrisierung von Kurven* und stellen außerdem die Grundlagen der sogenannten Frenet-Theorie zusammen.

Planare Kurven

Setting Wir betrachten in diesem Abschnitt parametrisierte Kurven $\mathbf{c} : I \rightarrow \mathbb{R}^2$, die mindestens zweimal stetig differenzierbar sind. Insbesondere wollen wir diesmal keine Kurven zulassen, die nur stückweise stetig differenzierbar sind und deshalb Ecken aufweisen. Die meisten Konzepte können aber auf solche Fälle verallgemeinert werden.

Erinnerung

1. Eine Kurve $\tilde{\mathbf{c}} : \tilde{I} \rightarrow \mathbb{R}^2$ wird Reparametrisierung von $\mathbf{c} : I \rightarrow \mathbb{R}^2$ genannt, wenn es monoton wachsende und bijektive Funktion $h : \tilde{I} \rightarrow I$ gibt, sodass

$$\tilde{\mathbf{c}}(\tilde{t}) = \mathbf{c}(t) \quad \text{mit} \quad t = h(\tilde{t})$$

für alle $t \in \tilde{I}$ gilt. Insbesondere gilt immer auch $\tilde{t} = \tilde{h}(t)$, wobei $\tilde{h} : I \rightarrow \tilde{I}$ die Umkehrfunktion zu h ist.

2. Für jede Kurve $\mathbf{c} : [t_a, t_e] \rightarrow \mathbb{R}^2$ gibt es eine ausgezeichnete Parametrisierung, nämlich die nach dem Bogenlängenparameter. Dieser wird üblicherweise mit s (und nicht mit \tilde{t}) bezeichnet und durch

$$s = l(t) = \int_{t_a}^t \ell(\tau) \, d\tau, \quad \ell(t) = |\dot{\mathbf{c}}(t)|$$

eingeführt, sodass symbolisch $ds = \ell(t) dt$ gilt. Die Bogenlängenfunktion l beschreibt, wie s von t abhängt, und bildet das Intervall I in das Intervall $J = [0, L(\mathbf{c})]$ ab, wobei $L(\mathbf{c}) = l(t_e)$ die Länge der Kurve bezeichnet. Das Intervall J bzw. die Funktion l übernehmen damit die Rolle von \tilde{I} bzw. $\tilde{h} = h^{-1}$. Die Umkehrfunktion $l^{-1} : J \rightarrow I$ gibt an, wie der Kurvenparameter t vom Bogenlängenparameter s abhängt, und entspricht der Funktion h . Wir schreiben auch $\mathbf{d}(s)$ statt $\tilde{\mathbf{c}}(\tilde{t})$.

Beispiele

1. Die Kurve

$$\mathbf{c}(t) = \begin{pmatrix} t - \sin(t) \\ 1 - \cos(t) \end{pmatrix}$$

wird *Zykloide* genannt und beschreibt, dass ein Kreis auf einer Geraden abgerollt wird (siehe WIKIPEDIA). Betrachten wir \mathbf{c} über dem Parametrisierungsintervall $I = [0, \pi]$, so überlegt man sich leicht, dass sowohl $x_1 = c_1(t)$ als auch $x_2 = c_2(t)$ monoton von der Zeit t abhängen und die Kurve daher auch nach x_1 oder x_2 parametrisiert werden kann. Außerdem kann man die Kurve auch durch ihre Bogenlänge parametrisieren.

- (a) Die Möglichkeit der Parametrisierung nach x_2 wird mit der allgemeinen bzw. abstrakten Notation durch

$$\tilde{\mathbf{c}}(\tilde{t}) = \begin{pmatrix} \arccos(1 - \tilde{t}) - \sqrt{2\tilde{t} - \tilde{t}^2} \\ \tilde{t} \end{pmatrix}, \quad \tilde{t} \in \tilde{I} = [0, 2]$$

und

$$h(\tilde{t}) = \arccos(1 - \tilde{t}), \quad \tilde{h}(t) = 1 - \cos(t)$$

beschrieben, denn dann gilt ja $x_2 = \tilde{c}_2(\tilde{t}) = \tilde{t}$, d.h. der neue Parameter \tilde{t} ist gerade die x_2 -Komponente entlang der Kurve. In den Anwendungswissenschaften würde man dies eher als

$$x_1 = \arccos(1 - x_2) + \sqrt{2x_2 - x_2^2}$$

schreiben, wobei dann die Formeln

$$t = \arccos(1 - x_2), \quad x_1 - t = -\sqrt{1 - \cos^2(t)} = \sqrt{1 - (1 - x_2)^2}$$

zu Grunde liegen. Beachte, dass diese Parametrisierung nur für $t \in I$, aber zum Beispiel nicht für $t \in \mathbb{R}$ verwendet werden kann. Denn im Allgemeinen gibt es für jeden Wert von $x_2 \in [0, 2]$ mehrere Zeiten t mit $x_2 = c_2(t)$.

- (b) Es gibt keine einfache Formel für die Parametrisierung nach x_1 , eben weil man die Gleichung $x_1 = t - \sin(t)$ nicht explizit nach t auflösen kann. Es ist aber nützlich zu wissen, dass es eine solche Parametrisierung im Prinzip gibt.
- (c) Die Bogenlängenparametrisierung kann in diesem Beispiel explizit berechnet werden. Wegen $\sin(\frac{1}{2}t) > 0$ für $t \in I$ und mit

$$\ell(t) = |\dot{\mathbf{c}}(t)| = \sqrt{2 - 2 \cos(t)} = \sqrt{4 \sin^2(\frac{1}{2}t)} = 2 \sin(\frac{1}{2}t)$$

ergibt sich

$$s = l(t) = \int_0^t \ell(\tau) d\tau = \int_0^t 2 \sin(\frac{1}{2}\tau) d\tau = 4 - 4 \cos(\frac{1}{2}t)$$

bzw.

$$t = l^{-1}(s) = 2 \arccos\left(1 - \frac{1}{4}s\right).$$

Dabei ist l^{-1} , die Umkehrfunktion von l , auf dem Intervall

$$J = [l(0), l(\pi)] = [0, 4]$$

definiert, da 4 gerade die Länge der Kurve $\mathbf{c} : I \rightarrow \mathbb{R}^2$ ist. Durch Einsetzen und Ausnutzen von Additionstheoremen erhalten wir

$$\mathbf{d}(s) = \begin{pmatrix} 2 \arccos\left(1 - \frac{1}{4}s\right) - \left(\frac{1}{8}s - \frac{1}{2}\right)\sqrt{8s - s^2} \\ s - \frac{1}{8}s^2 \end{pmatrix},$$

als Bogenlängenparametrisierung des betrachteten Zykloidenastes, wobei

$$|\mathbf{d}'(s)| = \left| \frac{d}{ds} \mathbf{d}(s) \right| = 1.$$

nach Konstruktion sichergestellt ist bzw. mithilfe direkter Rechnungen verifiziert werden kann. Insbesondere gilt

$$\mathbf{c}(t) = \mathbf{d}(s),$$

sofern s und t durch $s = l(t)$ bzw. $t = l^{-1}(s)$ gekoppelt sind. Oder anders gesagt: Die parametrisierten Kurven \mathbf{c} und \mathbf{d} beschreiben jeweils dieselbe geometrische Punktmenge, nämlich $\text{im}(\mathbf{c}) = \text{im}(\mathbf{d})$, nur einmal durch den Parameter t und das andere Mal mittels des Bogenlängenparameters s . Beachte auch, dass die Bogenlängenparametrisierung der Zykloide auf größeren Intervallen existiert, aber für $t > \pi$ müssen die Formeln entsprechend angepasst werden, da dann nur $l(t) = 2 \left| \sin\left(\frac{1}{2}t\right) \right|$ gilt.

2. Die *Neilsche Parabel* wird durch

$$\mathbf{c}(t) = \begin{pmatrix} t^2 \\ t^3 \end{pmatrix}, \quad t \geq 0$$

definiert, wobei wir $I = [0, \infty)$ als das Parametrisierungsintervall betrachten wollen und

$$\dot{\mathbf{c}}(t) = \begin{pmatrix} 2t \\ 3t^2 \end{pmatrix}, \quad \ell(t) = |\dot{\mathbf{c}}(t)| = t \sqrt{4 + 9t^2}$$

durch einfache Rechnungen finden. Um das entsprechende Bogenlängenintegral

$$l(t) = \int_0^t \ell(\tau) \, d\tau = \int_0^t \tau \sqrt{4 + 9\tau^2} \, d\tau.$$

auszuwerten, substituieren wir

$$\sigma = 4 + 9\tau^2, \quad d\sigma = 18\tau \, d\tau$$

und erhalten

$$l(t) = \frac{1}{18} \int_4^{4+9t^2} \sqrt{\sigma} \, d\sigma = \frac{1}{27} \left[\sigma^{3/2} \right]_{\sigma=4}^{\sigma=4+9t^2} = \frac{(4 + 9t^2)^{3/2} - 8}{27}.$$

Dies ist gerade die Formel für den Bogenlängenparameter

$$s = l(t)$$

und einfache Rechnungen liefern die Umkehrformel

$$t = l^{-1}(s) = \sqrt{\frac{(27s + 8)^{2/3} - 4}{9}}$$

für $s \in J = [0, \infty)$. Durch Einsetzen ergibt sich

$$\mathbf{d}(s) = \left(\begin{array}{c} \frac{(27s + 8)^{2/3} - 4}{9} \\ \left(\frac{(27s + 8)^{2/3} - 4}{9} \right)^{3/2} \end{array} \right)$$

für die Bogenlängenparametrisierung der Neilschen Parabel. Insbesondere gilt wieder $\mathbf{c}(t) = \mathbf{d}(s)$ mit $s = l(t)$ bzw. $t = h(s)$ sowie $|\mathbf{d}'(s)| = 1$ für alle $s \in J$.

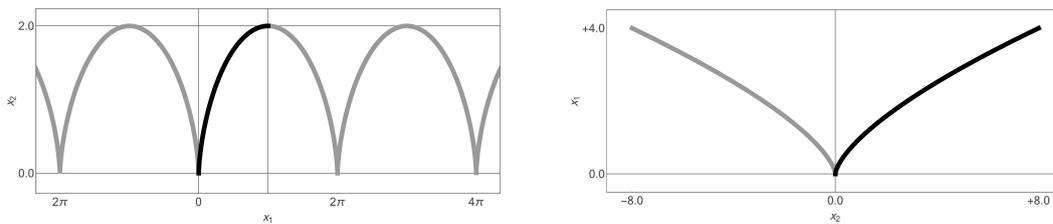


Abbildung Die Zyklode (links) und die Neilsche Parabel (rechts), wobei der in den Beispielen gerechnete Ast jeweils schwarz dargestellt ist und beide Male in $\mathbf{c}(0) = \mathbf{0}$ beginnt. Beachte, dass im rechten Bild die x_2 -Achse waagrecht liegt.

Bemerkung

1. Die Bogenlängenparametrisierung kann oftmals *nicht* in geschlossener Form angegeben werden (weil entweder das Integral für $l(t)$ nicht explizit berechnet oder die entstehende Funktion nicht explizit invertiert werden kann). Es ist aber ausgesprochen wichtig, dass es immer eine solche Parametrisierung gibt. Viele mathematische Theoreme über Kurven werden oftmals nur unter der Annahme einer Bogenlängenparametrisierung formuliert bzw. bewiesen, da dann viele Argumente deutlich einfacher werden.
2. Wir haben die Formeln für den Bogenlängenparameter s nur für Kurven mit ausgezeichnetem Anfangspunkt eingeführt. Man kann ein analoges Konzept für unendliche Kurven einführen, muss dann aber einen Referenzwert t_* bzw. einen Referenzpunkt $\mathbf{c}(t_*)$ auswählen und außerdem mit negativen Bogenlängen arbeiten. Bei der Zyklode und der Neilschen Parabel können wir zum Beispiel $t_* = 0$ mit $\mathbf{c}(t_*) = \mathbf{0}$ wählen und jeden Punkt $\mathbf{c}(t)$ mit $t < 0$ eindeutig durch Angabe einer negativen Bogenlänge charakterisieren.
3. Ganz allgemein spricht man von einer *Bogenlängenparametrisierung*, wenn die Ableitung nach dem Kurvenparameter immer die Länge 1 besitzt. Gilt also zum Beispiel $|\dot{\mathbf{c}}(t)| = 1$ für alle $t \in I$, so ist \mathbf{c} in diesem Sinne bereits nach Bogenlänge parametrisiert, denn es gilt $ds = dt$ und damit $s = t + c$ für eine geeignete Konstante c .
4. Im praktischen Leben besteht die Kunst darin, eine möglichst natürliche oder einfache Parametrisierung zu wählen, die für das jeweilige Problem am besten geeignet ist.

Kurven in der Ebene

Frenetsches Zweibein Für jede reguläre Kurve $\mathbf{c} : I \rightarrow \mathbb{R}^2$ gilt

$$\ell(t) = \sqrt{\dot{c}_1^2(t) + \dot{c}_2^2(t)} > 0,$$

sodass wir für jeden Parameterwert $t \in I$ den Frenetschen Tangentialvektor $\mathbf{b}_1(t)$ sowie den Frenetschen Normalenvektor $\mathbf{b}_2(t)$ durch

$$\mathbf{b}_1(t) = \frac{1}{\ell(t)} \begin{pmatrix} +\dot{c}_1(t) \\ +\dot{c}_2(t) \end{pmatrix}, \quad \mathbf{b}_2(t) = \frac{1}{\ell(t)} \begin{pmatrix} -\dot{c}_2(t) \\ +\dot{c}_1(t) \end{pmatrix}$$

definiert können. Insbesondere stehen diese zwei Vektoren immer senkrecht aufeinander, besitzen jeweils die Länge 1 und beschreiben eine orthonormale Basis des \mathbb{R}^2 , die entlang der parametrisierten Kurve \mathbf{c} mitbewegt wird (siehe Bild).

Alternative Notation In der Literatur finden sich oftmals andere Schreibweisen. Zum Beispiel wird $\mathbf{b}_1(t)$ auch mit $\mathbf{t}(t)$ und $\mathbf{b}_2(t)$ mit $\mathbf{n}(t)$ bezeichnet. Wie immer gilt: Nicht die Notation ist wichtig, sondern das zu Grunde liegende Konzept.

Bemerkung

- Wir hatten im Abschnitt über die Integralsätze ein leicht anderes Konzept von normierten Tangential- und Normalenvektoren eingeführt. Diese sind auf dem Rand ∂D einer Menge $D \subset \mathbb{R}^2$ definiert, wobei der normierte Normalenvektor $\boldsymbol{\nu}(\mathbf{x})$ in jedem Punkt $\mathbf{x} \in \partial D$ nach außen zeigt und der normierte Tangentialvektor $\boldsymbol{\tau}(\mathbf{x})$ dazu positiv orientiert ist. Beide Konzepte haben sehr viel miteinander zu tun. Insbesondere gilt

$$\boldsymbol{\tau}(\mathbf{c}(t)) = \pm \mathbf{b}_1(t), \quad \boldsymbol{\nu}(\mathbf{c}(t)) = \mp \mathbf{b}_2(t),$$

sofern \mathbf{c} den Rand der Menge D parametrisiert. Ein erster wichtiger Unterschied ist aber, dass $\mathbf{b}_1(t)$ und $\mathbf{b}_2(t)$ allein mit Hilfe der parametrisierten Kurve und ohne Bezug auf D eingeführt werden. Insbesondere macht die Unterscheidung zwischen *innen* und *außen* aus Sicht der Kurve keinen Sinn, d.h. allein durch Betrachtung von \mathbf{c} können wir nicht entscheiden, ob wir in den obigen Formeln nun jeweils $+$ oder $-$ wählen müssen. Ein zweiter wichtiger Unterschied betrifft die Verallgemeinerung auf höhere Dimensionen (siehe unten).

- Singuläre (bzw. entartete) Kurvenpunkte mit $\dot{\mathbf{c}}(t) = \mathbf{0}$ sind aus Sicht der Frenet-Theorie problematisch, da dort das Zweibein nicht eindeutig definiert werden kann.

Krümmung und Krümmungskreis Eine wichtige Rolle in der Theorie planarer Kurven spielt die skalare und vorzeichenbehaftete Krümmung

$$\kappa(t) = \frac{\dot{c}_1(t) \ddot{c}_2(t) - \ddot{c}_1(t) \dot{c}_2(t)}{\ell^3(t)}.$$

Diese beschreibt für jedes $t \in I$, wie stark sich gerade das Frenetsche Zweibein dreht. Dabei meint ein negatives bzw. positives Vorzeichen, dass sich das Zweibein mit bzw.

entgegen dem Uhrzeigersinn dreht. Man kann dies auch mit Hilfe der sogenannten *Krümmungskreise* beschreiben. Für jeden Punkt $\mathbf{c}(t)$ mit $\kappa(t) \neq 0$ sind durch

$$\varrho(t) = \frac{1}{|\kappa(t)|}, \quad \mathbf{m}(t) = \mathbf{c}(t) + \frac{1}{\kappa(t)} \mathbf{b}_2(t)$$

der entsprechende Radius $\varrho(t)$ sowie der entsprechende Mittelpunkt $\mathbf{m}(t)$ gegeben, wobei sich der Krümmungskreis im Punkt $\mathbf{c}(t)$ an die Kurve anschmiegt (siehe Bild). Für $\kappa(t) = 0$ fällt der Krümmungskreis mit der Tangentialgerade zusammen, die man als Kreis mit Radius ∞ um einen unendlich fernen Mittelpunkt interpretieren kann. Zu einer gegebenen Kurve \mathbf{c} wird durch \mathbf{m} eine weitere Kurve definiert, die man die *Evolute* von \mathbf{c} nennt (siehe Hausaufgaben).

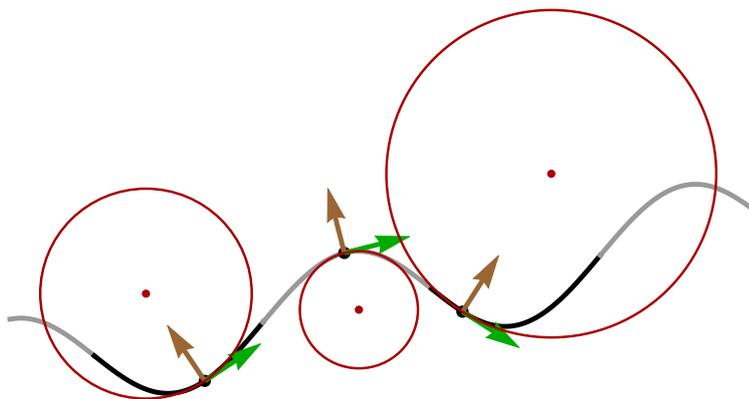


Abbildung Planare Kurve mit Frenetschem Zweibein (grün und braun) und Krümmungskreisen (rot), wobei der Wert von κ in den grauen bzw. schwarzen Kurvensegmenten negativ bzw. positiv ist. Die Mittelpunkte aller Krümmungskreise bilden zusammen die Evolute.

Frenet-Serret-Gleichungen Das Zweibein einer regulären planaren Kurve genügt den Differentialgleichungen von Frenet-Serret, d.h. es gilt

$$\dot{\mathbf{b}}_1(t) = +\ell(t) \kappa(t) \mathbf{b}_2(t), \quad \dot{\mathbf{b}}_2(t) = -\ell(t) \kappa(t) \mathbf{b}_1(t)$$

für alle $t \in I$. Wir können die Gültigkeit dieser Formeln einfach nachrechnen und sie außerdem benutzen, um Kurven mit vorgegebener Krümmung als Lösung geeigneter Anfangswertprobleme zu konstruieren.

Beispiele

1. Durch

$$\mathbf{c}(t) = \begin{pmatrix} \varrho \cos(\omega t) \\ \varrho \sin(\omega t) \end{pmatrix}$$

wird die Kreislinie mit vorgeschriebenem Radius ϱ parametrisiert, wobei die *Winkelgeschwindigkeit* ω negativ oder positiv sein kann (mit oder entgegen dem Uhrzeigersinn). Durch einfache Rechnungen (und wegen $\omega/\sqrt{\omega^2} = \text{sgn}(\omega)$) erhalten wir

$$\mathbf{b}_1(t) = \text{sgn}(\omega) \begin{pmatrix} -\sin(\omega t) \\ +\cos(\omega t) \end{pmatrix}, \quad \mathbf{b}_2(t) = \text{sgn}(\omega) \begin{pmatrix} -\cos(\omega t) \\ -\sin(\omega t) \end{pmatrix}$$

sowie

$$\kappa(t) = \frac{\operatorname{sgn}(\omega)}{\varrho}, \quad \varrho(t) = \varrho, \quad \mathbf{m}(t) = \mathbf{0}.$$

Insbesondere ist in diesem Beispiel die Krümmung konstant mit Betrag $1/\varrho$ (aber das Vorzeichen hängt vom Umlaufsinn ab) und der Krümmungskreis fällt immer mit dem Bild der parametrisierten Kurve zusammen. Die Evolute ist also eine konstante Kurve und damit entartet.

Bemerkung: Die Formel beschreibt genau dann eine Parametrisierung nach Bogenlänge, wenn $|\dot{\mathbf{c}}(t)| = 1$ für alle t gilt, d.h. im Fall von $\omega = \pm 1/\varrho$.

2. Für die Neilsche Parabel (siehe oben) mit $t > 0$ erhalten wir nach einfachen Rechnungen

$$\mathbf{b}_1(t) = \frac{1}{t(4+9t^2)^{1/2}} \begin{pmatrix} +2t \\ +3t^2 \end{pmatrix}, \quad \mathbf{b}_2(t) = \frac{1}{t(4+9t^2)^{1/2}} \begin{pmatrix} -3t^2 \\ +2t \end{pmatrix}$$

sowie

$$\kappa(t) = \frac{6}{t(4+9t^2)^{3/2}}, \quad \mathbf{m}(t) = \begin{pmatrix} t^2 \\ t^3 \end{pmatrix} + \frac{4+9t^2}{6} \begin{pmatrix} -3t^2 \\ +2t \end{pmatrix} = \begin{pmatrix} -t^2 - \frac{9}{2}t^4 \\ \frac{4}{3}t + 4t^3 \end{pmatrix}.$$

Beachte, dass einige Formeln im Limes $t \searrow 0$ singulär werden. Das liegt daran, dass $\mathbf{c}(0)$ wegen $\dot{\mathbf{c}}(0) = \mathbf{0}$ ein Entartungspunkt der Frenet-Theorie ist.

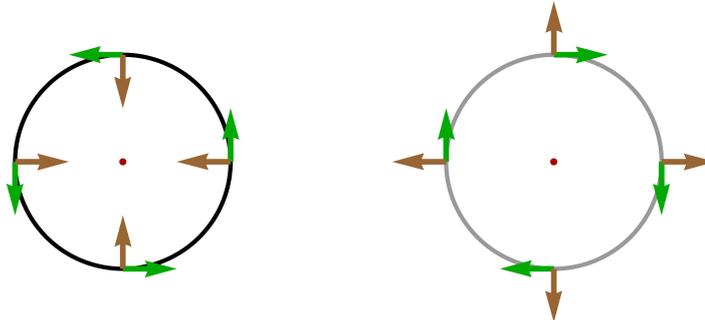


Abbildung Frenetsches Zweibein \mathbf{b}_1 (grün) und \mathbf{b}_2 (braun) auf einer Kreislinie, wobei die linke bzw. rechte Abbildung einer positiven bzw. negativen Krümmung entspricht, d.h. einem Umlauf entgegen bzw. mit dem Uhrzeigersinn. Beachte, dass hier der Krümmungskreis nicht vom Kurvenparameter t abhängt und mit der Kreislinie zusammenfällt.

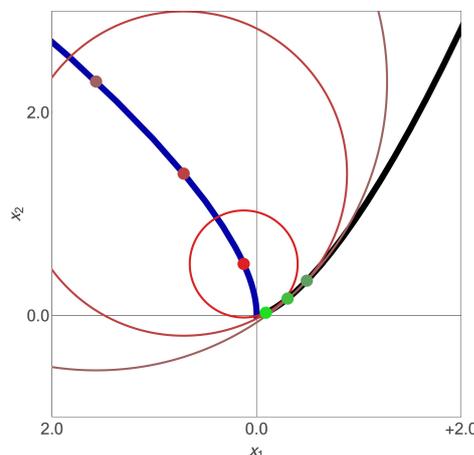


Abbildung Die Evolute (blau) der Neilschen Parabel (schwarz), wobei für drei Werte von t die Punkte $\mathbf{m}(t)$ (rot) und $\mathbf{c}(t)$ (grün) sowie der entsprechende Krümmungskreis dargestellt sind.

Theorem* (**Kurvenintegral der Krümmung**) Sei $\mathbf{c} : [t_a, t_e] \rightarrow \mathbb{R}^2$ eine stetig differenzierbare, reguläre und geschlossene Kurve ohne Doppelpunkt. Dann gilt

$$\int_{\mathbf{c}} \kappa \, ds = \pm 2\pi,$$

wobei das Vorzeichen negativ bzw. positiv zu wählen ist, sofern die Kurve mit oder entgegen dem Uhrzeigersinn durchlaufen wird.

Beweis* Ohne Beschränkung der Allgemeinheit können wir annehmen, dass die Kurve nach Bogenlänge parametrisiert ist, d.h. dass $\ell(t) = |\dot{\mathbf{c}}(t)| = 1$ für alle t . In diesem Fall gibt es eine Winkelfunktion $\theta : [t_a, t_e] \rightarrow \mathbb{R}$, sodass

$$\dot{\mathbf{c}}(t) = \begin{pmatrix} +\cos(\theta(t)) \\ +\sin(\theta(t)) \end{pmatrix}, \quad \ddot{\mathbf{c}}(t) = \dot{\theta}(t) \begin{pmatrix} -\sin(\theta(t)) \\ +\cos(\theta(t)) \end{pmatrix}$$

und die Annahmen an die Geometrie der Kurve implizieren, dass θ entweder monoton wachsend mit $\theta(t_e) = \theta(t_a) + 2\pi$ oder monoton fallend mit $\theta(t_e) = \theta(t_a) - 2\pi$ ist. Durch einfache Rechnungen zeigen wir

$$\kappa(t) = \dot{\theta}(t)$$

und erhalten via

$$\int_{\mathbf{c}} \kappa \, ds = \int_{t_a}^{t_e} \kappa(t) |\dot{\mathbf{c}}(t)| \, dt = \int_{t_a}^{t_e} \dot{\theta}(t) \, dt = \theta(t_e) - \theta(t_a)$$

das gewünschte Ergebnis. □

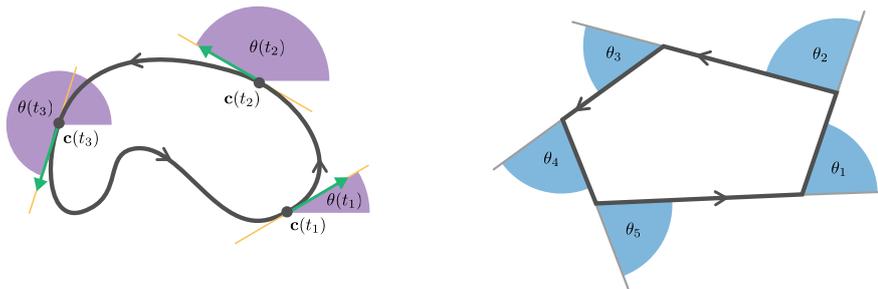


Abbildung *Links:* Die Krümmung κ einer stetig differenzierbaren planaren Kurve kann auch als infinitesimale Änderung des Tangentialwinkels θ entlang der Kurve interpretiert werden und muss sich daher bei einer geschlossenen Kurve ohne Doppelpunkt zu $\pm 2\pi$ integrieren. Bei Kurven mit einem oder mehreren Doppelpunkten wird man immer ein ganzzahliges Vielfaches von 2π erhalten, nämlich die sogenannte Umlaufzahl der Kurve. *Rechts:* Ein analoges Resultat für polygonale Kurven ohne Doppelpunkt besagt, dass sich die Winkelsprünge θ_j in den Eckpunkten immer zu $\pm 2\pi$ addieren (und zwar unabhängig von der Anzahl der Ecken).

Raumkurven

Setting Wir betrachten nun parametrisierte Kurven mit $n = 3$, d.h. Abbildungen $\mathbf{c} : I \rightarrow \mathbb{R}^3$. Wir wollen dabei voraussetzen, dass \mathbf{c} dreimal stetig differenzierbar ist und dass außerdem $\dot{\mathbf{c}}(t)$ und $\ddot{\mathbf{c}}(t)$ für jedes $t \in I$ linear unabhängig sind. Dies impliziert insbesondere $\dot{\mathbf{c}}(t) \neq \mathbf{0}$ für alle $t \in I$ und damit die Regularität der Kurve.

Bemerkung Die Formeln und Resultate bzgl. der *Länge*, der *Bogenlänge* und von *Reparametrisierungen* gelten ganz analog zum Fall $n = 2$ bzw. $n > 3$. Wir diskutieren im Folgenden die speziellen geometrischen Eigenschaften, die sich aus $n = 3$ ergeben.

Frenetsches Dreibein und Frenet-Serret-Gleichungen Zu jeder Raumkurve gibt es ein Dreibein — das als eine entlang der Kurve mitbewegte orthonormale Basis des \mathbb{R}^3 interpretiert werden kann — sowie im Allgemeinen zwei skalare Krümmungsgrößen. Es existieren mehrere sinnvolle Varianten von Dreibeinen und auch verschiedene Möglichkeiten, diese zu definieren bzw. einzuführen. Wir beschränken uns hier auf das Frenetsche Dreibein und starten diesmal mit der entsprechenden Differentialgleichung: Für jedes $t \in I$ gibt zwei Teilkrümmungen $\kappa_1(t)$ und $\kappa_2(t)$ sowie drei Vektoren $\mathbf{b}_1(t)$, $\mathbf{b}_2(t)$ und $\mathbf{b}_3(t)$, sodass die 3D-Variante der Gleichungen von Frenet-Serret, d.h. die Formeln

$$\begin{aligned}\dot{\mathbf{b}}_1(t) &= +\ell(t) \kappa_1(t) \mathbf{b}_2(t) \\ \dot{\mathbf{b}}_2(t) &= -\ell(t) \kappa_1(t) \mathbf{b}_1(t) + \ell(t) \kappa_2(t) \mathbf{b}_3(t) \\ \dot{\mathbf{b}}_3(t) &= -\ell(t) \kappa_2(t) \mathbf{b}_2(t)\end{aligned}$$

erfüllt sind, wobei wieder $\ell(t) = |\dot{\mathbf{c}}(t)|$ gilt und $\dot{\mathbf{b}}_j(t) = \frac{d}{dt} \mathbf{b}_j(t)$ die zeitliche Änderung von $\mathbf{b}_j(t)$ beschreibt.

Bezeichnung und alternative Notation Die drei Vektoren $\mathbf{b}_1(t)$ bzw. $\mathbf{b}_2(t)$ bzw. $\mathbf{b}_3(t)$ werden Frenetscher Tangentenvektor bzw. Hauptnormalenvektor bzw. Binormalenvektor genannt und in der Literatur auch mit $\mathbf{t}(t)$ bzw. $\mathbf{h}(t)$ bzw. $\mathbf{b}(t)$ bezeichnet. Die Frenetschen Teilkrümmungen $\kappa_1(t)$ bzw. $\kappa_2(t)$ werden dann als $\kappa(t)$ bzw. $\tau(t)$ geschrieben und Krümmung bzw. Torsion oder Windung genannt.

Bemerkungen

1. Es gilt stets

$$\langle \mathbf{b}_i(t), \mathbf{b}_j(t) \rangle = \delta_{ij},$$

wobei δ_{ij} das Kronecker-Delta ist und $\langle \cdot, \cdot \rangle$ das Skalarprodukt im \mathbb{R}^3 bezeichnet. Oder anders gesagt: Das Frenetsche Dreibein liefert wirklich für jedes $t \in I$ eine orthonormale Basis des \mathbb{R}^3 .

2. Es gilt auch stets

$$\mathbf{b}_3(t) = \mathbf{b}_1(t) \times \mathbf{b}_2(t), \quad \mathbf{b}_1(t) = \mathbf{b}_2(t) \times \mathbf{b}_3(t), \quad \mathbf{b}_2(t) = \mathbf{b}_3(t) \times \mathbf{b}_1(t)$$

sowie

$$\det \begin{pmatrix} | & | & | \\ \mathbf{b}_1(t) & \mathbf{b}_2(t) & \mathbf{b}_3(t) \\ | & | & | \end{pmatrix} = 1,$$

wobei \times das Kreuzprodukt im \mathbb{R}^3 bezeichnet und die Determinantengleichung in jedem Punkt sicherstellt, dass das Frenetsche Dreibein positiv orientiert ist und dass die *Rechte-Hand-Regel* im Sinne von

$$\mathbf{b}_1(t) = \text{„Daumen“}, \quad \mathbf{b}_2(t) = \text{„Zeigefinger“}, \quad \mathbf{b}_3(t) = \text{„Mittelfinger“}$$

erfüllt ist.

- Der erste Frenet-Vektor $\mathbf{b}_1(t)$ ist immer proportional zum Geschwindigkeitsvektor $\dot{\mathbf{c}}(t)$, der zweite Frenet-Vektor $\mathbf{b}_2(t)$ kann als Linearkombination von $\dot{\mathbf{c}}(t)$ und $\ddot{\mathbf{c}}(t)$ dargestellt werden.
- Analog zu den Krümmungskreisen kann man in 3D Krümmungskugeln einführen, aber die entsprechenden Formeln spielen in dieser Vorlesung keine Rolle.

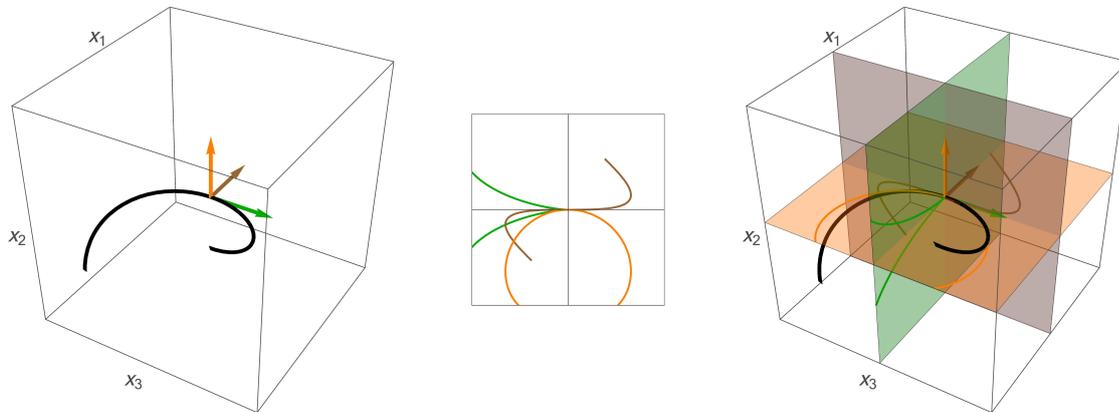


Abbildung Links: Frenetsches Dreibein in einem Punkt einer Raumkurve (schwarz, Teil einer Schraubenlinie), wobei $\mathbf{b}_1(t)$ bzw. $\mathbf{b}_2(t)$ bzw. $\mathbf{b}_3(t)$ durch den grünen bzw. braunen bzw. orangen Pfeil repräsentiert ist. Rechts: Die Normal- bzw. Streck- bzw. Schmiegebene, jeweils in der Farbe des senkrechten Vektors dargestellt, wobei die Ebenen im betrachteten Punkt achsenparallel liegen. Mitte: Projektionen der Raumkurve in die drei Ebenen.

Lokale Ebenen Man kann mit Hilfe des Dreibeins zu jedem Punkt $\mathbf{c}(t)$ in natürlicher Weise die folgenden affinen Ebenen einführen:

<u>Ebene</u>	<u>aufgespannt durch</u>	<u>senkrecht zu</u>
<u>Schmiegebene</u>	$\mathbf{b}_1(t), \mathbf{b}_2(t)$	$\mathbf{b}_3(t)$
<u>Normalebene</u>	$\mathbf{b}_2(t), \mathbf{b}_3(t)$	$\mathbf{b}_1(t)$
<u>Streckebene</u>	$\mathbf{b}_3(t), \mathbf{b}_1(t)$	$\mathbf{b}_2(t)$

Zum Beispiel besitzt die Schmiegebene im Kurvenpunkt $\mathbf{c}(t)$ die Punkt-Richtungs-darstellung

$$\mathbf{x} = \mathbf{c}(t) + \lambda_1 \mathbf{b}_1(t) + \lambda_2 \mathbf{b}_2(t),$$

wobei λ_1 und λ_2 die freien Parameter sind. Die entsprechende Tangentialgerade besitzt die Parameterdarstellung

$$\mathbf{x} = \mathbf{c}(t) + \lambda \mathbf{b}_1(t)$$

und liegt sowohl in der Schmiege- als auch in der Streckebene.

Beispiele

- Die Kurve

$$\mathbf{c}(t) = \begin{pmatrix} \varrho \cos(\omega t) \\ \varrho \sin(\omega t) \\ \eta \omega t \end{pmatrix}$$

beschreibt eine Schraubenlinie oder Helix mit Radius $\varrho > 0$, *Ganghöhe* $\eta \in \mathbb{R}$ und Winkelgeschwindigkeit ω , die hier jetzt positiv sein soll. Für diese wichtige Kurve impliziert das nachfolgende Theorem die Formeln

$$\begin{aligned}\mathbf{b}_1(t) &= \frac{1}{\sqrt{\varrho^2 + \eta^2}} \begin{pmatrix} -\varrho \sin(\omega t) \\ +\varrho \cos(\omega t) \\ \eta \end{pmatrix}, \\ \mathbf{b}_2(t) &= \begin{pmatrix} -\cos(\omega t) \\ -\sin(\omega t) \\ 0 \end{pmatrix}, \\ \mathbf{b}_3(t) &= \frac{1}{\sqrt{\varrho^2 + \eta^2}} \begin{pmatrix} +\eta \sin(\omega t) \\ -\eta \cos(\omega t) \\ \varrho \end{pmatrix}\end{aligned}$$

sowie

$$\kappa_1(t) = \frac{\varrho}{\varrho^2 + \eta^2}, \quad \kappa_2(t) = \frac{\eta}{\varrho^2 + \eta^2},$$

d.h. es handelt sich um eine *Kurve mit konstanter Krümmung und konstanter Torsion*.

Bemerkung: Es gilt $\ell(t) = |\dot{\mathbf{c}}(t)| = \sqrt{\omega^2} \sqrt{\varrho^2 + \eta^2}$, d.h. die Schraubenlinie ist genau dann nach Bogenlänge parametrisiert, wenn $\omega = 1/\sqrt{\varrho^2 + \eta^2}$ gilt.

2. Eine affine Kurve der Bauart

$$\mathbf{c}(t) = \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} + t \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$$

ist nicht von der dreidimensionalen Frenet-Theorie abgedeckt, eben weil die zwei Vektoren

$$\dot{\mathbf{c}}(t) = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}, \quad \ddot{\mathbf{c}}(t) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

nicht linear unabhängig sind. Dies ist auf den ersten Blick überraschend, aber kann wie folgt verstanden werden: In diesem Fall ist zwar klar, was $\mathbf{b}_1(t)$ sein soll, aber $\mathbf{b}_2(t)$ und $\mathbf{b}_3(t)$ können — da sich eine affine Raumkurve überhaupt nicht krümmt — *nicht eindeutig* festgelegt werden, sondern es gibt viele Möglichkeiten ein mitbewegtes Dreibein einzuführen.

3. Eine Raumkurve mit $c_3(t) = 0$ ist eigentlich eine planare Kurve in der (x_1, x_2) -Ebene. In diesem Fall gilt

$$\kappa_2(t) = 0, \quad \mathbf{b}_3(t) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad b_{1,3}(t) = b_{2,3}(t) = 0$$

sowie

$$\dot{\mathbf{b}}_1(t) = +\ell(t) \kappa_1(t) \mathbf{b}_2(t), \quad \dot{\mathbf{b}}_2(t) = -\ell(t) \kappa_1(t) \mathbf{b}_1(t),$$

wobei $b_{j,3}$ die x_3 -Komponente von \mathbf{b}_j meint. Oder anders gesagt: Die Vektoren $\mathbf{b}_1(t)$ und $\mathbf{b}_2(t)$ liegen immer in der (x_1, x_2) Ebene und beschreiben gerade das Zweibein an die planare Kurve, wobei κ_1 die Rolle der ebenen Krümmung übernimmt.

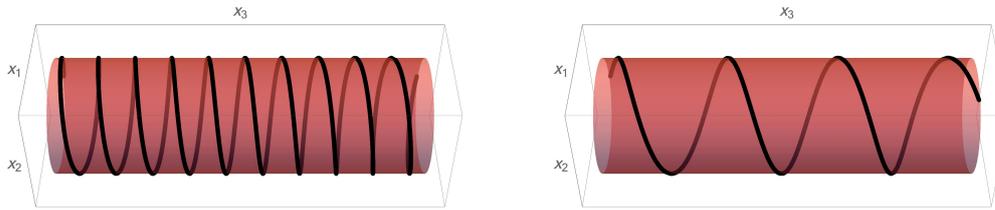


Abbildung Zwei Schraubenlinien (schwarz) aus dem Beispiel mit einem kleinen (links) bzw. großen Wert (rechts) von η , die auf demselben Kreiszyylinder (rot) mit Radius ϱ liegen. Beachte, dass die x_3 -Achse im Bild waagrecht liegt.

Bemerkung

1. Die Theorie der gewöhnlichen Differentialgleichungen impliziert, dass jede Kurve mit konstanter Krümmung und konstanter Torsion eine (im Allgemeinen gedrehte und verschobene) Schraubenlinie ist, wobei dann ϱ und η nach Umstellung mit den obigen Formeln aus κ_1 und κ_2 berechnet werden können.
2. Für eine allgemeine Raumkurve \mathbf{c} liefern $\kappa_1(t)$ und $\kappa_2(t)$ den Radius und die Ganghöhe einer Schraubenlinie, die sich lokal, d.h. in der Nähe von $\mathbf{c}(t)$, an die Kurve anschmiegt.
3. Der *Hauptsatz über Raumkurven* besagt ganz allgemein, dass eine Raumkurve bis auf Drehungen und Verschiebungen eindeutig durch ihre Krümmung und ihre Torsion, d.h. durch κ_1 und κ_2 festgelegt ist. Dies ergibt sich wieder als direkte Konsequenz der Differentialgleichungen von Frenet-Serret.

Theorem (vollständiger Formelsatz für die Frenet-Größen in 3D) Es gilt

$$\mathbf{b}_1(t) = \frac{\dot{\mathbf{c}}(t)}{|\dot{\mathbf{c}}(t)|}, \quad \mathbf{b}_2(t) = \mathbf{b}_3(t) \times \mathbf{b}_1(t), \quad \mathbf{b}_3(t) = \frac{\dot{\mathbf{c}}(t) \times \ddot{\mathbf{c}}(t)}{|\dot{\mathbf{c}}(t) \times \ddot{\mathbf{c}}(t)|}$$

sowie

$$\kappa_1(t) = \frac{|\dot{\mathbf{c}}(t) \times \ddot{\mathbf{c}}(t)|}{|\dot{\mathbf{c}}(t)|^3}, \quad \kappa_2(t) = \frac{\det \begin{pmatrix} | & | & | \\ \dot{\mathbf{c}}(t) & \ddot{\mathbf{c}}(t) & \ddot{\mathbf{c}}(t) \\ | & | & | \end{pmatrix}}{|\dot{\mathbf{c}}(t) \times \ddot{\mathbf{c}}(t)|^2},$$

d.h. für jedes $t \in I$ müssen die ersten drei Ableitungen von \mathbf{c} , zwei Kreuzprodukte, eine Determinante sowie zwei Normen berechnet werden.

Beweis Durch längere Rechnungen können wir zeigen, dass die Formeln aus der Behauptung wirklich den Differentialgleichungen von Frenet-Serret genügen, wobei man ohne Beschränkung der Allgemeinheit annehmen darf, dass die Kurve nach Bogenlänge parametrisiert ist (dann sind die Rechnungen zwar immer noch lang, aber doch deutlich einfacher als im allgemeinen Fall). Alternativ kann man sie auch aus geometrischen Überlegungen ableiten, siehe zum Beispiel [ABHKLS, Seite 960ff]. \square

Bemerkung In 3D ist $\kappa_1(t)$ immer positiv, wohingegen es keine Einschränkung an das Vorzeichen von $\kappa_2(t)$ gibt. Analog kann in 2D die Größe κ jedes Vorzeichen annehmen.

Vorlesung 10-2 : 20. Januar

3.4 Grundlagen der Flächentheorie

Vorbemerkung Wir wollen in diesem Abschnitt Flächen im dreidimensionalen Raum beschreiben, wobei es viele Analogien zur Theorie der Kurven gibt. Salopp können wir sagen: Kurven können durch einen Parameter, Flächen jedoch durch zwei reelle Parameter beschrieben werden. In diesem Sinne sind Kurven bzw. Flächen *ein-* bzw. *zweidimensional*, da sie ein bzw. zwei *innere Freiheitsgrade* aufweisen.

Definition Eine stetig differenzierbare Abbildung $\mathbf{p} : D \rightarrow \mathbb{R}^3$ mit $D \subset \mathbb{R}^2$ wird parametrisierte Fläche (oder auch parametrisiertes Flächenstück) genannt, wobei wir im Folgenden die Elemente aus D mit \mathbf{u} bezeichnen und $\mathbf{x} = \mathbf{p}(\mathbf{u})$ schreiben. Die Fläche wird regulär genannt, wenn die 3D-Vektoren

$$\partial_{u_1} \mathbf{p}(\mathbf{u}) = \begin{pmatrix} \partial_{u_1} p_1(\mathbf{u}) \\ \partial_{u_1} p_2(\mathbf{u}) \\ \partial_{u_1} p_3(\mathbf{u}) \end{pmatrix}, \quad \partial_{u_2} \mathbf{p}(\mathbf{u}) = \begin{pmatrix} \partial_{u_2} p_1(\mathbf{u}) \\ \partial_{u_2} p_2(\mathbf{u}) \\ \partial_{u_2} p_3(\mathbf{u}) \end{pmatrix}$$

für jedes $\mathbf{u} \in D$ linear unabhängig sind.

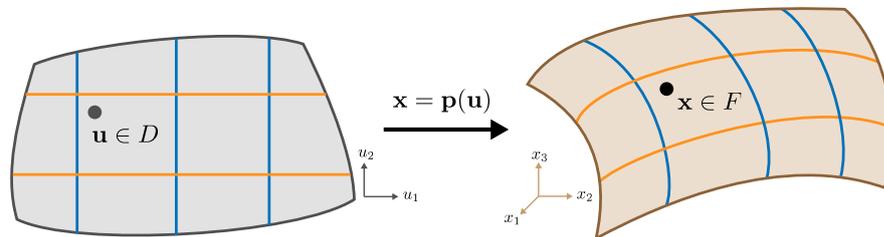


Abbildung Eine parametrisierte Fläche ist eine Abbildung \mathbf{p} , die eine *flache* Teilmenge $D \subset \mathbb{R}^2$ (grau) auf eine *gekrümmte* Teilmenge $F \subset \mathbb{R}^3$ (braun) abbildet. Die blauen bzw. orangen Linien repräsentieren die Kurven $u_1 = \text{const}$ bzw. $u_2 = \text{const}$ und liefern in D bzw. F das sogenannte *Parameter-Netz* bzgl. \mathbf{u} (siehe dazu weiter unten).

Begriffsbildung

1. Die Punktmenge

$$F = \text{im}(\mathbf{p}) = \left\{ \mathbf{p}(\mathbf{u}) : \mathbf{u} \in D \right\}$$

ist gerade das Bild von \mathbf{p} und entspricht dem, was wir landläufig eine Fläche nennen. Man sagt auch, \mathbf{p} sei eine Parametrisierung von F , und nennt F manchmal auch unparametrisierte Fläche.

2. Die Komponenten u_1 und u_2 von \mathbf{u} werden *Parameter* oder auch (krummlinige Koordinaten) auf F genannt. Die Idee ist, dass man jeden Raumpunkt $\mathbf{x} \in F$ durch die Angabe der entsprechenden Werte von u_1 und u_2 beschreiben kann.

3. In der Mathematik werden sowohl \mathbf{p} als auch F häufig einfach *Fläche* (oder *Flächenstück*) genannt, wobei in aller Regel aus dem Kontext klar wird, ob gerade eine Abbildung oder eine Punktmenge — bzw. eine Parametrisierung oder ihr Bild — gemeint ist. Eine ähnliche Ambivalenz hatten wir schon bei Kurven kennengelernt.

Hinweis Auch in diesem Abschnitt müssen Sie unbedingt versuchen, die abstrakten Formeln mit Ihrer geometrischen Anschauung in Einklang zu bringen. Dann wird auch klar, dass alles gar nicht so kompliziert ist, wie es auf den ersten Blick aussieht.

Beispiele

1. Für jedes feste $\varrho > 0$ wird durch

$$\mathbf{p}(\mathbf{u}) = \begin{pmatrix} \varrho \cos(u_1) \\ \varrho \sin(u_1) \\ u_2 \end{pmatrix}, \quad \mathbf{u} \in D = \mathbb{R}^2$$

eine reguläre parametrisierte Fläche beschrieben, denn für jedes $\mathbf{u} \in \mathbb{R}^2$ sind die Vektoren

$$\partial_{u_1} \mathbf{p}(\mathbf{u}) = \begin{pmatrix} -\varrho \sin(u_1) \\ +\varrho \cos(u_1) \\ 0 \end{pmatrix}, \quad \partial_{u_2} \mathbf{p}(\mathbf{u}) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

offensichtlich linear unabhängig. Es gilt

$$F = \left\{ \mathbf{x} \in \mathbb{R}^3 : x_1^2 + x_2^2 = \varrho^2 \right\},$$

d.h. es handelt sich um eine Parametrisierung eines unendlichen Kreiszyinders mit Radius ϱ , dessen Mittellinie gerade die x_3 -Achse ist. Beachte, dass \mathbf{p} nicht injektiv ist, denn wegen $\mathbf{p}(u_1, u_2) = \mathbf{p}(u_1 + 2\pi, u_2)$ wird jeder Punkt in F mehrmals erreicht. Im Fall von

$$D = [0, 2\pi) \times [0, \eta]$$

ergibt sich jedoch die bijektive Parametrisierung eines endlichen Kreiszyinders mit Höhe η .

Bemerkung: In der Praxis wird man meist die Flächenparameter für F anders bezeichnen und zum Beispiel θ bzw. h statt u_1 bzw. u_2 schreiben.

2. Die Sphäre mit Radius ϱ und Mittelpunkt $\mathbf{0}$ kann durch

$$\mathbf{p}(\mathbf{u}) = \begin{pmatrix} \varrho \cos(u_1) \cos(u_2) \\ \varrho \sin(u_1) \cos(u_2) \\ \varrho \sin(u_2) \end{pmatrix}, \quad \mathbf{u} \in D = \mathbb{R}^2$$

parametrisiert werden, wobei u_1 und u_2 gerade die Euler-Winkel sind. Diese entsprechen den *Längen-* und *Breitengraden* in der Nautik bzw. Geodäsie. Es gibt viele weitere Parametrisierungen der Sphäre (zum Beispiel in der Kartographie), von denen wir einige noch kennenlernen werden.

3. Die Formeln

$$\mathbf{p}(\mathbf{u}) = \begin{pmatrix} u_1 \cos(u_2) \\ u_1 \sin(u_2) \\ u_2 \end{pmatrix}, \quad u_1 \in [0, \varrho], \quad u_2 \in [0, \eta]$$

liefern eine Parametrisierung des *Helikoiden* bzw. einer *Wendelfläche*.

4. Ein weiteres Standardbeispiel ist die parametrisierte Fläche

$$\mathbf{p}(\mathbf{u}) = \begin{pmatrix} \cos(u_1) (\varrho_2 + \varrho_1 \cos(u_2)) \\ \sin(u_1) (\varrho_2 + \varrho_1 \cos(u_2)) \\ \varrho_1 \sin(u_2) \end{pmatrix}, \quad \mathbf{u} \in D = \mathbb{R}^2$$

die für $0 < \varrho_1 < \varrho_2$ einen *Torus* beschreibt.

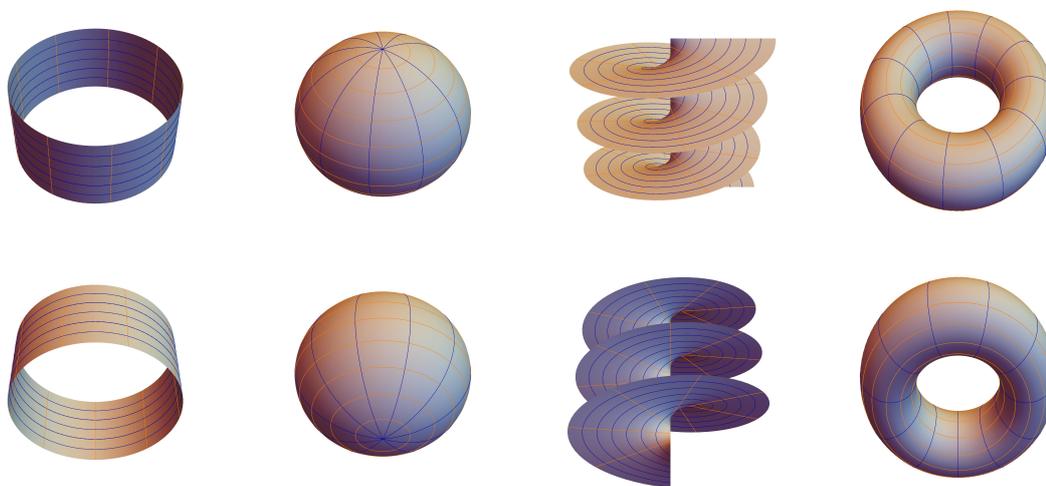


Abbildung Die vier Flächen aus den Beispielen. Die obere bzw. untere Zeile entspricht einem Blick von oben vorn bzw. links unten.

Bemerkung

1. Wir haben zunächst *nicht* vorausgesetzt, dass die Abbildung \mathbf{p} injektiv ist, werden dies aber später, vor allem bei der Berechnung von Flächenintegralen, zusätzlich fordern.
2. Regularität einer Fläche meint, dass die Jacobi-Matrix

$$\mathbf{J}\mathbf{p}(\mathbf{u}) = \begin{pmatrix} | & | \\ \partial_{u_1}\mathbf{p}(\mathbf{u}) & \partial_{u_2}\mathbf{p}(\mathbf{u}) \\ | & | \end{pmatrix} \in \mathbb{R}^{(3,2)}$$

für jedes $\mathbf{u} \in D$ den Rang 2 besitzt. Dies impliziert insbesondere $\partial_{u_1}\mathbf{p}(\mathbf{u}) \neq \mathbf{0}$ sowie $\partial_{u_2}\mathbf{p}(\mathbf{u}) \neq \mathbf{0}$.

3. Annahme: Wir werden in dieser Vorlesung immer reguläre Flächen studieren und keine Entartungspunkte zulassen. Außerdem haben wir der Einfachheit halber zunächst auch Abbildungen \mathbf{p} ausgeschlossen, die nur stückweise stetig differenzierbar sind. Solche Abbildungen sind aber in den Anwendungen durchaus wichtig, zum Beispiel um die Randkanten von Quadern richtig zu beschreiben.

4. Verallgemeinerung*: Eine stetig differenzierbare Abbildung $\mathbf{p} : U \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$ wird *parametrisierte m -dimensionale Mannigfaltigkeit* im \mathbb{R}^n genannt. Kurven bzw. Flächen sind also ein- bzw. zweidimensionale Mannigfaltigkeiten. In der Allgemeinen Relativitätstheorie studiert man zum Beispiel vierdimensionale Mannigfaltigkeiten (die gekrümmte *Raumzeit*).

Tangentialvektoren und Tangentialebene Für jeden Punkt $\mathbf{x}_* = \mathbf{p}(\mathbf{u}_*) \in F$ wird der zweidimensionale Vektorraum

$$T_{\mathbf{x}_*}F = \left\{ \lambda_1 \partial_{u_1} \mathbf{p}(\mathbf{u}_*) + \lambda_2 \partial_{u_2} \mathbf{p}(\mathbf{u}_*) : \lambda_1, \lambda_2 \in \mathbb{R} \right\} = \text{span} \{ \partial_{u_1} \mathbf{p}(\mathbf{u}_*), \partial_{u_2} \mathbf{p}(\mathbf{u}_*) \}$$

der lineare Tangentialraum an F im Punkt \mathbf{x}_* genannt und seine Elemente werden als Tangentialvektoren bezeichnet. Es gibt darüberhinaus den affinen Tangentialraum $E_{\mathbf{x}_*}F$, der durch die Bedingung

$$\mathbf{x} \in E_{\mathbf{x}_*}F \quad \Leftrightarrow \quad \mathbf{x} - \mathbf{x}_* \in T_{\mathbf{x}_*}F$$

charakterisiert ist. Oder anders gesagt: Der affine Tangentialraum entsteht, wenn der lineare Tangentialraum in den Punkt \mathbf{x}_* verschoben wird. Er wird auch die Tangentialebene an F in \mathbf{x}_* genannt.

Geometrische Interpretation: $E_{\mathbf{x}_*}F$ ist gerade die Ebene im \mathbb{R}^3 , die sich im Punkt \mathbf{x}_* an F anschmiegt (siehe Bild).

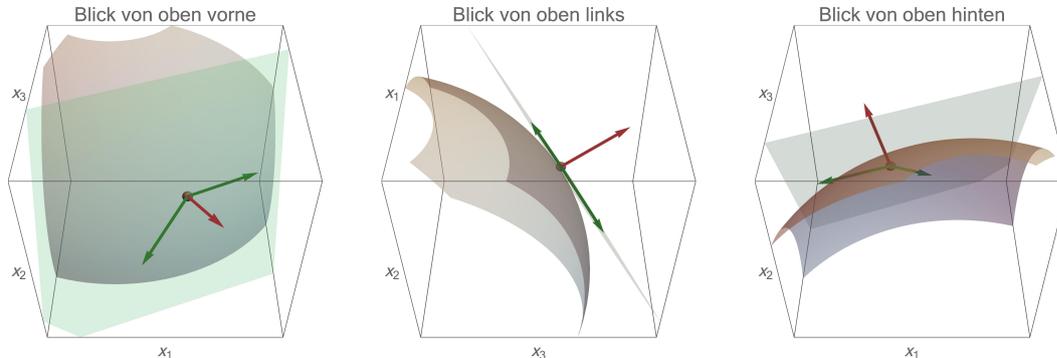


Abbildung Links: Der affine Tangentialraum (grün) in einem Punkt (grau) einer Fläche (braun) sowie zwei Tangentialvektoren (grün) und ein Normalenvektor (rot). Die drei Bilder entsprechen dabei verschiedenen Blickrichtungen.

Flächennormalenvektor In jedem Punkt $\mathbf{x}_* = \mathbf{p}(\mathbf{u}_*) \in F$ existieren immer zwei Kandidaten für einen *normierten Normalenvektor* $\boldsymbol{\nu}(\mathbf{x}_*)$, sodass stets

$$T_{\mathbf{x}_*}F = \left\{ \mathbf{x} \in \mathbb{R}^3 : \langle \mathbf{x}, \boldsymbol{\nu}(\mathbf{x}_*) \rangle = 0 \right\}, \quad |\boldsymbol{\nu}(\mathbf{x}_*)| = 1$$

gilt. Man kann diese Vektoren mit Hilfe einer Parametrisierung durch

$$\boldsymbol{\nu}(\mathbf{p}(\mathbf{u}_*)) = \pm \frac{\partial_{u_1} \mathbf{p}(\mathbf{u}_*) \times \partial_{u_2} \mathbf{p}(\mathbf{u}_*)}{|\partial_{u_1} \mathbf{p}(\mathbf{u}_*) \times \partial_{u_2} \mathbf{p}(\mathbf{u}_*)|}$$

berechnen, wobei \times das Kreuzprodukt im \mathbb{R}^3 ist.

Beispiele

1. Für die bereits oben angegebene Parametrisierung der Sphäre mit Radius ϱ und Mittelpunkt $\mathbf{0}$ ergeben sich für jeden Punkt

$$F \in \mathbf{x}_* = \mathbf{p}(\mathbf{u}_*) = \varrho \begin{pmatrix} \cos(u_{*,1}) \cos(u_{*,2}) \\ \sin(u_{*,1}) \cos(u_{*,2}) \\ \sin(u_{*,2}) \end{pmatrix}$$

die Formeln

$$\partial_{u_1} \mathbf{p}(\mathbf{u}_*) = \varrho \begin{pmatrix} -\sin(u_{*,1}) \cos(u_{*,2}) \\ +\cos(u_{*,1}) \cos(u_{*,2}) \\ 0 \end{pmatrix}, \quad \partial_{u_2} \mathbf{p}(\mathbf{u}_*) = \varrho \begin{pmatrix} -\cos(u_{*,1}) \sin(u_{*,2}) \\ -\sin(u_{*,1}) \sin(u_{*,2}) \\ +\cos(u_{*,2}) \end{pmatrix}.$$

Man sieht nun leicht, dass diese beiden tangentialen Basisvektoren nicht nur linear unabhängig sind, sondern auch senkrecht auf \mathbf{x}_* stehen. Insbesondere kann

$$\nu(\mathbf{x}_*) = \mathbf{x}_* / |\mathbf{x}_*|$$

gewählt werden, wobei wir dies alternativ auch mit Hilfe des Kreuzproduktes hätten ausrechnen können. Damit ergibt sich

$$T_{\mathbf{x}_*} F = \left\{ \mathbf{x} \in \mathbb{R}^3 : \langle \mathbf{x}, \mathbf{x}_* \rangle = 0 \right\},$$

d.h. bei einer Sphäre mit Mittelpunkt $\mathbf{0}$ besteht der lineare Tangentialraum aus allen Vektoren \mathbf{x} , die senkrecht auf \mathbf{x}_* stehen. Für den affinen Tangentialraum erhalten wir schließlich

$$E_{\mathbf{x}_*} F = \left\{ \mathbf{x} \in \mathbb{R}^3 : \langle \mathbf{x} - \mathbf{x}_*, \mathbf{x}_* \rangle = 0 \right\}.$$

2. Für den Helikoiden berechnen wir mit Hilfe der oben angegebenen Parametrisierung für jedes $\mathbf{u} \in D$ die Tangentialvektoren

$$\partial_{u_1} \mathbf{p}(\mathbf{u}) = \begin{pmatrix} +\cos(u_2) \\ +\sin(u_2) \\ 0 \end{pmatrix}, \quad \partial_{u_2} \mathbf{p}(\mathbf{u}) = \begin{pmatrix} -u_1 \sin(u_2) \\ +u_1 \cos(u_2) \\ 1 \end{pmatrix}$$

und wegen

$$\partial_{u_1} \mathbf{p}(\mathbf{u}) \times \partial_{u_2} \mathbf{p}(\mathbf{u}) = \begin{pmatrix} +\sin(u_2) \\ -\cos(u_2) \\ u_1 \end{pmatrix}$$

erhalten wir

$$\nu(\mathbf{p}(\mathbf{u})) = \frac{\pm 1}{\sqrt{1+u_1^2}} \begin{pmatrix} +\sin(u_2) \\ -\cos(u_2) \\ u_1 \end{pmatrix}.$$

In diesem Beispiel können wir die Formeln nicht weiter vereinfachen.

Bemerkung

1. Bei $T_{\mathbf{x}_*}F$ handelt sich immer um einen *linearen Unterraum* des \mathbb{R}^3 , der wegen der Regularitätsforderung die Dimension 2 besitzt und von $\partial_{u_1}\mathbf{p}(\mathbf{u}_*)$ und $\partial_{u_2}\mathbf{p}(\mathbf{u}_*)$ aufgespannt wird. Insbesondere gilt immer $\mathbf{0} \in T_{\mathbf{x}_*}F$.
2. Der Tangentialraum $T_{\mathbf{x}_*}F$ hängt von \mathbf{x}_* ab, d.h. jeder Punkt von F besitzt im Allgemeinen einen anderen Tangentialraum. Beachte auch, dass die Basisvektoren $\partial_{u_1}\mathbf{p}(\mathbf{u}_*)$ und $\partial_{u_2}\mathbf{p}(\mathbf{u}_*)$ im Allgemeinen weder senkrecht aufeinander stehen noch die Länge 1 besitzen.
3. Die affine Tangentialebene $E_{\mathbf{x}_*}F$ enthält immer den Punkt \mathbf{x}_* und besitzt die Punkt-Richtungsgleichung

$$\mathbf{x} = \mathbf{x}_* + \lambda_1 \partial_{u_1}\mathbf{p}(\mathbf{u}_*) + \lambda_2 \partial_{u_2}\mathbf{p}(\mathbf{u}_*) = \mathbf{x}_* + \mathbf{Jp}(\mathbf{u}_*) \cdot \boldsymbol{\lambda},$$

wobei die Komponenten λ_j von $\boldsymbol{\lambda} \in \mathbb{R}^2$ die freien Parameter sind und $\mathbf{Jp}(\mathbf{u}_*)$ die Jacobi-Matrix von \mathbf{p} in \mathbf{u}_* bezeichnet (die drei Zeilen sowie zwei Spalten, nämlich die beiden Vektoren $\partial_{u_j}\mathbf{p}(\mathbf{u}_*)$ besitzt).

*Zusatz**: Die rechte Seite der Punkt-Richtungsgleichung kann auch als eine parametrisierte Fläche verstanden werden, wobei λ_1 und λ_2 die beiden Parameter sind. Interpretiert man diese als die beiden Komponenten des Differenzvektors $\boldsymbol{\lambda} = \mathbf{u} - \mathbf{u}_*$, so kann die rechte Seite als $\mathbf{q}(\mathbf{u})$ geschrieben werden, wobei $\mathbf{q} : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ durch

$$\mathbf{q}(\mathbf{u}) = \mathbf{p}(\mathbf{u}_*) + \mathbf{Jp}(\mathbf{u}_*) \cdot (\mathbf{u} - \mathbf{u}_*)$$

gegeben ist. Man bezeichnet \mathbf{q} auch als Linearisierung von \mathbf{p} im Entwicklungspunkt \mathbf{u}_* , wobei es sich letztlich um eine vektorwertige Variante des Taylor-Polynoms erster Ordnung handelt und die Formeln

$$\mathbf{q}(\mathbf{u}_*) = \mathbf{p}(\mathbf{u}_*), \quad \partial_{u_1}\mathbf{q}(\mathbf{u}_*) = \partial_{u_1}\mathbf{p}(\mathbf{u}_*), \quad \partial_{u_2}\mathbf{q}(\mathbf{u}_*) = \partial_{u_2}\mathbf{p}(\mathbf{u}_*)$$

nach Konstruktion gelten. Solche Linearisierungen spielen in allen Bereichen der Mathematik und der Anwendungswissenschaften eine wesentliche Rolle.

4. In der Literatur wird nicht immer klar zwischen dem linearen und dem affinen Tangentialraum unterschieden und die Notation $T_{\mathbf{x}_*}F$ wird manchmal für beide verwendet. In der Regel wird aber aus dem Kontext klar, welche Variante des Tangentialraumes gerade gemeint ist bzw. ob nun $\mathbf{0}$ oder \mathbf{x}_* in diesem Raum liegt.
5. Der lineare und der affine Tangentialraum reflektieren geometrische Eigenschaften von $F = \text{im}(\mathbf{p})$ und daher existieren sie unabhängig von der Parametrisierung (siehe dazu auch die Diskussion weiter unten). Durch eine Parametrisierung wird aber eine Basis ausgezeichnet, nämlich die Vektoren $\partial_{u_1}\mathbf{p}(\mathbf{u}_*)$ und $\partial_{u_2}\mathbf{p}(\mathbf{u}_*)$, die als partielle Ableitungen von \mathbf{p} meist leicht berechnet werden können.
6. Die Frage, ob es in der Kreuzprodukt-Formel für $\boldsymbol{\nu}(\mathbf{p}(\mathbf{u}_*))$ eine sinnvolle Wahl für das Vorzeichen gibt, wird uns weiter unten noch beschäftigen, wenn wir die *Orientierbarkeit von Flächen* diskutieren.
7. Die Abbildung $\mathbf{u} \in D \mapsto \boldsymbol{\nu}(\mathbf{p}(\mathbf{u}))$ wird *Gauß-Abbildung* zu \mathbf{p} genannt und spielt in der mathematischen Theorie gekrümmter Flächen eine herausgehobene Rolle. Insbesondere können die partiellen Ableitungen dieser Abbildung benutzt werden, um die beiden *Hauptkrümmungen von Flächen* zu verstehen bzw. zu berechnen.

Graphen skalarer Funktionen mit zwei Variablen Ist $D \subset \mathbb{R}^2$ eine Menge und $f : D \rightarrow \mathbb{R}$ eine Funktion mit zwei Variablen, so kann der Graph von f mittels der natürlichen Parametrisierung

$$\mathbf{p}(\mathbf{u}) = \begin{pmatrix} u_1 \\ u_2 \\ f(u_1, u_2) \end{pmatrix}$$

als Fläche betrachtet werden. Mit $\mathbf{x} = \mathbf{p}(\mathbf{u})$ ergibt sich

$$x_1 = u_1, \quad x_2 = u_2, \quad x_3 = f(u_1, u_2)$$

und man schreibt deshalb einfach x_1 statt u_1 und x_2 statt u_2 sowie

$$F = \text{graph } f = \left\{ \mathbf{x} \in \mathbb{R}^3 : x_3 = f(x_1, x_2), (x_1, x_2) \in D \right\}.$$

Die entsprechenden Tangentialvektoren können mittels

$$\partial_{x_1} \mathbf{p}(x_1, x_2) = \begin{pmatrix} 1 \\ 0 \\ \partial_{x_1} f(x_1, x_2) \end{pmatrix}, \quad \partial_{x_2} \mathbf{p}(x_1, x_2) = \begin{pmatrix} 0 \\ 1 \\ \partial_{x_2} f(x_1, x_2) \end{pmatrix}$$

berechnet werden und für jedes $(x_{*,1}, x_{*,2}) \in D$ erhalten wir

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_{*,1} \\ x_{*,2} \\ f(x_{*,1}, x_{*,2}) \end{pmatrix} + \lambda_1 \begin{pmatrix} 1 \\ 0 \\ \partial_{x_1} f(x_{*,1}, x_{*,2}) \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ 1 \\ \partial_{x_2} f(x_{*,1}, x_{*,2}) \end{pmatrix}$$

als Punkt-Richtungsdarstellung der Tangentialebene $E_{(x_{*,1}, x_{*,2}, x_{*,3})} F$, wobei wir die konsistente Abkürzung $x_{*,3} = f(x_{*,1}, x_{*,2})$ verwendet haben. Der entsprechende Normalenvektor ist durch

$$\boldsymbol{\nu}(x_{*,1}, x_{*,2}, x_{*,3}) = \frac{\pm 1}{\sqrt{1 + |\text{grad } f(x_{*,1}, x_{*,2})|^2}} \begin{pmatrix} -\partial_{x_1} f(x_{*,1}, x_{*,2}) \\ -\partial_{x_2} f(x_{*,1}, x_{*,2}) \\ +1 \end{pmatrix}$$

bis auf sein Vorzeichen bestimmt.

*Zusatz**: Die ersten beiden Gleichungen für die Tangentialebene liefern $\lambda_1 = x_1 - x_{*,1}$ sowie $\lambda_2 = x_2 - x_{*,2}$ und nach Einsetzen in die dritte Gleichung erhalten wir

$$x_3 = f(x_{*,1}, x_{*,2}) + \partial_{x_1} f(x_{*,1}, x_{*,2})(x_1 - x_{*,1}) + \partial_{x_2} f(x_{*,1}, x_{*,2})(x_2 - x_{*,2}).$$

Wir können diese Formeln wie folgt verstehen: Die Tangentialebene an der Graphen von f im Punkt $(x_{*,1}, x_{*,2}, x_{*,3})$ ist gerade der Graph des ersten Taylor-Polynoms von f im Entwicklungspunkt $(x_{*,1}, x_{*,2})$.

Beispiel Der Graph der Funktion

$$f(x_1, x_2) = x_1^2 + x_2^2, \quad (x_1, x_2) \in \mathbb{R}^2$$

ist ein Rotationsparaboloid und mit den obigen Notationen erhalten wir

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_{*,1} + \lambda_1 \\ x_{*,2} + \lambda_2 \\ x_{*,1}^2 + x_{*,2}^2 + 2\lambda_1 x_{*,1} + 2\lambda_2 x_{*,2} \end{pmatrix}$$

als Parameterdarstellung der Tangentialebene an den Graphen von f im Punkt $(x_{*,1}, x_{*,2}, x_{*,3})$ mit $x_{*,3} = x_{*,1}^2 + x_{*,2}^2$. Außerdem gilt

$$\boldsymbol{\nu}(x_{*,1}, x_{*,2}, x_{*,3}) = \frac{\pm 1}{\sqrt{1 + 4x_{*,1}^2 + 4x_{*,2}^2}} \begin{pmatrix} -2x_{*,1} \\ -2x_{*,2} \\ 1 \end{pmatrix}$$

für einen entsprechenden Normalenvektor.

Flächenkurven Ist $\mathbf{a} : I \rightarrow D$ eine gegebene planare Kurve in $D \subset \mathbb{R}^2$, so wird durch

$$\mathbf{c} := \mathbf{p} \circ \mathbf{a} \quad \text{bzw.} \quad \mathbf{c}(t) := \mathbf{p}(\mathbf{a}(t))$$

in natürlicher Weise eine Raumkurve $\mathbf{c} : I \rightarrow \mathbb{R}^3$ definiert, die man die *Liftung von \mathbf{a} durch \mathbf{p}* nennt. Dabei handelt es sich um ein Kurve in F , da $\mathbf{c}(t) \in F$ für alle $t \in I$ gilt. Die Kettenregel impliziert außerdem

$$\dot{\mathbf{c}}(t) = \mathbf{J}\mathbf{p}(\mathbf{a}(t)) \cdot \dot{\mathbf{a}}(t) = \dot{a}_1(t) \partial_{u_1} \mathbf{p}(\mathbf{a}(t)) + \dot{a}_2(t) \partial_{u_2} \mathbf{p}(\mathbf{a}(t)),$$

d.h. der momentane Tangentialvektor $\dot{\mathbf{c}}(t)$ an die Kurve liegt immer im linearen Tangentialraum $T_{\mathbf{c}(t)}F$ der Fläche, wobei die Komponenten $\dot{a}_j(t)$ gerade Koeffizienten bzgl. der Basisvektoren $\partial_{u_j} \mathbf{p}(\mathbf{a}(t))$ sind.

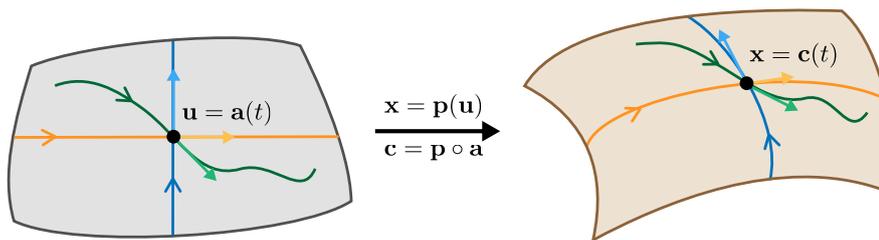


Abbildung Mittels p kann jede planare Kurve \mathbf{a} in D (grün links) in eine Raumkurve $\mathbf{c} = \mathbf{p} \circ \mathbf{a}$ (grün rechts) überführt (bzw. *geliftet*) werden, die in der Fläche F liegt. Die Liftungen der kartesischen Kurven in D (blau und orange links) erzeugen dabei das \mathbf{u} -Parameter-Netz in F (blau und orange rechts). Die farbigen Pfeile repräsentieren Tangentialvektoren, die im \mathbb{R}^2 (links) bzw. im linearen Tangentialraum der Fläche liegen.

Bemerkung Man kann — zumindest in einem lokalen Sinne — jede Flächenkurve $\mathbf{c} : I \rightarrow F$ in F als Liftung einer planaren Kurve $\mathbf{a} : I \rightarrow D$ betrachten. Dies folgt aus dem Satz über implizite Funktionen, aber wir wollen hier auf eine formale Herleitung verzichten.

Parameter-Netze Ist $\mathbf{u}_* \in D$ ein festgehaltener Punkt in D und $\mathbf{x}_* = \mathbf{p}(\mathbf{u}_*) \in F$ sein Bild unter \mathbf{p} , so verlaufen die kartesischen Kurven

$$\mathbf{a}^{(1)}(t) = \mathbf{u}_* + \begin{pmatrix} t \\ 0 \end{pmatrix}, \quad \mathbf{a}^{(2)}(t) = \mathbf{u}_* + \begin{pmatrix} 0 \\ t \end{pmatrix}$$

parallel zu den Koordinatenachsen in D und passieren zur Zeit $t = 0$ den Punkt \mathbf{u}_* . Insbesondere gilt $u_2 = \text{const} = u_{*,2}$ bzw. $u_1 = \text{const} = u_{*,1}$ entlang von $\mathbf{a}^{(1)}$ bzw. $\mathbf{a}^{(2)}$ und

$$\dot{\mathbf{a}}^{(1)}(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \dot{\mathbf{a}}^{(2)}(0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

sind die entsprechenden, zweidimensionalen Tangentialvektoren. Für die gelifteten Kurven gilt

$$\mathbf{c}^{(2)}(0) = \mathbf{x}_* = \mathbf{c}^{(2)}(0), \quad \dot{\mathbf{c}}^{(1)}(0) = \partial_{u_1} \mathbf{p}(\mathbf{u}_*), \quad \dot{\mathbf{c}}^{(2)}(0) = \partial_{u_2} \mathbf{p}(\mathbf{u}_*),$$

d.h. die durch die Parametrisierung \mathbf{p} ausgezeichneten Basisvektoren in $T_{\mathbf{x}_*} F$ sind gerade die Tangentialvektoren der gelifteten Kurven $\mathbf{c}^{(1)}$ und $\mathbf{c}^{(2)}$. Beachte, dass diese im Allgemeinen kein ON-System bilden.

Man kann die kartesischen Kurven $\mathbf{a}^{(1)}$ bzw. ihre Liftung $\mathbf{c}^{(1)}$ auch als

$$u_1 \mapsto \begin{pmatrix} u_1 \\ u_{2,*} \end{pmatrix} \quad \text{bzw.} \quad u_1 \mapsto \mathbf{p}(u_1, u_{2,*}) = \begin{pmatrix} p_1(u_1, u_{2,*}) \\ p_2(u_1, u_{2,*}) \\ p_3(u_1, u_{2,*}) \end{pmatrix}$$

schreiben. Mit dieser kompakten Notation wird klar ersichtlich, dass beide Kurven der Bedingung $u_2 = u_{2,*}$ entsprechen und in natürlicher Weise durch u_1 parametrisiert werden können. Analoge Formeln können für $\mathbf{a}^{(2)}$ bzw. $\mathbf{c}^{(2)}$ abgeleitet werden.

Werden die Liftungen von hinreichend vielen kartesischen Kurven gezeichnet, entsteht in F insgesamt ein Netz von Kurven, die jeweils einem festen Wert der Kurvenparameter u_1 oder u_2 entsprechen (siehe die Bilder). Mit Hilfe solcher Parameter-Netze kann die Bedeutung der Flächenparameter auf intuitive Weise verstanden und dargestellt werden. Außerdem kann man mit solchen Netzen bzw. den zu Grunde liegenden Flächenkurven auch die Krümmung einer Fläche charakterisieren, aber dies können wir in dieser Vorlesung nicht vertiefen.

Reparametrisierung von Flächen Analog zur Kurventheorie gibt es auch in der Flächentheorie Parameterwechsel und Reparametrisierungen. Wir nennen $\tilde{\mathbf{p}} : \tilde{D} \rightarrow \mathbb{R}^3$ eine Reparametrisierung von $\mathbf{p} : D \rightarrow \mathbb{R}^3$, sofern es *bijektive* und *zueinander inverse* Abbildungen $\Phi : \tilde{D} \rightarrow D$ und $\tilde{\Phi} : D \rightarrow \tilde{D}$ gibt, sodass

$$\tilde{\mathbf{p}}(\tilde{\mathbf{u}}) = \mathbf{p}(\mathbf{u}) \quad \text{mit} \quad \mathbf{u} = \Phi(\tilde{\mathbf{u}}) \quad \text{bzw.} \quad \tilde{\mathbf{u}} = \tilde{\Phi}(\mathbf{u})$$

für alle $\tilde{\mathbf{u}} \in \tilde{D}$ bzw. $\mathbf{u} \in D$ gilt. Die Abbildungen Φ und $\tilde{\Phi}$ sind dabei gerade die Parameterwechsel.

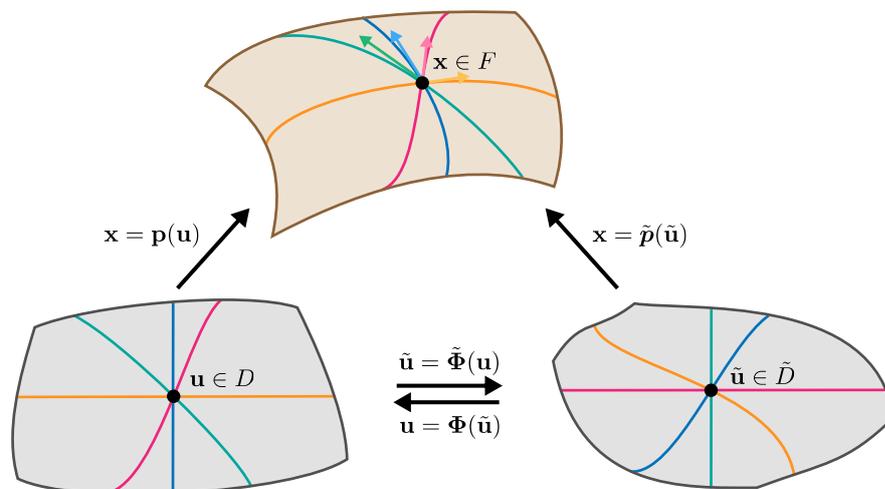


Abbildung Schematische Darstellung von Reparametrisierungen bzw. von Parameterwechseln. Die orangen und blauen bzw. rosanen und türkisen Kurven illustrieren das \mathbf{u} -Netz bzw. das $\tilde{\mathbf{u}}$ -Netz in F , aber analog auch in D und \tilde{D} . Der blaue und der orange bzw. der rosane und türkise Vektor beschreiben die Basisvektoren im Tangentialraum des schwarzen Punktes in F , die durch die partiellen Ableitungen von \mathbf{p} bzw. $\tilde{\mathbf{p}}$ festgelegt sind.

Bemerkung

1. Ganz ähnliche Konzepte hatten wir (mit anderen Notationen) schon bei der Kettenregel der Differentialrechnung sowie bei der Transformationsformel für Integrale verwendet.
2. Mit den obigen Notationen ist auch $\mathbf{p} : D \rightarrow \mathbb{R}^3$ eine Reparametrisierung von $\tilde{\mathbf{p}} : \tilde{D} \rightarrow \mathbb{R}^3$, d.h. es gibt hier (wie auch bei Kurven) eine natürliche Symmetrie in der Begriffsbildung.
3. Es gilt $\text{im}(\tilde{\mathbf{p}}) = \text{im}(\mathbf{p})$, d.h. $\tilde{\mathbf{p}}$ und \mathbf{p} beschreiben dieselbe Punktmenge $F \subset \mathbb{R}^3$, nur mit Hilfe unterschiedlicher Parameter. Außerdem ist durch

$$T_{\tilde{\mathbf{p}}(\tilde{\mathbf{u}})}F = T_{\mathbf{p}(\mathbf{u})}F, \quad E_{\tilde{\mathbf{p}}(\tilde{\mathbf{u}})}F = E_{\mathbf{p}(\mathbf{u})}F$$

sichergestellt, dass heißt ein Parameterwechsel ändert nicht die Tangentialräume. Beachte aber, dass im Allgemeinen

$$\partial_{u_j}\mathbf{p}(\mathbf{u}) \neq \partial_{\tilde{u}_j}\tilde{\mathbf{p}}(\tilde{\mathbf{u}})$$

für $j = 1, 2$ gelten wird, d.h. die partiellen Ableitungen von \mathbf{p} und $\tilde{\mathbf{p}}$ liefern in jedem Punkt von F eine andere Basis im linearen Tangentialraum. Die entsprechenden Umrechnungsformeln ergeben sich dabei aus der Kettenregel. Genauer gesagt: Mit den obigen Notationen gilt

$$\mathbf{J}\tilde{\mathbf{p}}(\tilde{\mathbf{u}}) = \mathbf{J}\mathbf{p}(\mathbf{u}) \cdot \mathbf{J}\tilde{\Phi}(\tilde{\mathbf{u}}) \quad \text{bzw.} \quad \mathbf{J}\mathbf{p}(\mathbf{u}) = \mathbf{J}\tilde{\mathbf{p}}(\tilde{\mathbf{u}}) \cdot \tilde{\Phi}(\mathbf{u})$$

für die $(2, 2)$ -Matrizen $\mathbf{J}\tilde{\Phi}(\tilde{\mathbf{u}})$ und $\mathbf{J}\tilde{\Phi}(\mathbf{u})$ sowie die $(3, 2)$ -Matrizen $\mathbf{J}\tilde{\mathbf{p}}(\tilde{\mathbf{u}})$ und $\mathbf{J}\mathbf{p}(\mathbf{u})$, wobei die Spalten der letzteren die durch $\tilde{\mathbf{p}}$ und \mathbf{p} gewählten Basisvektoren im Tangentialraum liefern.

4. Man kann sich Reparametrisierungen — bzw. den Übergang von den Parametern \mathbf{u} zu den Parametern $\tilde{\mathbf{u}}$ — anschaulich als einen Wechsel des Parameter-Netzes auf F vorstellen.

Beispiel Die Viertelsphäre

$$F = \{(x_1, x_2, x_3) : x_1^2 + x_2^2 + x_3^2 = 1, x_1 > 0, x_3 > 0\}$$

kann zum einen mittels der bereits eingeführten Euler-Winkel via

$$\mathbf{p}(\mathbf{u}) = \begin{pmatrix} \cos(u_1) \cos(u_2) \\ \sin(u_1) \cos(u_2) \\ \sin(u_2) \end{pmatrix}, \quad \mathbf{u} \in D := \left(-\frac{1}{2}\pi, -\frac{1}{2}\pi\right) \times \left(0, \frac{1}{2}\pi\right)$$

parametrisiert werden. Andererseits gelingt dies auch mit

$$\tilde{\mathbf{p}}(\tilde{\mathbf{u}}) = \begin{pmatrix} \tilde{u}_1 \\ \tilde{u}_2 \\ \sqrt{1 - \tilde{u}_1^2 - \tilde{u}_2^2} \end{pmatrix}, \quad \tilde{\mathbf{u}} \in \tilde{D} = \{(\tilde{u}_1, \tilde{u}_2) : \tilde{u}_1 > 0, \tilde{u}_1^2 + \tilde{u}_2^2 < 1\},$$

wobei \tilde{u}_1 bzw. \tilde{u}_2 via $\mathbf{x} = \tilde{\mathbf{p}}(\tilde{\mathbf{u}})$ gerade die x_1 - bzw. die x_2 -Koordinate eines Punktes $\mathbf{x} \in F$ beschreiben. Der Parameterwechsel $\tilde{\Phi}$ kann in diesem Beispiel einfach abgelesen werden und wir erhalten

$$\tilde{\Phi}(\mathbf{u}) = \begin{pmatrix} \cos(u_1) \cos(u_2) \\ \sin(u_1) \cos(u_2) \end{pmatrix}$$

als Formel, mit der $\tilde{\mathbf{u}}$ aus \mathbf{u} berechnet werden kann. Die Umkehrabbildung ergibt sich nach kleineren Rechnungen zu

$$\Phi(\tilde{\mathbf{u}}) = \begin{pmatrix} \arctan\left(\frac{\tilde{u}_2}{\tilde{u}_1}\right) \\ \arccos\left(\sqrt{\tilde{u}_1^2 + \tilde{u}_2^2}\right) \end{pmatrix},$$

wobei die Wahl von \tilde{D} sicherstellt, dass die rechte Seite für jedes $\tilde{\mathbf{u}} \in \tilde{D}$ wohldefiniert ist und einen Punkt in D liefert.

Parameter-Netze: Die Liftungen der kartesischen Kurven bzgl. der Euler-Winkel können als

$$u_1 \mapsto \begin{pmatrix} \cos(u_1) \cos(u_{*,2}) \\ \sin(u_1) \cos(u_{*,2}) \\ \sin(u_{*,2}) \end{pmatrix} \quad \text{bzw.} \quad u_2 \mapsto \begin{pmatrix} \cos(u_{*,1}) \cos(u_2) \\ \sin(u_{*,1}) \cos(u_2) \\ \sin(u_2) \end{pmatrix}$$

geschrieben werden und entsprechen gerade den Breiten- bzw. Längengraden auf der Erde, wobei $u_{*,2}$ bzw. $u_{*,1}$ den Zahlenwert des jeweiligen Grades festlegen. Die entsprechenden Formeln

$$\tilde{u}_1 \mapsto \begin{pmatrix} \tilde{u}_1 \\ \tilde{u}_{*,2} \\ \sqrt{1 - \tilde{u}_1^2 - \tilde{u}_{*,2}^2} \end{pmatrix} \quad \text{bzw.} \quad \tilde{u}_2 \mapsto \begin{pmatrix} \tilde{u}_{*,1} \\ \tilde{u}_2 \\ \sqrt{1 - \tilde{u}_{*,1}^2 - \tilde{u}_2^2} \end{pmatrix}$$

bzgl. der $\tilde{\mathbf{u}}$ -Koordinaten besitzen jedoch keine praktische Bedeutung.

Bemerkung: Wir hatten schon gesehen, dass \mathbf{p} mühelos auf ganz \mathbb{R}^2 ausgedehnt werden kann und dann jeden Punkt der dreidimensionalen Vollsphäre sogar mehrmals abdeckt. Die Parametrisierung $\tilde{\mathbf{p}}$ kann maximal auf der abgeschlossenen Einheitskreisscheibe $\{\tilde{\mathbf{u}} : |\tilde{\mathbf{u}}| \leq 1\}$ betrachtet werden, wobei sie auf dem Rand derselben nicht differenzierbar ist und außerdem nur Raumpunkte in der oberen Hemisphäre liefert. Um die Formeln für die Parameterwechsel möglichst einfach zu halten, haben wir unsere Betrachtungen auf eine Viertelsphäre ohne Rand eingeschränkt, da wir andernfalls Fallunterscheidungen bei der Formel für Φ hätten treffen müssen oder Probleme mit der Differenzierbarkeit bekommen hätten. Ganz allgemein kann man sagen, dass man bei Flächenparametrisierungen (bzw. dem Wechsel zwischen solchen) besonders aufpassen muss, damit alle Ausdrücke wohldefiniert und differenzierbar sind. In praktischen Fällen hilft in der Regel die geometrische Anschauung dabei, den Überblick zu behalten.

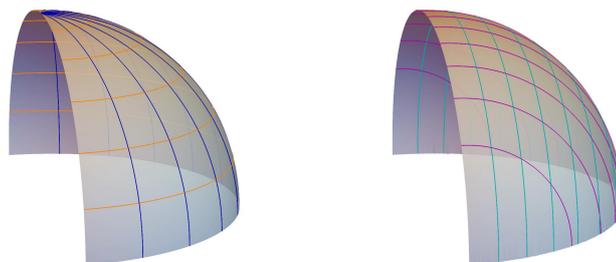


Abbildung Die Parameter-Netze der im Beispiel angegebenen Parametrisierungen der Viertelsphäre.

3.5 Flächenintegrale

Ziel In diesem Abschnitt definieren wir Integrale von Funktionen auf gekrümmten Flächen und verallgemeinern so das Konzept *Gebietsintegral* aus dem letzten Kapitel. Als Spezialfall werden wir lernen, wie Flächeninhalte gekrümmter Flächen durch Integrale berechnet werden können.

Setting und Voraussetzung Wir betrachten eine reguläre parametrisierte Fläche $\mathbf{p} : D \rightarrow \mathbb{R}^3$, die auf einer *messbaren* Teilmenge $D \subset \mathbb{R}^2$ definiert ist, und setzen wieder $F = \text{im}(\mathbf{p})$. Als Standardannahme wollen wir dabei zusätzlich voraussetzen, dass \mathbf{p} injektiv ist, obwohl wir auch leichte Abschwächungen dieser Bedingung zulassen werden.

Bemerkungen

1. Erinnerung: Die Menge D ist genau dann messbar, wenn sie in sinnvoller Weise von innen und von außen durch endlich viele zweidimensionale Quader approximiert werden kann. Bei uns sind messbare Mengen immer kompakt, aber nicht jede kompakte Teilmenge des \mathbb{R}^2 ist messbar. Die in der Praxis auftretenden Mengen werden in aller Regel messbar sein und Sie dürfen in den Hausaufgaben immer stillschweigend Messbarkeit voraussetzen.
2. Erinnerung: \mathbf{p} heißt injektiv, sofern die Implikation

$$\mathbf{u} \neq \tilde{\mathbf{u}} \implies \mathbf{p}(\mathbf{u}) \neq \mathbf{p}(\tilde{\mathbf{u}})$$

gilt, d.h. wenn verschiedene Punkte in D auf verschiedene Punkte in F abgebildet werden. Man kann die Injektivität in aller Regel durch Verkleinerung des Definitionsbereichs D von \mathbf{p} erreichen.

3. Ist \mathbf{p} injektiv, so bildet \mathbf{p} die Menge D bijektiv auf $F = \text{im}(\mathbf{p})$ ab, d.h. jeder Punkt $\mathbf{u} \in D$ entspricht via $\mathbf{x} = \mathbf{p}(\mathbf{u})$ genau einem Punkt $\mathbf{x} \in F$ und umgekehrt. Insbesondere ist die Umkehrabbildung $\mathbf{p}^{-1} : F \rightarrow D$ wohldefiniert und wird auch Karte von F oder Projektion von F genannt. Sie bildet jeden Punkt \mathbf{x} der gekrümmten Fläche F auf einen Punkt \mathbf{u} der flachen Fläche D ab.

Flächeninhalte

Flächeninhalt parametrisierter Flächen Der Flächeninhalt von $F = \text{im}(\mathbf{p})$ kann in sinnvoller Weise durch

$$\text{vol}(F) := \int_F 1 \, d\sigma := \int_D |\partial_{u_1} \mathbf{p}(\mathbf{u}) \times \partial_{u_2} \mathbf{p}(\mathbf{u})| \, d\mathbf{u}$$

definiert werden, wobei die mittlere Formel das unten eingeführte Flächenintegral 1. Art der Einsfunktion ist und $d\sigma$ das infinitesimale Flächenelement bezeichnet.

Motivation und Heuristik* Die Formel für $\text{vol}(F)$ ist sehr naheliegend. Ist D zum Beispiel ein Quader, so können wir diesen in viele Teilquader zerlegen und die Abbildung \mathbf{p} auf jedem Teilquader Q_* mit Mittelpunkt \mathbf{u}_* durch ihre lokale *Linearisierung*

$$\mathbf{q}(\mathbf{u}) = \mathbf{J}\mathbf{p}(\mathbf{u}_*) \cdot (\mathbf{u} - \mathbf{u}_*) = (u_1 - u_{*,1})\partial_{u_1}\mathbf{p}(\mathbf{u}_*) + (u_2 - u_{*,2})\partial_{u_2}\mathbf{p}(\mathbf{u}_*)$$

ersetzen (wobei \mathbf{u}_* der entsprechende *Entwicklungspunkt* ist). Diese Abbildung \mathbf{q} bildet den Teilquader Q_* auf ein schief im Raum liegendes Parallelogramm ab, dessen Kantenvektoren parallel zu den Tangentialvektoren $\partial_{u_1}\mathbf{p}(\mathbf{u}_*)$ und $\partial_{u_2}\mathbf{p}(\mathbf{u}_*)$ sind und dessen Flächeninhalt mittels

$$|\partial_{u_1}\mathbf{p}(\mathbf{u}_*) \times \partial_{u_2}\mathbf{p}(\mathbf{u}_*)| \text{ vol}(Q_*)$$

exakt berechnet werden kann (dies ist eine der Eigenschaften des Kreuzproduktes). Die Vereinigung aller dieser Raum-Parallelogramme approximiert F (siehe das Bild) und die Integralformel für $\text{vol}(F)$ beschreibt den Grenzwert der approximativen Flächeninhalte, sofern die Zerlegung von D immer feiner wählt wird. Dieselbe Strategie hatten wir schon bei der Definition von Gebietsintegralen sowie bei der Herleitung der entsprechenden Transformationsformel verwendet (ein großer Teil Mathematik besteht darin, bekannte Grundideen in immer neuen Kontexten anzuwenden und dabei leicht zu variieren). Wir könnten mit unserem Wissenstand diese Idee mathematisch präzise formulieren und auch die entsprechenden Approximationsfehler mit Hilfe des Satzes von Taylor kontrollieren, aber die Details wären sehr technisch und sollen hier nicht ausgeführt werden.

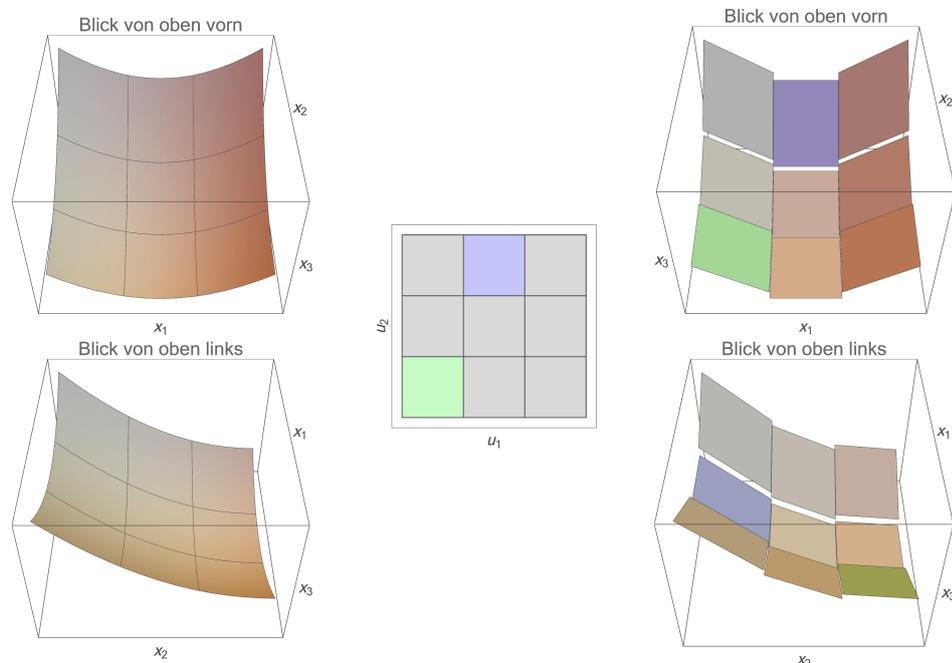


Abbildung Eine über einem Quader D (Mitte) parametrisierte glatte Fläche F (links) kann durch schief im Raum liegende Parallelogramme (rechts) approximiert werden. Dabei wird D in viele kleine Quader zerlegt (dargestellt ist eine sehr grobe Zerlegung mit nur 9 Teilquadern) und \mathbf{p} wird auf jedem Teilquader durch eine lokale Linearisierung ersetzt. Die approximierende Fläche (rechts) ist stückweise flach und weist Unstetigkeiten auf, konvergiert aber bei immer feiner werdender Zerlegung von D gegen die glatte Fläche F . Wenn D kein Quader ist, kann analog argumentiert werden. Es gibt dann weitere Approximationsfehler, die aber auch im Limes verschwinden.

Bemerkungen

1. Die Definition des Flächeninhalts stimmt mit unserer intuitiven Vorstellung überein. Insbesondere quantifiziert $\text{vol}(F)$ wirklich die Menge an Farbe, die wir brauchen, um F anzustreichen. Wir werden auch sehen, dass die abstrakte Integraldarstellung viele bekannte Oberflächenformeln als Spezialfall enthält.
2. Die Injektivität von \mathbf{p} ist wichtig, da sonst Teile der Fläche mehrmals gezählt werden. In der Praxis reicht es jedoch meist aus, wenn \mathbf{p} nur „fast injektiv“ ist. Siehe dazu das erste Beispiel zur Halbsphäre.
3. Auch die stetige Differenzierbarkeit von \mathbf{p} ist wichtig, denn andernfalls wäre der Integrand des \mathbf{u} -Integrals nicht definiert oder nicht stetig. Am Rand von D kann man aber oftmals einen Verlust von Differenzierbarkeit hinnehmen. Siehe dazu das zweite Beispiel zur Halbsphäre.
4. Statt $\text{vol}(F)$ wird oft auch $\text{area}(F)$ geschrieben.

Beispiele

1. Die obere Halbsphäre mit Radius $\varrho > 0$ und Mittelpunkt $\mathbf{0}$ kann mittels der Euler-Winkel durch

$$\mathbf{p}(\mathbf{u}) = \begin{pmatrix} \varrho \cos(u_1) \cos(u_2) \\ \varrho \sin(u_1) \cos(u_2) \\ \varrho \sin(u_2) \end{pmatrix}, \quad \mathbf{u} \in D = [0, 2\pi] \times [0, \pi/2]$$

parametrisiert werden. Mit diesen Formeln ergibt sich

$$\begin{aligned} \partial_{u_1} \mathbf{p}(\mathbf{u}) \times \partial_{u_2} \mathbf{p}(\mathbf{u}) &= \begin{pmatrix} -\varrho \sin(u_1) \cos(u_2) \\ +\varrho \cos(u_1) \cos(u_2) \\ 0 \end{pmatrix} \times \begin{pmatrix} -\varrho \cos(u_1) \sin(u_2) \\ -\varrho \sin(u_1) \sin(u_2) \\ +\varrho \cos(u_2) \end{pmatrix} \\ &= \begin{pmatrix} \varrho^2 \cos(u_1) \cos^2(u_2) \\ \varrho^2 \sin(u_1) \cos^2(u_2) \\ \varrho^2 \sin(u_2) \cos(u_2) \end{pmatrix} \end{aligned}$$

und damit

$$\begin{aligned} |\partial_{u_1} \mathbf{p}(\mathbf{u}) \times \partial_{u_2} \mathbf{p}(\mathbf{u})|^2 &= \varrho^4 \cos^2(u_1) \cos^4(u_2) + \varrho^4 \sin^2(u_1) \cos^4(u_2) \\ &\quad + \varrho^4 \sin^2(u_2) \cos^2(u_2) \\ &= \varrho^4 \cos^4(u_2) + \varrho^4 \sin^2(u_2) \cos^2(u_2) \\ &= \varrho^4 \cos^2(u_2). \end{aligned}$$

Insbesondere erhalten wir via

$$\text{vol}(F) = \varrho^2 \int_0^{2\pi} \int_0^{\pi/2} \cos(u_2) \, du_2 \, du_1 = \varrho^2 \int_0^{2\pi} 1 \, du_1 = 2 \varrho^2 \pi$$

die bekannte Formel für die Fläche einer halben Sphäre. Der Flächeninhalt der Vollsphäre ergibt sich nach Multiplikation mit 2 und wir können dieses Ergebnis alternativ auch mittels eines eindimensionalen Integrals aus der Theorie für Rotationskörper herleiten (siehe *Analysis 2*).

Bemerkungen:

- (a) Die betrachtete Parametrisierung ist streng genommen nicht injektiv, da ja $\mathbf{p}(0, u_2) = \mathbf{p}(2\pi, u_2)$ für alle u_2 sowie $\mathbf{p}(u_1, \pi/2) = (0, 0, 1)$ gilt. Die Parametrisierung ist jedoch „fast injektiv“, da die doppelt oder mehrfach abgedeckten Punkte aus F alle auf einer eindimensionalen Menge (Kurve) liegen und daher zur Fläche nichts beitragen. Generell gilt: Man muss bei praktischen Rechnungen nicht allzu pingelig sein, sofern einem die geometrische Anschauung dies erlaubt oder gar nahelegt.
- (b) Wir können durch die Wahl $D = [0, 2\pi] \times [-\pi/2, +\pi/2]$ auch die Fläche der Vollsphäre berechnen und erhalten $4\pi\rho^2$.
- (c) Mit $D = [0, 4\pi] \times [-\pi/2, +\pi/2]$ würden wir durch Berechnung des entsprechenden Integrals den doppelten Flächeninhalt der Sphäre erhalten, weil dann wegen $u_1 \in [0, 4\pi]$ fast jeder Punkt der Sphäre doppelt abgedeckt ist (und einige wenige sogar mehrfach). Insbesondere ist dann \mathbf{p} nicht mehr „fast injektiv“.
2. Um die Mantelfläche eines Kreiszylinders mit Radius $\rho > 0$ und Höhe $\eta > 0$ zu berechnen, betrachten wir die parametrisierte Fläche

$$\mathbf{p}(\varphi, h) = \begin{pmatrix} \rho \cos(\varphi) \\ \rho \sin(\varphi) \\ h \end{pmatrix}, \quad \varphi \in [0, 2\pi], \quad h \in [0, \eta],$$

wobei wir in diesem Beispiel die Flächenparameter nicht mit u_1 und u_2 , sondern mit φ (Winkel) und h (Höhe) bezeichnen wollen. Mit

$$\partial_\varphi \mathbf{p}(\varphi, h) \times \partial_h \mathbf{p}(\varphi, h) = \begin{pmatrix} -\rho \sin(\varphi) \\ +\rho \cos(\varphi) \\ 0 \end{pmatrix} \times \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \rho \cos(\varphi) \\ \rho \sin(\varphi) \\ 0 \end{pmatrix}$$

erhalten wir die aus der Schule bekannte Formel

$$\text{vol}(F) = \int_0^{2\pi} \int_0^\eta \rho \, dh \, d\varphi = 2\pi \rho \eta.$$

Bemerkungen:

- (a) Für die Gesamtoberfläche des Kreiszylinders müssen wir noch die Grund- und Deckfläche (jeweils $\pi\rho^2$) addieren, denn die entsprechenden Punkte werden ja nicht durch \mathbf{p} abgedeckt.
- (b) Auch hier ist \mathbf{p} wieder „fast injektiv“, denn die Mehrdeutigkeit betrifft nur das Geradenstück $\{(1, 0, h) : 0 \leq h \leq \eta\}$, die als echte eindimensionale Menge bei der Berechnung von Flächen keine Probleme bereitet.
- (c) Es gilt $do = \rho \, d\varphi \, dh$, wobei der Vorfaktor auf der rechten Seite zwar vom Radius ρ , aber nicht von den Flächenparametern φ und h abhängt.

Ausblick*: Dies ist eine sehr spezielle Eigenschaft des Kreiszylinders und hat damit zu tun, dass man diesen verzerrungsfrei in der Ebene abrollen kann. Mathematiker sagen deshalb, der Kreiszylinder besitzt zwar eine *extrinsische*, aber keine *intrinsische* Krümmung. Insbesondere wird eine

Ameise, die sich nur auf einem kleinen Stück des Kreiszyinders bewegt, den Eindruck haben, sie reise durch den flachen zweidimensionalen Raum. Auf der Sphäre ist dies wegen der vorhandenen intrinsischen Krümmung anders. Gauß konnte zum Beispiel die Krümmung der Erde (bzw. den Erdradius) durch lokale Messungen bestimmen, d.h. ohne um die ganze Welt reisen zu müssen.

3. Wir hatten schon weiter oben gesehen, dass der Graph einer skalaren Funktion $f : D \rightarrow \mathbb{R}$ als reguläre Fläche $F = \text{graph } f$ interpretiert und durch

$$\mathbf{p}(\mathbf{u}) = \begin{pmatrix} u_1 \\ u_2 \\ f(u_1, u_2) \end{pmatrix}, \quad \mathbf{u} \in D$$

parametrisiert werden kann. Wegen

$$\partial_{u_1} \mathbf{p}(\mathbf{u}) \times \partial_{u_2} \mathbf{p}(\mathbf{u}) = \begin{pmatrix} 1 \\ 0 \\ \partial_{u_1} f(\mathbf{u}) \end{pmatrix} \times \begin{pmatrix} 0 \\ 1 \\ \partial_{u_2} f(\mathbf{u}) \end{pmatrix} = \begin{pmatrix} -\partial_{u_1} f(\mathbf{u}) \\ -\partial_{u_2} f(\mathbf{u}) \\ 1 \end{pmatrix}$$

ergibt sich

$$\text{vol}(F) = \int_D \sqrt{1 + |\mathbf{grad} f(\mathbf{u})|^2} \, d\mathbf{u}$$

für den Flächeninhalt des Graphen von F . Im Spezialfall

$$D = \overline{B}_\varrho(\mathbf{0}) = \{(u_1, u_2) : u_1^2 + u_2^2 \leq \varrho^2\}, \quad f(u_1, u_2) = \sqrt{\varrho^2 - u_1^2 - u_2^2}$$

gilt

$$\mathbf{grad} f(u_1, u_2) = \frac{1}{\sqrt{\varrho^2 - u_1^2 - u_2^2}} \begin{pmatrix} -u_1 \\ -u_2 \end{pmatrix}$$

sowie

$$\sqrt{1 + |\mathbf{grad} f(\mathbf{u})|^2} = \sqrt{1 + \frac{u_1^2 + u_2^2}{\varrho^2 - u_1^2 - u_2^2}} = \frac{\varrho}{\sqrt{\varrho^2 - u_1^2 - u_2^2}},$$

und wir erhalten mit

$$\begin{aligned} \text{vol}(\text{graph}(f)) &= \int_D \frac{\varrho}{\sqrt{\varrho^2 - u_1^2 - u_2^2}} \, d\mathbf{u} \\ &= \int_0^\varrho \int_0^{2\pi} \frac{\varrho r}{\sqrt{\varrho^2 - r^2}} \, d\varphi \, dr \\ &= 2\pi \varrho \left[-\sqrt{\varrho^2 - r^2} \right]_{r=0}^{r=\varrho} = 2\pi \varrho^2 \end{aligned}$$

wieder den Wert von oben, wobei wir das \mathbf{u} -Integral diesmal mit Hilfe der ebenen Polarkoordinaten $u_1 = r \cos(\varphi)$, $u_2 = r \sin(\varphi)$ berechnet haben.

Bemerkungen:

- (a) Wir hätten in allen Formeln zu $\text{graph } f$ auch wieder (d.h. wie schon weiter oben) x_1 bzw. x_2 statt u_1 bzw. u_2 schreiben können.
- (b) Es ist kein Zufall, dass wir für den Flächeninhalt der Halbsphäre denselben Wert wie oben erhalten, denn das Flächenintegral ist *invariant unter Reparametrisierungen* (siehe das Theorem unten). Oder anders gesagt: Wir brauchen in der Regel eine Parametrisierung \mathbf{p} von F , um den Flächeninhalt von F überhaupt ausrechnen zu können. Es ist aber egal, welche Parametrisierung wir wählen (sie sollte nur injektiv sein und wirklich F beschreiben).
- (c) Die in diesem Beispiel verwendete Parametrisierung \mathbf{p} ist in Randpunkten von D *nicht* differenzierbar, aber wir haben keine Schwierigkeit, das Integral zu berechnen. Woran genau das liegt, können wir hier nicht diskutieren, wollen aber anmerken, dass die Nichtdifferenzierbarkeit in diesem Beispiel hinreichend harmlos ist (und das Integral streng genommen nur in einem uneigentlichen Sinn existiert). Im Allgemeinen müssen wir bei Singularitäten in den Ableitungen aber besonders aufpassen.

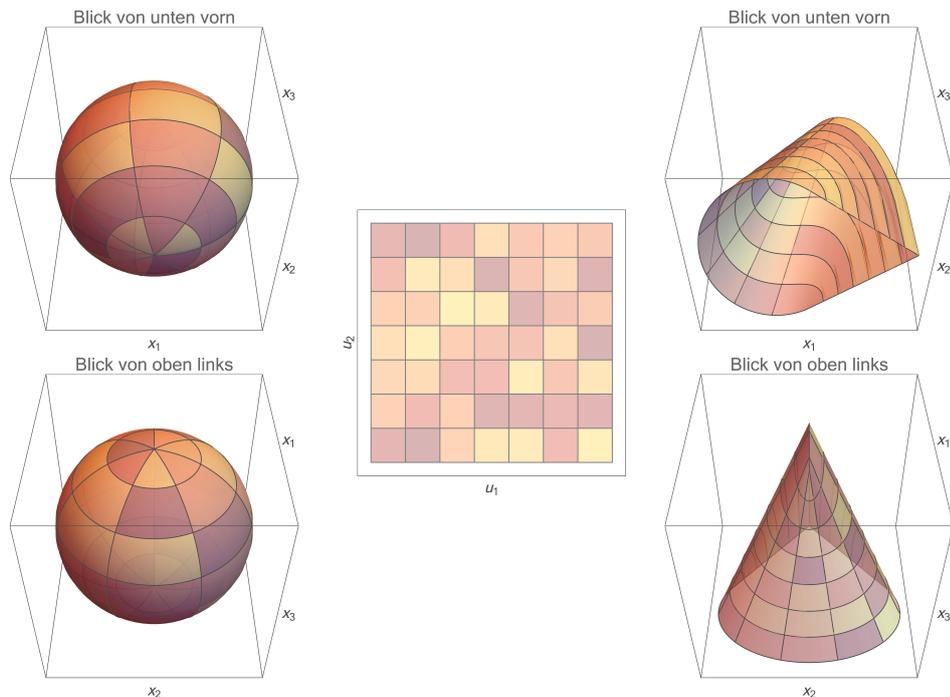


Abbildung Zur Interpretation des Flächenintegrals bzw. des infinitesimalen Flächenelements: Eine Zerlegung von D in gleich große Teile (Mitte) entspricht einer Zerlegung von F (links bzw. rechts), wobei aber die Teilflächen von F im Allgemeinen nicht mehr gleich groß sind. Der Term $|\partial_{u_1}\mathbf{p}(\mathbf{u}) \times \partial_{u_2}\mathbf{p}(\mathbf{u})|$ im Flächenintegral stellt sicher, dass diese Volumenänderung richtig berücksichtigt wird. Die Fläche rechts wird übrigens *Oloid* genannt. Sie kann vollständig abgerollt werden und spielt in Technischen Mechanik eine Rolle (siehe WIKIPEDIA).

Integrale von Funktionen über Flächen

Flächenintegrale der 1. Art und der 2. Art Ist $f : F \rightarrow \mathbb{R}$ eine hinreichend gute skalare Funktion auf F , so wird

$$\int_F f(\mathbf{x}) \, d\sigma := \int_D f(\mathbf{p}(\mathbf{u})) |\partial_{u_1}\mathbf{p}(\mathbf{u}) \times \partial_{u_2}\mathbf{p}(\mathbf{u})| \, d\mathbf{u}$$

als das Flächenintegral der 1. Art von f über F bezeichnet. Ist $\mathbf{f} : F \rightarrow \mathbb{R}^3$ ein Vektorfeld auf D , so nennt man

$$\int_F \mathbf{f}(\mathbf{x}) \cdot d\boldsymbol{\sigma} := \int_D \mathbf{f}(\mathbf{p}(\mathbf{u})) \cdot (\partial_{u_1}\mathbf{p}(\mathbf{u}) \times \partial_{u_2}\mathbf{p}(\mathbf{u})) \, d\mathbf{u}$$

das Oberflächenintegral der 2. Art von \mathbf{f} über F . Beachte, dass beide Integrale immer eine reelle Zahl (und nicht etwa einen Vektor liefern).

1. Flächenintegrale werden manchmal auch *Oberflächenintegrale* genannt. Es gibt darüberhinaus wieder verschiedene Notationen, aber die zu Grunde liegenden Ideen und Rechenschritte sind immer dieselben.
2. Die Funktionen f und \mathbf{f} werden meist nicht nur auf F , sondern auf einer größeren Teilmenge des \mathbb{R}^3 definiert sein. „Hinreichend gut“ meint letztlich, dass $f \circ \mathbf{p}$ bzw. die Komponenten von $\mathbf{f} \circ \mathbf{p}$ integrierbare Funktionen auf D sein müssen, wobei Stetigkeit wieder eine hinreichende Bedingung ist.
3. Unsere Definitionen können als die formalen Substitutionsregeln

$$d\sigma = |\partial_{u_1}\mathbf{p}(\mathbf{u}) \times \partial_{u_2}\mathbf{p}(\mathbf{u})| \, du_1 \, du_2$$

und

$$d\boldsymbol{\sigma} = (\partial_{u_1}\mathbf{p}(\mathbf{u}) \times \partial_{u_2}\mathbf{p}(\mathbf{u})) \, du_1 \, du_2$$

interpretiert werden, wobei $d\sigma$ bzw. $d\boldsymbol{\sigma}$ als eine „infinitesimale skalare“ bzw. als eine „infinitesimale vektorielle“ Größe zu betrachten ist. Die analogen Formeln für Kurvenintegrale sind $ds = |\dot{\mathbf{c}}(t)| \, dt$ und $d\mathbf{x} = \dot{\mathbf{c}}(t) \, dt$.

4. Es gilt

$$d\boldsymbol{\sigma} = \pm \boldsymbol{\nu}(\mathbf{p}(\mathbf{u})) \, d\sigma,$$

wobei

$$\boldsymbol{\nu}(\mathbf{p}(\mathbf{u})) = \pm \frac{\partial_{u_1}\mathbf{p}(\mathbf{u}) \times \partial_{u_2}\mathbf{p}(\mathbf{u})}{|\partial_{u_1}\mathbf{p}(\mathbf{u}) \times \partial_{u_2}\mathbf{p}(\mathbf{u})|}$$

den bereits oben eingeführten normierten Normalenvektor auf F bezeichnet, für den aber eine Vorzeichenwahl getroffen werden kann. Insbesondere ist jedes Flächenintegral der 2. Art via

$$\int_F \mathbf{f}(\mathbf{x}) \cdot d\boldsymbol{\sigma} = \pm \int_F \langle \mathbf{f}(\mathbf{x}), \boldsymbol{\nu}(\mathbf{x}) \rangle \, d\sigma$$

immer ein spezielles Flächenintegral der 1. Art, wobei der Wert von der Wahl des Vorzeichens für $\boldsymbol{\nu}$ abhängt.

Interpretation Das Flächenintegral der 1. Art ist die natürliche Verallgemeinerung der im letzten Kapitel studierten zweidimensionalen Gebietsintegrale, wobei der Definitionsbereich der Funktion f diesmal keine flache Teilmenge D des \mathbb{R}^2 sein muss, sondern eine gekrümmte Teilmenge F des \mathbb{R}^3 sein darf. Das Flächenintegral der 2. Art quantifiziert den *Durchfluss* des Vektorfeldes \mathbf{f} durch die Fläche, d.h. wir integrieren die Länge der *normalen Komponente* von \mathbf{f} über F .

Beispiele

1. Um die Funktion

$$f(\mathbf{x}) = x_1 x_2 + x_3$$

auf dem Kreiszylinder

$$F = \{\mathbf{x} : x_1^2 + x_2^2 = \varrho^2, 0 \leq x_3 \leq \eta\}$$

mit Radius ϱ und Höhe η zu integrieren, benutzen wir die Flächenformeln von oben mit den Parametern φ und h . Wegen $do = \varrho \, d\varphi \, dh$ erhalten wir

$$\begin{aligned} \int_F f(\mathbf{x}) \, do &= \int_0^{2\pi} \int_0^\eta f(\varrho \cos(\varphi), \varrho \sin(\varphi), h) \varrho \, dh \, d\varphi \\ &= \int_0^{2\pi} \int_0^\eta (\varrho^3 \cos(\varphi) \sin(\varphi) + \varrho h) \, dh \, d\varphi \\ &= \int_0^{2\pi} (\varrho^3 \eta \cos(\varphi) \sin(\varphi) + \frac{1}{2} \varrho \eta^2) \, d\varphi = \pi \varrho \eta^2, \end{aligned}$$

wobei wir auch $\int_0^{2\pi} \cos(\varphi) \sin(\varphi) \, d\varphi = \frac{1}{2} \int_0^{2\pi} \sin(2\varphi) \, d\varphi = 0$ benutzt haben.

2. Wir wollen das Vektorfeld

$$\mathbf{f}(\mathbf{x}) = \begin{pmatrix} x_1 x_2 \\ 0 \\ x_3 \end{pmatrix}$$

mit Hilfe der Euler-Winkel auf der halben Einheitskugel

$$F = \{\mathbf{x} : x_1^2 + x_2^2 + x_3^2 = 1, x_3 \geq 0\}$$

integrieren. Wir hatten dazu oben schon ausgerechnet, dass

$$d\boldsymbol{\sigma} = \begin{pmatrix} \cos(u_1) \cos^2(u_2) \\ \sin(u_1) \cos^2(u_2) \\ \sin(u_2) \cos(u_2) \end{pmatrix} du_1 \, du_2$$

gilt, wobei $u_1 \in [0, 2\pi]$ und $u_2 \in [0, \pi/2]$ die eulerschen Winkelparameter für F sind. Durch Einsetzen der Parametrisierung $\mathbf{x} = \mathbf{p}(\mathbf{u})$ in die Definition von \mathbf{f} erhalten wir

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{p}(\mathbf{u})) = \begin{pmatrix} \cos(u_1) \sin(u_1) \cos^2(u_2) \\ 0 \\ \sin(u_2) \end{pmatrix}$$

und damit

$$\mathbf{f}(\mathbf{x}) \cdot d\boldsymbol{\sigma} = (\sin(u_1) \cos^2(u_1) \cos^4(u_2) + \sin^2(u_2) \cos(u_2)) \, du_1 \, du_2$$

nach Berechnung des Skalarproduktes. Wir müssen nun die rechte Seite in dieser Formel bzgl. u_1 und u_2 integrieren. Mit dem Satz von Fubini und elementaren Rechenregeln erhalten wir

$$\int_0^{2\pi} \int_0^{\pi/2} \sin(u_1) \cos^2(u_1) \cos^4(u_2) du_2 du_1 = 0$$

sowie

$$\int_0^{2\pi} \int_0^{\pi/2} \sin^2(u_2) \cos(u_2) du_2 du_1 = 2\pi \int_0^{\pi/2} \sin^2(u_2) \cos(u_2) du_2 = \frac{2}{3}\pi.$$

Insgesamt ergibt sich

$$\int_F \mathbf{f}(\mathbf{x}) \cdot d\boldsymbol{\sigma} = \frac{2}{3}\pi$$

als der gesuchte Wert des Flächenintegrals der 2. Art von \mathbf{f} über F .

Bemerkungen*

1. Das infinitesimale Flächenelement kann alternativ auch durch

$$d\sigma = g(\mathbf{u}) du_1 du_2, \quad g(\mathbf{u}) = \sqrt{\det(\mathbf{G}(\mathbf{u}))}$$

bestimmt werden, wobei die symmetrische $(2, 2)$ -Matrix

$$\mathbf{G}(\mathbf{u}) = \mathbf{J}\mathbf{p}(\mathbf{u})^T \cdot \mathbf{J}\mathbf{p}(\mathbf{u}) = \begin{pmatrix} \langle \partial_{u_1}\mathbf{p}(\mathbf{u}), \partial_{u_1}\mathbf{p}(\mathbf{u}) \rangle & \langle \partial_{u_1}\mathbf{p}(\mathbf{u}), \partial_{u_2}\mathbf{p}(\mathbf{u}) \rangle \\ \langle \partial_{u_1}\mathbf{p}(\mathbf{u}), \partial_{u_2}\mathbf{p}(\mathbf{u}) \rangle & \langle \partial_{u_2}\mathbf{p}(\mathbf{u}), \partial_{u_2}\mathbf{p}(\mathbf{u}) \rangle \end{pmatrix}$$

auch *metrischer Tensor* oder *Erste Fundamentalform* genannt wird. Insbesondere können wir also an Stelle des Kreuzproduktes und seiner Norm auch die drei Skalarprodukte zwischen den Tangentialvektoren $\partial_{u_1}\mathbf{p}(\mathbf{u})$ und $\partial_{u_2}\mathbf{p}(\mathbf{u})$ sowie eine Determinante und eine Wurzel berechnen.

2. Im Spezialfall $p_3(\mathbf{u}) = 0$ für alle $\mathbf{u} \in D$ kann F selbst als Teilmenge des \mathbb{R}^2 betrachtet werden und kleinere Rechnungen zeigen

$$g(\mathbf{u}) = \left| \det \begin{pmatrix} \partial_{u_1}p_1(\mathbf{u}) & \partial_{u_2}p_1(\mathbf{u}) \\ \partial_{u_1}p_2(\mathbf{u}) & \partial_{u_2}p_2(\mathbf{u}) \end{pmatrix} \right| = |\partial_{u_1}p_1(\mathbf{u}) \partial_{u_2}p_2(\mathbf{u}) - \partial_{u_2}p_1(\mathbf{u}) \partial_{u_1}p_2(\mathbf{u})|.$$

Insbesondere wird dann aus der Definition des Flächenintegrals ein Spezialfall der Transformationsformel für Gebietsintegrale.

Theorem (Invarianz unter Reparametrisierung) Bei einem Wechsel der Parametrisierung von F ändert sich ein Flächenintegral der 1. Art nicht. Ein Flächenintegral der 2. Art kann höchstens sein Vorzeichen ändern.

Beweis Vorbemerkung: Wir betrachten zwei parametrisierte Flächen $\mathbf{p} : D \rightarrow \mathbb{R}^3$ und $\tilde{\mathbf{p}} : \tilde{D} \rightarrow \mathbb{R}^3$, sodass $F = \text{im}(\mathbf{p}) = \text{im}(\tilde{\mathbf{p}})$ sowie

$$\tilde{\mathbf{p}}(\tilde{\mathbf{u}}) = \mathbf{p}(\mathbf{u}), \quad \mathbf{u} = \Phi(\tilde{\mathbf{u}}), \quad \tilde{\mathbf{u}} = \tilde{\Phi}(\mathbf{u})$$

gelten, wobei die Abbildungen $\Phi : \tilde{D} \rightarrow D$ und $\tilde{\Phi} : D \rightarrow \tilde{D}$ die Parameterwechsel beschreiben und zueinander invers sind.

Teil 1: Die Kettenregel liefert

$$\mathbf{J}\tilde{\mathbf{p}}(\tilde{\mathbf{u}}) = \mathbf{J}\mathbf{p}(\mathbf{u}) \cdot \mathbf{J}\Phi(\tilde{\mathbf{u}}),$$

wobei dies auch als

$$\begin{pmatrix} \left. \begin{array}{c} \partial_{\tilde{u}_1} \tilde{\mathbf{p}}(\tilde{\mathbf{u}}) \\ \partial_{\tilde{u}_2} \tilde{\mathbf{p}}(\tilde{\mathbf{u}}) \end{array} \right| \end{pmatrix} = \begin{pmatrix} \left. \begin{array}{c} \partial_{u_1} \mathbf{p}(\mathbf{u}) \\ \partial_{u_2} \mathbf{p}(\mathbf{u}) \end{array} \right| \end{pmatrix} \cdot \begin{pmatrix} \left. \begin{array}{c} \partial_{\tilde{u}_1} \Phi_1(\tilde{\mathbf{u}}) \\ \partial_{\tilde{u}_1} \Phi_2(\tilde{\mathbf{u}}) \end{array} \right| \quad \left. \begin{array}{c} \partial_{\tilde{u}_2} \Phi_1(\tilde{\mathbf{u}}) \\ \partial_{\tilde{u}_2} \Phi_2(\tilde{\mathbf{u}}) \end{array} \right| \end{pmatrix}$$

geschrieben werden kann. Die Eigenschaften der Matrizenmultiplikation implizieren nun die Darstellungsformeln

$$\begin{aligned} \partial_{\tilde{u}_1} \tilde{\mathbf{p}}(\tilde{\mathbf{u}}) &= \partial_{\tilde{u}_1} \Phi_1(\tilde{\mathbf{u}}) \partial_{u_1} \mathbf{p}(\mathbf{u}) + \partial_{\tilde{u}_1} \Phi_2(\tilde{\mathbf{u}}) \partial_{u_2} \mathbf{p}(\mathbf{u}), \\ \partial_{\tilde{u}_2} \tilde{\mathbf{p}}(\tilde{\mathbf{u}}) &= \partial_{\tilde{u}_2} \Phi_1(\tilde{\mathbf{u}}) \partial_{u_1} \mathbf{p}(\mathbf{u}) + \partial_{\tilde{u}_2} \Phi_2(\tilde{\mathbf{u}}) \partial_{u_2} \mathbf{p}(\mathbf{u}), \end{aligned}$$

wobei $\partial_{\tilde{u}_j} \tilde{\mathbf{p}}(\tilde{\mathbf{u}})$ sowie $\partial_{u_j} \mathbf{p}(\mathbf{u})$ dreidimensionale Vektoren und $\partial_{\tilde{u}_j} \Phi_i(\tilde{\mathbf{u}})$ reelle Zahlen sind. Unter Ausnutzung der Bilinearität sowie der Antisymmetrie des Kreuzproduktes erhalten wir deshalb

$$\partial_{\tilde{u}_1} \tilde{\mathbf{p}}(\tilde{\mathbf{u}}) \times \partial_{\tilde{u}_2} \tilde{\mathbf{p}}(\tilde{\mathbf{u}}) = \left(\partial_{\tilde{u}_1} \Phi_1(\tilde{\mathbf{u}}) \partial_{\tilde{u}_2} \Phi_2(\tilde{\mathbf{u}}) - \partial_{\tilde{u}_1} \Phi_2(\tilde{\mathbf{u}}) \partial_{\tilde{u}_2} \Phi_1(\tilde{\mathbf{u}}) \right) \partial_{u_1} \mathbf{p}(\mathbf{u}) \times \partial_{u_2} \mathbf{p}(\mathbf{u})$$

und damit auch

$$\left| \partial_{\tilde{u}_1} \tilde{\mathbf{p}}(\tilde{\mathbf{u}}) \times \partial_{\tilde{u}_2} \tilde{\mathbf{p}}(\tilde{\mathbf{u}}) \right| = \left| \det \mathbf{J}\Phi(\tilde{\mathbf{u}}) \right| \left| \partial_{u_1} \mathbf{p}(\mathbf{u}) \times \partial_{u_2} \mathbf{p}(\mathbf{u}) \right|.$$

Teil 2: Die Transformationsformel für Gebietsintegrale impliziert

$$d\tilde{u}_1 d\tilde{u}_2 = \left| \det \mathbf{J}\tilde{\Phi}(\mathbf{u}) \right| du_1 du_2$$

und durch Kombination der bisherigen Teilresultate ergibt sich

$$d\tilde{o} = \left| \det \mathbf{J}\Phi(\tilde{\mathbf{u}}) \right| \left| \det \mathbf{J}\tilde{\Phi}(\mathbf{u}) \right| d\mathbf{o},$$

wobei $d\tilde{o}$ bzw. $d\mathbf{o}$ das infinitesimale Flächenelement bzgl. $\tilde{\mathbf{p}}$ bzw. \mathbf{p} meint.

Teil 3: Die Jacobi-Matrizen $\mathbf{J}\Phi(\tilde{\mathbf{u}})$ und $\mathbf{J}\tilde{\Phi}(\mathbf{u})$ sind zueinander invers (dies ist wieder eine Folge der Kettenregel) und ihre Determinanten müssen sich daher zu 1 multiplizieren. Insbesondere gilt

$$\int_{\tilde{D}} f(\tilde{\mathbf{p}}(\tilde{\mathbf{u}})) d\tilde{o} = \int_D f(\mathbf{p}(\mathbf{u})) d\mathbf{o}$$

für jede Funktion $f : F \rightarrow \mathbb{R}$, und wir haben damit die erste Behauptung bewiesen.

Schlussbemerkung: Wie wir gesehen haben, folgt die Invarianz des Integrals der 1. Art direkt aus der Kettenregel sowie der Transformationsformel für zweidimensionale Gebietsintegrale. Oder anders gesagt: Das Flächenintegral ist gerade so definiert, dass es invariant unter Reparametrisierung ist und damit ein wirkliches geometrisches Konzept darstellt. Die Argumente für Integrale der 2. Art sind analog und liefern

$$d\tilde{\sigma} = (\det \mathbf{J}\Phi(\tilde{\mathbf{u}})) |\det \mathbf{J}\tilde{\Phi}(\mathbf{u})| d\sigma = \pm d\sigma,$$

wobei das Vorzeichen $\operatorname{sgn}(\det \mathbf{J}\Phi(\tilde{\mathbf{u}})) = \operatorname{sgn}(\det \mathbf{J}\tilde{\Phi}(\mathbf{u}))$ darüber entscheidet, ob die Formel mit $+$ oder $-$ zu benutzen ist. \square

3.6 Integralsätze von Gauß und Stokes in 3D

Vorlesung 11-3: 01. Februar

Ziel Wir verallgemeinern die Sätze von Gauß und Stokes von 2D auf 3D. Die geometrische Interpretation und die physikalische Bedeutung beider Theoreme wird sich dabei nicht ändern, aber aus dem Anwachsen der Raumdimension ergibt sich ein wesentlicher Unterschied:

1. Der Satz von Gauß wird diesmal ein dreidimensionales Gebietsintegral mit einem zweidimensionalen Oberflächenintegral in Beziehung setzen.
2. Der Satz von Stokes wird auch in 3D ein zweidimensionales Flächenintegral mit einem eindimensionalen Kurvenintegral in Beziehung setzen.

Formulierung des Satzes von Gauß

Setting In diesem Abschnitt betrachten wir eine Menge $G \subset \mathbb{R}^3$, die durch endlich viele Flächen berandet sind, wobei wir die Anzahl der Randflächen mit K bezeichnen, sowie Vektorfelder auf G .

1. $G \subset \mathbb{R}^3$ ist eine messbare Menge des \mathbb{R}^3 .
2. Es existieren K parametrisierte Flächen $\mathbf{p}_k : D_k \rightarrow \mathbb{R}^3$, so dass

$$\partial G = \bigcup_{k=1}^K F_k, \quad F_k = \operatorname{im}(\mathbf{p}_k).$$

3. Wir nehmen außerdem an, dass sich die Randmengen F_k nur in endlich vielen Kurven überlappen, da wir andernfalls wesentliche Teile von ∂G doppelt abdecken würden. Diese Kurven können, aber müssen keine Knickkanten sein.
4. In jedem Punkt $\mathbf{x} \in \partial G$ (mit Ausnahme von endlich vielen Knickkurven oder Eckpunkten) existiert ein normierter äußerer Normalenvektor $\boldsymbol{\nu} : \partial G \rightarrow \mathbb{R}^3$, sodass
 - (a) $|\boldsymbol{\nu}(\mathbf{x})| = 1$,
 - (b) $\boldsymbol{\nu}(\mathbf{x})$ steht senkrecht auf jedem Flächentangententialvektor in \mathbf{x} .
 - (c) $\boldsymbol{\nu}(\mathbf{x})$ zeigt aus Sicht von G immer nach außen.
5. $\mathbf{f} : G \rightarrow \mathbb{R}$ ist ein stetig differenzierbares Vektorfeld.

Beispiele

1. Der Würfel $G = [a_1, b_1] \times [a_2, b_2] \times [a_3, b_3]$ besitzt die 6 Randflächen

$$F_h, \quad F_v, \quad F_l, \quad F_r, \quad F_u, \quad F_o$$

deren Indizes für „hinten“, „vorne“, „links“, „rechts“, „unten“ und „oben“ stehen. Diese Teilflächen werden in naheliegender Weise über den 2D-Quadern

$$D_{h/v} = [a_2, b_2] \times [a_3, b_3], \quad D_{l/r} = [a_1, b_1] \times [a_3, b_3], \quad D_{u/o} = [a_1, b_1] \times [a_2, b_2]$$

durch die Formeln

$$\mathbf{p}_h(u_2, u_3) = \begin{pmatrix} a_1 \\ u_2 \\ u_3 \end{pmatrix}, \quad \mathbf{p}_l(u_1, u_3) = \begin{pmatrix} u_1 \\ a_2 \\ u_3 \end{pmatrix}, \quad \mathbf{p}_u(u_1, u_2) = \begin{pmatrix} u_1 \\ u_2 \\ a_3 \end{pmatrix}$$

und

$$\mathbf{p}_v(u_2, u_3) = \begin{pmatrix} b_1 \\ u_2 \\ u_3 \end{pmatrix}, \quad \mathbf{p}_r(u_1, u_3) = \begin{pmatrix} u_1 \\ b_2 \\ u_3 \end{pmatrix}, \quad \mathbf{p}_o(u_1, u_2) = \begin{pmatrix} u_1 \\ u_2 \\ b_3 \end{pmatrix}$$

parametrisiert, wobei immer $u_j \in [a_j, b_j]$ gilt. Die entsprechenden äußeren Normalenvektoren der Teilränder sind durch

$$\boldsymbol{\nu}_h = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}, \quad \boldsymbol{\nu}_l = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}, \quad \boldsymbol{\nu}_u = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}$$

sowie

$$\boldsymbol{\nu}_v = \begin{pmatrix} +1 \\ 0 \\ 0 \end{pmatrix}, \quad \boldsymbol{\nu}_r = \begin{pmatrix} 0 \\ +1 \\ 0 \end{pmatrix}, \quad \boldsymbol{\nu}_o = \begin{pmatrix} 0 \\ 0 \\ +1 \end{pmatrix}$$

gegeben und hängen nicht von \mathbf{x} ab.

2. Die Vollkugel

$$G = \overline{B}_\varrho(\mathbf{0}) = \{\mathbf{x} : x_1^2 + x_2^2 + x_3^2 \leq \varrho^2\}$$

besitzt den äußeren Normalenvektor

$$\boldsymbol{\nu}(\mathbf{x}) = +\frac{\mathbf{x}}{|\mathbf{x}|} \quad \text{für alle } \mathbf{x} \in \partial G = \{\mathbf{x} : |\mathbf{x}| = \varrho\}.$$

Wir können den Rand ∂G durch eine einzelne Parametrisierung (also mit $K = 1$) beschreiben, indem wir zum Beispiel die bereits oben eingeführten Euler-Winkel mit $D = [0, 2\pi] \times [-\pi/2, +\pi/2]$ verwenden. Alternativ können wir $K = 2$ setzen und ∂G durch die zwei Parametrisierungen

$$\mathbf{p}_\pm(u_1, u_2) = \begin{pmatrix} u_1 \\ u_2 \\ \pm\sqrt{\varrho^2 - u_1^2 - u_2^2} \end{pmatrix}, \quad \mathbf{u} \in D = \{\mathbf{u} : u_1^2 + u_2^2 \leq \varrho^2\}$$

beschreiben, die wir in natürlicher Weise mit „–“ bzw. „+“ (und nicht mit 1 bzw. 2) indizieren, und die die Vollsphäre als Vereinigung zweier Graphen über einer zweidimensionalen Kreisscheibe darstellen. Diese beiden Graphen entsprechen der nördlichen bzw. der südlichen Hemisphäre und überlappen sich gerade auf der Äquatorlinie

$$\{(x_1, x_2, x_3) : x_1^2 + x_2^2 = \varrho^2, x_3 = 0\}.$$

Es handelt sich dabei aber *nicht* um eine Knickkurve.

3. Ein *Volltorus* mit den Radien $0 < \varrho_1 < \varrho_2$ besteht aus allen Punkten des \mathbb{R}^3 , deren Abstand von einer gegebenen Kreislinie mit Radius ϱ_2 den Wert ϱ_1 nicht überschreitet. Ein achsenparalleles Beispiel wird durch die Formel

$$G = \left\{ (x_1, x_2, x_3) : \left(\varrho_2 - \sqrt{x_1^2 + x_2^2} \right)^2 + x_3^2 = \varrho_1^2 \right\}$$

beschrieben, wobei

$$K = \left\{ (x_1, x_2, x_3) : x_1^2 + x_2^2 = \varrho_2^2, x_3 = 0 \right\}$$

die zu Grunde liegende Kreislinie vom Radius ϱ_2 ist und in der (x_1, x_2) -Ebene liegt. Der Rand ∂G ist ein Torus, der durch die schon weiter oben eingeführte Formel

$$\mathbf{p}(\mathbf{u}) = \begin{pmatrix} \cos(u_1) (\varrho_2 + \varrho_1 \cos(u_2)) \\ \sin(u_1) (\varrho_2 + \varrho_1 \cos(u_2)) \\ \varrho_1 \sin(u_2) \end{pmatrix}, \quad u_1 \in [0, 2\pi], \quad u_2 \in [0, 2\pi]$$

in „fast injektiver“ Weise parametrisiert werden kann, und für den durch

$$\boldsymbol{\nu}(\mathbf{p}(\mathbf{u})) = + \frac{\partial_{u_1} \mathbf{p}(\mathbf{u}) \times \partial_{u_2} \mathbf{p}(\mathbf{u})}{|\partial_{u_1} \mathbf{p}(\mathbf{u}) \times \partial_{u_2} \mathbf{p}(\mathbf{u})|}$$

ein globaler äußerer Normalenvektor definiert wird. Direkte Rechnungen liefern

$$\partial_{u_1} \mathbf{p}(\mathbf{u}) \times \partial_{u_2} \mathbf{p}(\mathbf{u}) = \begin{pmatrix} \varrho_1 \cos(u_1) \cos(u_2) (\varrho_2 + \varrho_1 \cos(u_2)) \\ \varrho_1 \sin(u_1) \cos(u_2) (\varrho_2 + \varrho_1 \cos(u_2)) \\ \varrho_1 \sin(u_2) (\varrho_2 + \varrho_1 \cos(u_2)) \end{pmatrix}$$

und damit auch

$$|\partial_{u_1} \mathbf{p}(\mathbf{u}) \times \partial_{u_2} \mathbf{p}(\mathbf{u})| = \varrho_1 (\varrho_2 + \varrho_1 \cos(u_2))$$

bzw.

$$\boldsymbol{\nu}(\mathbf{p}(\mathbf{u})) = \begin{pmatrix} \cos(u_1) \cos(u_2) \\ \sin(u_1) \cos(u_2) \\ \sin(u_2) \end{pmatrix}.$$

Bemerkung: Die richtige Vorzeichenwahl in der Formel für $\boldsymbol{\nu}(\mathbf{p}(\mathbf{u}))$ kann übrigens recht einfach in einem festen Punkt verifiziert werden. Mit

$$\mathbf{p}(\mathbf{0}) = \begin{pmatrix} \varrho_1 + \varrho_2 \\ 0 \\ 0 \end{pmatrix}, \quad \partial_{u_1} \mathbf{p}(\mathbf{0}) \times \partial_{u_2} \mathbf{p}(\mathbf{0}) = \begin{pmatrix} \varrho_1 \\ 0 \\ 0 \end{pmatrix}$$

sehen wir zum Beispiel, dass unsere Vorzeichenwahl sicherstellt, dass $\boldsymbol{\nu}(\mathbf{p}(\mathbf{0}))$ im Flächenpunkt $\mathbf{p}(\mathbf{0})$ wirklich nach außen zeigt.

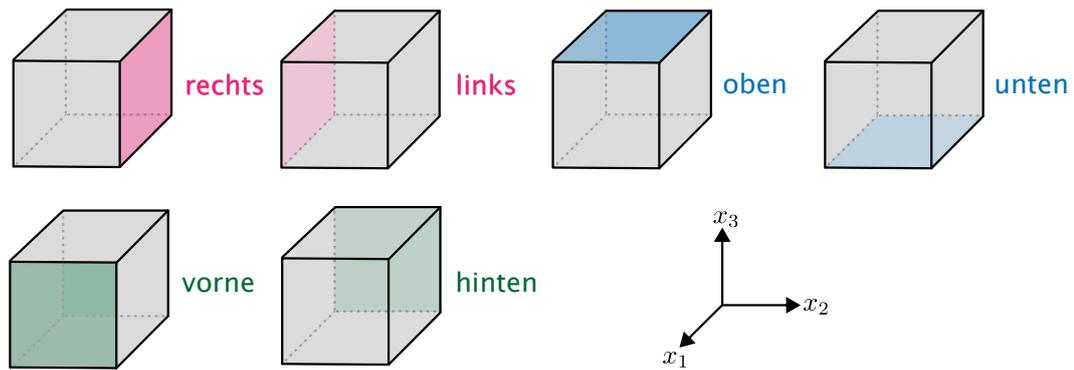


Abbildung Ein dreidimensionaler, achsenparalleler Quader besitzt 6 Randflächen, wobei der äußere Normalenvektor auf den 12 Kanten und in den 8 Eckpunkten nicht definiert ist, sonst aber immer in Richtung einer der Koordinatenachsen zeigt.

Theorem (Integralsatz von Gauß in 3D) Mit den obigen Notationen und Annahmen gilt

$$\int_G \operatorname{div} \mathbf{f}(\mathbf{x}) \, d\mathbf{x} = \sum_{k=1}^K \int_{F_k} \langle \mathbf{f}(\mathbf{x}), \boldsymbol{\nu}(\mathbf{x}) \rangle \, d\sigma,$$

wobei links das dreidimensionale Gebietsintegral der Divergenz von \mathbf{f} steht und rechts die zweidimensionalen Flächenintegrale über die Teilränder von G auftauchen.

Beweisidee Man beweist den Satz von Gauß analog zu 2D in drei Schritten:

1. Auf einem Quader kann die Formel einfach nachgerechnet werden (Hausaufgabe).
2. Für eine Vereinigung endlich vieler Quader kann die Formel zunächst auf jedem Teilquader separat ausgewertet werden. Werden die Teilergebnisse anschließend zusammengesetzt, so treten Auslöschungseffekte auf allen Rechtecken auf, die den gemeinsamen Rand zweier benachbarter Teilquader darstellen, und dies impliziert wieder die Behauptung. Siehe dazu den 2D-Beweis sowie die entsprechenden Bilder.
3. Allgemeinere Mengen G werden durch Quaderzerlegungen approximiert.

Die Schwierigkeit besteht darin, die Approximationsfehler im dritten Beweisschritt abzuschätzen. \square

Interpretation Analog zum zweidimensionalen Fall gilt: Das Flächenintegral auf der rechten Seite beschreibt den *effektiven Ausfluss* — oder, je nach Kontext, den *effektiven Durchfluss* — des Vektorfeldes über die Oberfläche ∂G , wobei in jedem Randpunkt $\mathbf{x} \in \partial G$ der Term $\langle \mathbf{f}(\mathbf{x}), \boldsymbol{\nu}(\mathbf{x}) \rangle$ die vorzeichenbehaftete Länge der Normalkomponente von $\mathbf{f}(\mathbf{x})$ liefert. Ein positiver bzw. negativer Beitrag meint dabei, dass $\mathbf{f}(\mathbf{x})$ in diesem Punkt nach außen bzw. nach innen zeigt, wohingegen ein verschwindender Beitrag impliziert, dass $\mathbf{f}(\mathbf{x})$ im linearen Tangentialraum $T_{\mathbf{x}}\partial G$ liegt. Die tiefe Erkenntnis ist nun, dass dieser effektive Ausfluss auch durch das Gebietsintegral von $\operatorname{div} \mathbf{f}$ über G beschrieben werden kann.

Bemerkungen

1. Die rechte Seite kann auch als

$$\int_{\partial G} \langle \mathbf{f}(\mathbf{x}), \boldsymbol{\nu}(\mathbf{x}) \rangle d\sigma$$

geschrieben werden, aber für $K > 1$ müssen wir dieses Oberflächenintegral als Summe von K Summanden auswerten, wobei der k -te Summand mit Hilfe der Parametrisierung \mathbf{p}_k berechnet wird.

2. Es gilt

$$\int_{F_k} \langle \mathbf{f}(\mathbf{x}), \boldsymbol{\nu}(\mathbf{x}) \rangle d\sigma = \pm \int_{D_k} \mathbf{f}(\mathbf{x}) \cdot d\boldsymbol{\sigma},$$

d.h. wir können das Oberflächenintegral über den k -ten Teilrand von G auch als Flächenintegral der 2. Art berechnen. Allerdings muss man sich dann immer genau überlegen, welches Vorzeichen das Richtige ist, d.h. ob bei der gegebenen Parametrisierung \mathbf{p}_k von F_k der Vektor $\partial_{u_1}\mathbf{p}_k(\mathbf{u}) \times \partial_{u_2}\mathbf{p}_k(\mathbf{u})$ aus Sicht von G nach außen oder nach innen zeigt.

3. Ein wichtiger Spezialfall betrifft Vektorfelder, deren j -Komponente durch eine skalare Funktion $f : G \rightarrow \mathbb{R}$ gegeben ist und deren restlichen Komponenten verschwinden (also $f_j = f$ und $f_i = 0$ für $i \neq j$). In diesem Fall reduziert sich der Gaußsche Satz auf die Formel

$$\int_G \partial_{x_j} f(\mathbf{x}) d\mathbf{x} = \int_{\partial G} f(\mathbf{x}) \nu_j(\mathbf{x}) d\sigma,$$

die als Verallgemeinerung des Hauptsatzes der Differential- und Integralrechnung verstanden werden kann. Dabei bezeichnet ν_j die j -te Komponente von $\boldsymbol{\nu}$.

4. Wird diese Formel für das Produkt zweier Funktionen $f, g : G \rightarrow \mathbb{R}$ ausgewertet, so ergibt sich mit

$$\int_G f(\mathbf{x}) \partial_{x_j} g(\mathbf{x}) d\mathbf{x} = - \int_G \partial_{x_j} f(\mathbf{x}) g(\mathbf{x}) d\mathbf{x} + \int_{\partial G} f(\mathbf{x}) g(\mathbf{x}) \nu_j(\mathbf{x}) d\sigma,$$

ein Analogon zur partiellen Integration.

Anwendungen des Satzes von Gauß

Beispiele

1. Wir wollen für das lineare Vektorfeld $\mathbf{f} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ mit

$$\mathbf{f}(\mathbf{x}) = \mathbf{S} \cdot \mathbf{x}$$

den effektiven Durchfluss durch den Quader $G = [a_1, b_1] \times [a_2, b_2] \times [a_3, b_3]$ berechnen, wobei

$$\mathbf{S} = \begin{pmatrix} s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{23} \\ s_{31} & s_{32} & s_{33} \end{pmatrix}$$

eine gegebene $(3, 3)$ -Matrix ist. Um den gesuchten Wert zu bestimmen, können wir auf den sechs zweidimensionalen Randquadern jeweils das entsprechende Flächenintegral berechnen und anschließend alle Teilergebnisse aufsummieren. Es ist aber einfacher, den gesuchten Wert mit Hilfe des Satzes von Gauß durch ein einziges Gebietsintegral zu bestimmen. Insbesondere gilt

$$\operatorname{div} \mathbf{f}(\mathbf{x}) = s_{11} + s_{22} + s_{33} = \operatorname{tr}(\mathbf{S})$$

und damit

$$\begin{aligned} \int_{\partial G} \langle \mathbf{f}(\mathbf{x}), \boldsymbol{\nu}(\mathbf{x}) \rangle d\sigma &= \operatorname{tr}(\mathbf{S}) \cdot \operatorname{vol}(G) \\ &= (s_{11} + s_{22} + s_{33}) \cdot (b_1 - a_1) \cdot (b_2 - a_2) \cdot (b_3 - a_3), \end{aligned}$$

wobei hier die Berechnung des Gebietsintegrals besonders einfach war, da $\operatorname{div} \mathbf{f}(\mathbf{x})$ gar nicht von \mathbf{x} abhängt und $\operatorname{vol}(G)$ bekannt ist.

2. Wir betrachten das dreidimensionale rotationssymmetrische Zentralfeld

$$\mathbf{f}(\mathbf{x}) = \phi\left(\sqrt{x_1^2 + x_2^2 + x_3^2}\right) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad \mathbf{x} \neq 0$$

auf der Hohlkugel

$$G = \{\mathbf{x} \in \mathbb{R}^3 : \varrho_i \leq |\mathbf{x}| \leq \varrho_a\},$$

wobei $0 < \varrho_i < \varrho_a < \infty$ gilt und $\phi : (0, \infty) \rightarrow \mathbb{R}$ eine gegebene und stetig differenzierbare Funktion ist. Der Rand von G besteht offensichtlich aus zwei Teilen, nämlich der Innensphäre bzw. Außensphäre

$$F_i = \{\mathbf{x} \in \mathbb{R}^3 : |\mathbf{x}| = \varrho_i\} \quad \text{bzw.} \quad F_a = \{\mathbf{x} \in \mathbb{R}^3 : |\mathbf{x}| = \varrho_a\}$$

und es gilt

$$\boldsymbol{\nu}(\mathbf{x}) = -\frac{\mathbf{x}}{|\mathbf{x}|} \quad \text{für} \quad \mathbf{x} \in F_i, \quad \boldsymbol{\nu}(\mathbf{x}) = +\frac{\mathbf{x}}{|\mathbf{x}|} \quad \text{für} \quad \mathbf{x} \in F_a.$$

Wir wollen nun die beiden Oberflächenintegrale im Gaußschen Satz berechnen. In diesem konkreten Beispiel müssen wir gar keine Parametrisierung für ∂G wählen, denn das Skalarprodukt von $\mathbf{f}(\mathbf{x})$ und $\boldsymbol{\nu}(\mathbf{x})$ hängt sowohl auf F_i als auch auf F_a nicht von \mathbf{x} ab. Insbesondere ergibt sich

$$\int_{F_i} \langle \mathbf{f}(\mathbf{x}), \boldsymbol{\nu}(\mathbf{x}) \rangle d\sigma = - \int_{F_i} \phi(\varrho_i) \varrho_i d\sigma = -\phi(\varrho_i) \varrho_i \operatorname{vol}(F_i) = 4\pi \phi(\varrho_i) \varrho_i^3$$

sowie eine analoge Formel mit vertauschtem Vorzeichen und ϱ_a statt ϱ_i für F_a . Durch Addition erhalten wir schließlich

$$\int_{\partial G} \langle \mathbf{f}(\mathbf{x}), \boldsymbol{\nu}(\mathbf{x}) \rangle d\sigma = 4\pi (\phi(\varrho_a) \varrho_a^3 - \phi(\varrho_i) \varrho_i^3).$$

Mit dem Satz von Gauß können wir nun schließen, dass das Vektorfeld nur im Fall von

$$\phi(r) = \frac{c}{r^3}$$

divergenzfrei ist, da es andernfalls einen effektiven Durchfluss durch G geben würde. Insbesondere kann nur in diesem Fall \mathbf{f} als Geschwindigkeitsfeld einer inkompressiblen Flüssigkeit verstanden werden. Natürlich könnten wir dies auch aus den Formeln

$$\partial_{x_j} \mathbf{f}(\mathbf{x}) = \phi(|\mathbf{x}|) + \phi'(|\mathbf{x}|) x_j^2, \quad \operatorname{div} \mathbf{f}(\mathbf{x}) = 3\phi(|\mathbf{x}|) + \phi'(|\mathbf{x}|) |\mathbf{x}|^2$$

sowie einem Differentialgleichungsargument ableiten.

Theorem (Greensche Formeln) Für alle hinreichend guten skalaren Funktionen $f, g : G \rightarrow \mathbb{R}$ gilt

$$\int_G (f(\mathbf{x}) \Delta g(\mathbf{x}) + \langle \mathbf{grad} f(\mathbf{x}), \mathbf{grad} g(\mathbf{x}) \rangle) d\mathbf{x} = \int_{\partial G} f(\mathbf{x}) \langle \mathbf{grad} g(\mathbf{x}), \boldsymbol{\nu}(\mathbf{x}) \rangle d\sigma$$

sowie

$$\int_G (f(\mathbf{x}) \Delta g(\mathbf{x}) - g(\mathbf{x}) \Delta f(\mathbf{x})) d\mathbf{x} = \int_{\partial G} (f(\mathbf{x}) \langle \mathbf{grad} g(\mathbf{x}), \boldsymbol{\nu}(\mathbf{x}) \rangle - g(\mathbf{x}) \langle \mathbf{grad} f(\mathbf{x}), \boldsymbol{\nu}(\mathbf{x}) \rangle) d\sigma,$$

wobei das Oberflächenintegral über ∂G im Fall von $K > 1$ wieder als Summe von Teilintegralen auszuwerten ist.

Beweis Die erste Greensche Formel ist gerade der Gaußsche Integralsatz für das Vektorfeld

$$\mathbf{f}(\mathbf{x}) = f(\mathbf{x}) \mathbf{grad} g(\mathbf{x}),$$

wobei

$$\operatorname{div} \mathbf{f}(\mathbf{x}) = \langle \mathbf{grad} f(\mathbf{x}), \mathbf{grad} g(\mathbf{x}) \rangle + f(\mathbf{x}) \Delta g(\mathbf{x})$$

aus der Produktregel für Ableitungen folgt und komponentenweise nachgerechnet werden kann. Durch Vertauschung von f und g erhalten wir eine Variante der ersten Greenschen Formel und wenn wir diese von der Originalvariante abziehen, ergibt sich gerade die zweite Greensche Formel. \square

Bemerkungen

1. Die Greenschen Formeln sehen sehr kompliziert aus, aber sie werden in der Mathematik und in den Anwendungswissenschaften sehr oft benutzt. Am besten ist, Sie memorieren nicht die Formeln, sondern wie diese aus dem Gaußschen Integralsatz abgeleitet werden können. Dieses Wissen könnte auch in einer Klausur hilfreich sein.

2. Δ ist der Laplace-Operator mit

$$\Delta g(\mathbf{x}) = \operatorname{div} \mathbf{grad} g(\mathbf{x}) = \partial_{x_1}^2 g(\mathbf{x}) + \partial_{x_2}^2 g(\mathbf{x}) + \partial_{x_3}^2 g(\mathbf{x}).$$

Er spielt zum Beispiel in der Theorie der partiellen Differentialgleichungen sowie in der Maxwell'schen Theorie des Elektromagnetismus eine wichtige Rolle.

3. Die Greenschen Formeln können als dreidimensionales Analogon der partiellen Integration verstanden werden. Die entsprechenden eindimensionalen Formeln auf einem Intervall $I = [a, b]$ sind

$$\int_a^b (f(x) g''(x) + f'(x) g'(x)) dx = [f(x) g'(x)]_{x=a}^{x=b}$$

und

$$\int_a^b (f(x) g''(x) - f''(x) g(x)) dx = [f(x) g'(x) - f'(x) g(x)]_{x=a}^{x=b},$$

wobei die rechten Seiten das Analogon zu den Randintegralen darstellen und auf dem Rand von I , d.h. in den Punkten $x = a$ und $x = b$ ausgewertet werden.

4. In der Literatur existieren die alternativen Notationen

$$\langle \boldsymbol{\nu}(\mathbf{x}), \mathbf{grad} g(\mathbf{x}) \rangle = \partial_{\boldsymbol{\nu}(\mathbf{x})} g(\mathbf{x}) = \frac{\partial g}{\partial \boldsymbol{\nu}}(\mathbf{x}),$$

wobei jeder Term die *Richtungsableitung von g in Richtung des Normalenvektors* beschreibt.

Vorlesung 11-4: 3. Februar

Der Satz von Stokes

Setting In diesem Abschnitt betrachten wir dreidimensionale Vektorfelder auf einer gekrümmten Fläche, die von einer *geschlossenen* Raumkurve berandet wird:

1. $\mathbf{p} : D \rightarrow \mathbb{R}^3$ ist eine Parametrisierung von $F = \operatorname{im}(\mathbf{p}) \subset \mathbb{R}^3$, wobei $D \subset \mathbb{R}^2$ gilt.
2. Der Flächenrand von F ist das Bild $C = \operatorname{im}(\mathbf{c})$ einer parametrisierten Raumkurve $\mathbf{c} : I \rightarrow \mathbb{R}^3$, wobei diese geschlossen ist und keine Doppelpunkte aufweist.
3. F ist orientierbar, d.h. es gibt zwei Seiten von F . Diese sehr wichtige geometrische Bedingung ist nicht verhandelbar und wird gleich genauer diskutiert.
4. Es gibt einen normierten Tangentialvektor $\boldsymbol{\tau}$ auf C und einen normierten Normalenvektor $\boldsymbol{\nu}$ auf F , die zueinander positiv orientiert sind. Es gibt mehrere Möglichkeiten, dies zu beschreiben:
 - (a) Die Kurve \mathbf{c} umrundet die Flächennormalenvektoren in mathematisch positiver Weise.
 - (b) Die Parametrisierung \mathbf{c} von C erfolgt so, dass F immer „links“ von C liegt, wobei dann stillschweigend vorausgesetzt wird, dass $\boldsymbol{\nu}$ nach „oben“ zeigt.

- (c) $\boldsymbol{\nu}$ und $\boldsymbol{\tau}$ beschreiben eine *Rechtsschraubung*. Siehe dazu das Bild.
 (d) In jedem Punkt $\mathbf{x} \in C$ zeigt der der Vektor $\boldsymbol{\nu}(\mathbf{x}) \times \boldsymbol{\tau}(\mathbf{x})$ in das Innere von F .

Mit etwas Übung kann man die richtige Orientierung in konkreten Fällen meist sehr einfach von der falschen unterscheiden.

5. \mathbf{f} ist ein stetig differenzierbares 3D-Vektorfeld, dass in der Nähe von F definiert ist.

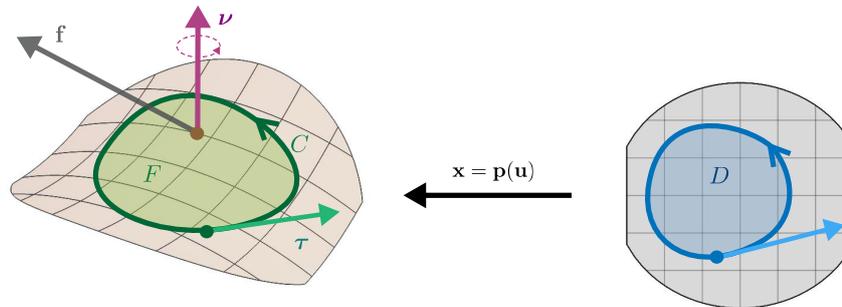


Abbildung Die typische Situation für den Integralsatz von Stokes mit einer gekrümmten Fläche F (hellgrün), deren Flächenrand C (dunkelgrün) als parametrisierte Raumkurve beschrieben werden kann. Der grüne Pfeil illustriert den Tangentialvektor $\boldsymbol{\tau}(\mathbf{x})$, den es in jedem Punkt $\mathbf{x} \in C$ gibt, wohingegen der violette Pfeil den Normalenvektor $\boldsymbol{\nu}(\mathbf{x})$ an die Fläche darstellt, der in jedem Punkt $\mathbf{x} \in F$ existiert. Rechts ist in blau die entsprechende planare Konfiguration gezeichnet, die mittels \mathbf{p} in die räumliche überführt wird. Die braune und die hellgraue Menge wurden nur zur besseren Anschauung gezeichnet und der dunkelgraue Pfeil repräsentiert das Vektorfeld \mathbf{f} . *Zur Orientierung:* Die Vektoren $\boldsymbol{\nu}$ und $\boldsymbol{\tau}$ sind zueinander positiv orientiert. Richten wir zum Beispiel eine Baumarkt-Schraube entlang des violetten Normalenvektors so aus, dass der Pfeil zur Spitze der Schraube zeigt, so entspricht sie Parametrisierung von C dem *Eindrehen der Schraube*.

Bemerkung

1. Die *positive* (also *richtige*) Orientierung von $\boldsymbol{\nu}$ und $\boldsymbol{\tau}$ ist insofern wichtig, als das wir bei falscher Wahl des Flächennormalenvektors oder des Durchlaufsinns von C das linke oder das rechte Vorzeichen im Satz von Stokes ändern müssten. Eine ähnliche Bedingung hatten wir schon in 2D kennen gelernt.
2. Die Bedingung der *Orientierbarkeit* hat keine Entsprechung in 2D. Sie stellt sicher, dass man überhaupt von richtiger oder falscher Orientierung reden kann.
3. Wir wollen immer annehmen, dass \mathbf{p} und \mathbf{c} jeweils stetig differenzierbar sowie regulär sind. Im Prinzip können F bzw. C Knickkurven bzw. Knickpunkte aufweisen, aber wir werden hier von dieser Möglichkeit keinen Gebrauch machen.
4. Der Flächenrand von F kann in aller Regel intuitiv verstanden bzw. identifiziert werden. Formal handelt es sich um die Menge $\mathbf{p}(\partial D)$, also um das Bild von ∂D unter der Parametrisierungsabbildung \mathbf{p} . Beachte, dass bei nicht-regulären Parametrisierungen gewisse Entartungen auftreten können, die wir aber hier nicht diskutieren wollen.
5. In jedem Punkt $\mathbf{x} \in C$ stehen die Vektoren $\boldsymbol{\nu}(\mathbf{x})$ und $\boldsymbol{\tau}(\mathbf{x})$ senkrecht aufeinander und besitzen jeweils die Länge 1. Die Vektoren $\boldsymbol{\tau}(\mathbf{x})$ und $\boldsymbol{\nu}(\mathbf{x}) \times \boldsymbol{\tau}(\mathbf{x})$ spannen dabei den Tangentialraum $T_{\mathbf{x}}F$ auf.

Orientierbarkeit von Flächen Die Frage, ob eine gegebene (unparametrisierte) Fläche $F \subset \mathbb{R}^3$ eigentlich immer „zwei Seiten“ besitzt, ist viel subtiler als sie auf den ersten Blick aussieht. Die erste spontane Antwort ist sicherlich ja, denn wir können uns zum Beispiel sehr gut vorstellen, dass eine Kugel oder die Oberfläche eines Kreiszylinders von innen grün und von außen blau angestrichen wird. Bei genauerer Betrachtung wird aber klar, dass dies nicht immer so ist. Als Standardbeispiele betrachten wir die parametrisierten Flächen

$$\mathbf{p}(u_1, u_2) = \begin{pmatrix} \cos(u_1) \left(1 - u_2 \sin\left(\frac{1}{2} m u_1\right)\right) \\ \sin(u_1) \left(1 + u_2 \sin\left(\frac{1}{2} m u_1\right)\right) \\ u_2 \cos\left(\frac{1}{2} m u_1\right) \end{pmatrix}, \quad u_1 \in \mathbb{R}, \quad u_2 \in [-\eta, \eta]$$

mit ganzzahligem Parameter m und reellem Parameter $\eta > 0$. Die entsprechende Fläche besitzt für jeden geraden Wert von m wirklich zwei Seiten, aber für jeden ungeraden Wert von m nur eine Seite (siehe das Bild und seine Beschreibung). Im ersten Fall sprechen Mathematiker von einer *orientierbaren Fläche*, im zweiten nicht.

Achtung: Der Satz von Stokes gilt nur auf Flächen, die zwei Seiten haben, also hier nur für gerade m .

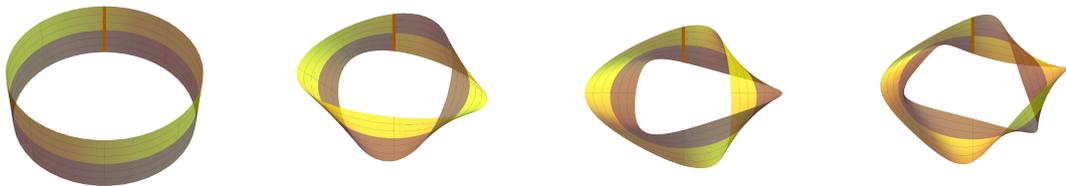


Abbildung Die parametrisierte Fläche von eben für $m = 0$ (Zylinderband), $m = 1$ (Möbiusband) sowie für $m = 2$ und $m = 3$, wobei die Flächen abwechselnd zwei bzw. eine Seite besitzen und daher als orientierbar bzw. nicht orientierbar klassifiziert werden. Sie können diese Flächen sehr einfach zu Hause aus einem schmalen Papierstreifen nachbauen, in dem Sie die kurzen Enden des Streifens direkt verkleben ($m = 0$) oder zunächst um 180 Grad ($m = 1$), 360 Grad ($m = 2$) bzw. 540 Grad gegeneinander drehen und dann verkleben. Für $m = 0$ und $m = 2$ entsteht dadurch eine Fläche mit zwei Seiten und Sie können mit dem Finger entweder die Innenseite oder die Außenseite vollständig abfahren, ohne dabei die Seite zu wechseln. Alternativ können Sie die eine Seite vollständig blau, die andere vollständig grün färben. Für $m = 1$ und $m = 3$ geht das alles nicht, denn es gibt nur eine Seite. Probieren Sie es aus! Ein zweiter wichtiger Unterschied betrifft den Flächenrand: Für $m = 0$ und $m = 2$ wird die Fläche durch *zwei* geschlossene Raumkurven berandet, aber für $m = 1$ und $m = 3$ besteht der Flächenrand aus *einer* einzelnen geschlossenen Raumkurve. Auch das kann man sehen bzw. sich mit verklebten Papierstreifen klar machen. Im Bild ist die Klebekante rot markiert und zur besseren Illustration haben wir angenommen, dass der Streifen teils aus braunem und teils aus gelbem Papier besteht.

Theorem (Integralsatz von Stokes in 3D) Mit den obigen Notationen und Annahmen gilt

$$\int_F \langle \mathbf{rot} \mathbf{f}(\mathbf{x}), \boldsymbol{\nu}(\mathbf{x}) \rangle d\sigma = \int_{\partial F} \langle \mathbf{f}(\mathbf{x}), \boldsymbol{\tau}(\mathbf{x}) \rangle ds,$$

wobei $\langle \cdot, \cdot \rangle$ immer das Skalarprodukt im \mathbb{R}^3 meint und $\partial F = C$ den Flächenrand von F meint.

Beweisidee Die 3D-Variante kann mithilfe der Parametrisierung \mathbf{p} aus der bereits bewiesenen 2D-Variante abgeleitet werden, wobei F gerade das Bild der Menge D ist

und $\mathbf{c} = \mathbf{p} \circ \mathbf{a}$ als Liftung einer planaren Kurve \mathbf{a} betrachtet werden kann, die ∂D parametrisiert (siehe Bild). Außerdem wird durch $\mathbf{g} = \mathbf{f} \circ \mathbf{p}$ ein Vektorfeld auf D definiert. Man schreibt nun die 2D-Variante des Satzes von Stokes für das Gebiet D mit Randkurve \mathbf{a} sowie das Vektorfeld \mathbf{g} hin und transformiert abschließend sowohl das Stokessche 2D-Gebietsintegrals als auch das Stokes 2D-Kurvenintegral mittels \mathbf{p} . \square

Interpretation Der dreidimensionale Satz von Stokes ist das direkte Analogon zur 2D-Variante, nur dass nun die Fläche gekrümmt sein darf und durch eine Raumkurve berandet wird.

Bemerkungen

1. Der Satz von Stokes erlaubt es, die Zirkulation eines Vektorfeldes entlang einer geschlossenen Kurve alternativ als Flächenintegral seiner Rotation auszudrücken und umgekehrt. Darüber hinaus ist er von fundamentalem theoretischen Interesse, zum Beispiel in der Hydrodynamik oder in der Maxwell'schen Theorie des Elektromagnetismus.
2. Die Integralsätze von Stokes in 2D und 3D sind eigentlich nur Spezialfälle eines sehr viel allgemeineren Satzes von Stokes, der in allen Raumdimensionen gilt und auch alle Varianten des Integralsatzes von Gauß als Spezialfälle enthält. Dieses allgemeine Theorem kann als

$$\int_{\Omega} d\omega = \int_{\partial\Omega} \omega$$

geschrieben werden und etabliert die Dualität zwischen dem *Randoperator* ∂ und dem *Cartanschen Differentialoperator* d , wobei ω eine *Differentialform* auf der Mannigfaltigkeit Ω ist. Wir können dies in dieser Vorlesung nicht mal ansatzweise vertiefen, obwohl die allgemeine Stokes-Formel und die zu Grunde liegenden Konzepte auch in der theoretischen Elektrodynamik eine wichtige Rolle spielen. Zum Beispiel kann man auch die Kirchhoffschen Regeln für elektrische Netzwerke als Konkretisierung des allgemeinen Satzes von Stokes interpretieren.

3. Die 2D-Variante von Stokes kann als Spezialfall der 3D Version betrachtet werden. Das sieht man zum Beispiel mit der Wahl

$$\mathbf{p}(\mathbf{u}) = \begin{pmatrix} u_1 \\ u_2 \\ 0 \end{pmatrix}, \quad \mathbf{f}(\mathbf{x}) = \begin{pmatrix} f_1(x_1, x_2) \\ f_2(x_1, x_2) \\ 0 \end{pmatrix},$$

denn dann gilt $x_1 = u_1$, $x_2 = u_2$, $x_3 = 0$ sowie $F = D$ und

$$\boldsymbol{\nu}(\mathbf{x}) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{rot} \mathbf{f}(\mathbf{x}) = \begin{pmatrix} 0 \\ 0 \\ \partial_{x_1} f_2(x_1, x_2) - \partial_{x_2} f_1(x_1, x_2) \end{pmatrix}.$$

Beispiele

1. Ein sehr einfaches Beispiel ist die obere Hemisphäre der Einheitskugel

$$F = \{(x_1, x_2, x_3) : x_1^2 + x_2^2 + x_3^2 = 1, x_3 \geq 0\}.$$

Der entsprechende Flächenrand ist gerade die Kreislinie

$$C = \{(x_1, x_2, x_3) : x_1^2 + x_2^2 = 1, x_3 = 0\}$$

und durch

$$\boldsymbol{\nu}(x_1, x_2, x_3) = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad \boldsymbol{\tau}(x_1, x_2, x_3) = \begin{pmatrix} -x_2 \\ +x_1 \\ 0 \end{pmatrix}$$

können wir in zulässiger Weise einen normierten Normalenvektor $\boldsymbol{\nu}$ auf F sowie einen normierten Tangentialvektor $\boldsymbol{\tau}$ auf C einführen. Betrachten wir außerdem das Vektorfeld $\mathbf{f} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ mit

$$\mathbf{f}(\mathbf{x}) = \begin{pmatrix} -x_2 \\ +x_1 \\ x_3 \end{pmatrix},$$

so ergibt sich

$$\langle \mathbf{f}(\mathbf{x}), \boldsymbol{\tau}(\mathbf{x}) \rangle = x_1^2 + x_2^2 = 1 \quad \text{für } \mathbf{x} \in C$$

und damit auch

$$\int_C \langle \mathbf{f}(\mathbf{x}), \boldsymbol{\tau}(\mathbf{x}) \rangle ds = \int_C 1 ds = \text{len}(C) = 2\pi,$$

wobei $\text{len}(C)$ die Länge von C ist. Der Satz von Stokes impliziert nun

$$\int_F \langle \text{rot } \mathbf{f}(\mathbf{x}), \boldsymbol{\nu}(\mathbf{x}) \rangle d\sigma = 2\pi,$$

aber wir wollen dies noch einmal nachprüfen. Dazu berechnen wir

$$\text{rot } \mathbf{f}(\mathbf{x}) = \begin{pmatrix} \partial_{x_2} f_3(\mathbf{x}) - \partial_{x_3} f_2(\mathbf{x}) \\ \partial_{x_3} f_1(\mathbf{x}) - \partial_{x_1} f_3(\mathbf{x}) \\ \partial_{x_1} f_2(\mathbf{x}) - \partial_{x_2} f_1(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}$$

und erhalten

$$\langle \text{rot } \mathbf{f}(\mathbf{x}), \boldsymbol{\nu}(\mathbf{x}) \rangle = 2x_3 \quad \text{für } \mathbf{x} \in F$$

als die zu integrierende Funktion auf F . Parametrisieren wir F zum Beispiel durch Euler-Winkel $u_1 \in [0, \pi]$, $u_2 \in [0, \pi/2]$, so gilt $d\sigma = \cos(u_2) du_1 du_2$ (siehe oben) und wir erhalten mit

$$\begin{aligned} \int_F \langle \text{rot } \mathbf{f}(\mathbf{x}), \boldsymbol{\nu}(\mathbf{x}) \rangle d\sigma &= \int_0^{2\pi} \int_0^{\pi/2} 2 \sin(u_2) \cos(u_2) du_2 du_1 \\ &= 2\pi \int_0^{\pi/2} \sin(2u_2) du_2 = 2\pi, \end{aligned}$$

das gewünschte Ergebnis.

2. Wir betrachten das Vektorfeld

$$\mathbf{f}(\mathbf{x}) = \begin{pmatrix} -x_2^3 \\ +x_1^3 \\ -x_3^3 \end{pmatrix}, \quad \mathbf{x} \in G$$

auf dem unendlich ausgedehnten Kreiszyylinder

$$G = \{(x_1, x_2, x_3) : x_1^2 + x_2^2 \leq 1\},$$

wobei \mathbf{f} eine turbulente Rohrströmung beschreibt. Wir wollen nun die Zirkulation von \mathbf{f} entlang der schief im Raum liegenden Ellipsenlinie

$$C = \{(x_1, x_2, x_3) : x_1^2 + x_2^2 = 1, x_3 = 1 - x_1 - x_2\},$$

berechnen, wobei K so durchlaufen werden soll, dass der Einheitskreis der (x_1, x_2) -Ebene (Projektion von C) entgegen dem Uhrzeigersinn durchlaufen wird.

Lösung durch Kurvenintegral: Wir können die Kreislinie K durch

$$\mathbf{c} = \begin{pmatrix} \cos(t) \\ \sin(t) \\ 1 - \cos(t) - \sin(t) \end{pmatrix}, \quad t \in [0, 2\pi]$$

parametrisieren und das Stokesche Kurvenintegral mittels

$$\int_K \langle \mathbf{f}(\mathbf{x}), \boldsymbol{\tau}(\mathbf{x}) \rangle ds = \int_{\mathbf{x}} \mathbf{f}(\mathbf{x}) \cdot d\mathbf{x} = \int_0^{2\pi} \langle \mathbf{f}(\mathbf{c}(t)), \dot{\mathbf{c}}(t) \rangle dt$$

berechnen. Dabei haben wir ausgenutzt, dass $\boldsymbol{\tau}(\mathbf{c}(t))$ immer in die gleiche (und nicht die entgegengesetzte) Richtung $\dot{\mathbf{c}}(t)$ zeigt, da wir andernfalls ein Vorzeichen ändern müssen. Mit direkten Rechnungen erhalten wir

$$\dot{\mathbf{c}}(t) = \begin{pmatrix} -\sin(t) \\ +\cos(t) \\ +\sin(t) - \cos(t) \end{pmatrix}, \quad \boldsymbol{\tau}(\mathbf{x}) = + \frac{\dot{\mathbf{c}}(t)}{|\dot{\mathbf{c}}(t)|}$$

sowie

$$\langle \mathbf{f}(\mathbf{c}(t)), \dot{\mathbf{c}}(t) \rangle = \sin^4(t) + \cos^4(t) + (\cos(t) - \sin(t))(1 - \cos(t) - \sin(t))^3,$$

und schließlich ergibt sich

$$\int_K \langle \mathbf{f}(\mathbf{x}), \boldsymbol{\tau}(\mathbf{x}) \rangle ds = \frac{2}{3} \pi$$

nach Einsetzen und Berechnung der Integrale.

Lösung durch Gebietsintegral: Wir können die Menge C als den Flächenrand der Schnittfläche $F = G \cap E$ zwischen dem Zylinder G und der Ebene

$$E = \{(x_1, x_2, x_3) : x_1 + x_2 + x_3 = 1\}$$

betrachten, wobei wir F zum Beispiel als Graph einer Funktion mittels

$$\mathbf{p}(\mathbf{u}) = \begin{pmatrix} u_1 \\ u_2 \\ 1 - u_1 - u_2 \end{pmatrix}, \quad \mathbf{u} \in D := \{(u_1, u_2) : u_1^2 + u_2^2 \leq 1\}$$

parametrisieren können. Der Satz von Stokes erlaubt es, die gesuchte Zirkulation durch ein Flächenintegral zu ermitteln. Dazu berechnen wir

$$\mathbf{rot} \mathbf{f}(\mathbf{x}) = \begin{pmatrix} \partial_{x_2} f_3(\mathbf{x}) - \partial_{x_3} f_2(\mathbf{x}) \\ \partial_{x_3} f_1(\mathbf{x}) - \partial_{x_1} f_3(\mathbf{x}) \\ \partial_{x_1} f_2(\mathbf{x}) - \partial_{x_2} f_1(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 3x_1^2 + 3x_2^2 \end{pmatrix}$$

sowie

$$\partial_{u_1} \mathbf{p}(\mathbf{u}) \times \partial_{u_2} \mathbf{p}(\mathbf{u}) = \begin{pmatrix} +1 \\ 0 \\ -1 \end{pmatrix} \times \begin{pmatrix} 0 \\ +1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

wobei die zweite Formel

$$d\sigma = \sqrt{3} du_1 du_2$$

liefert sowie

$$\boldsymbol{\nu}(\mathbf{p}(\mathbf{u})) = \pm \frac{1}{3} \begin{pmatrix} \sqrt{3} \\ \sqrt{3} \\ \sqrt{3} \end{pmatrix}.$$

Wir müssen nun aber noch das richtige Vorzeichen wählen, damit $\boldsymbol{\nu}$ und $\boldsymbol{\tau}$ zueinander passen. Im konkreten Fall ist $+$ die richtige Wahl. Man kann dies mit Hilfe der geometrischen Anschauung oder einem 3D-Plot begründen (siehe Bild). Alternativ können wir einzelne Punkte in C betrachten. Mit $\mathbf{x}_* = \mathbf{p}(\frac{1}{4}\pi)$ gilt etwa

$$\mathbf{x}_* = \frac{1}{2} \begin{pmatrix} \sqrt{2} \\ \sqrt{2} \\ 2 - 2\sqrt{2} \end{pmatrix}, \quad \boldsymbol{\tau}(\mathbf{x}_*) = \frac{1}{2} \begin{pmatrix} -\sqrt{2} \\ +\sqrt{2} \\ 0 \end{pmatrix}, \quad \boldsymbol{\nu}(\mathbf{x}_*) \times \boldsymbol{\tau}(\mathbf{x}_*) = \frac{1}{6} \begin{pmatrix} -\sqrt{6} \\ -\sqrt{6} \\ 2\sqrt{6} \end{pmatrix}$$

und man sieht die positive Orientierung (zum Beispiel, weil $\boldsymbol{\nu}(\mathbf{x}_*) \times \boldsymbol{\tau}(\mathbf{x}_*)$ im Punkts \mathbf{x}_* ins Innere von F zeigt). Da wir nun $\boldsymbol{\nu}$ identifiziert haben, berechnen wir

$$\mathbf{rot} \mathbf{f}(\mathbf{p}(\mathbf{u})) = \begin{pmatrix} 0 \\ 0 \\ 3(u_1^2 + u_2^2) \end{pmatrix}, \quad \langle \mathbf{rot} \mathbf{f}(\mathbf{p}(\mathbf{u})), \boldsymbol{\nu}(\mathbf{p}(\mathbf{u})) \rangle = \sqrt{3} (u_1^2 + u_2^2)$$

und erhalten

$$\begin{aligned} \int_F \langle \mathbf{f}(\mathbf{x}), \boldsymbol{\nu}(\mathbf{x}) \rangle d\sigma &= \int_F \mathbf{f}(\mathbf{x}) \cdot d\boldsymbol{\sigma} = \int_D \sqrt{3} (u_1^2 + u_2^2) \sqrt{3} du \\ &= 3 \int_0^1 r^3 \int_0^{2\pi} d\varphi dr = \frac{3}{2} \pi, \end{aligned}$$

wobei wir zur Berechnung des zweidimensionalen Gebietsintegrals über D die ebenen Polarkoordinaten in der (u_1, u_2) -Ebene benutzt haben.

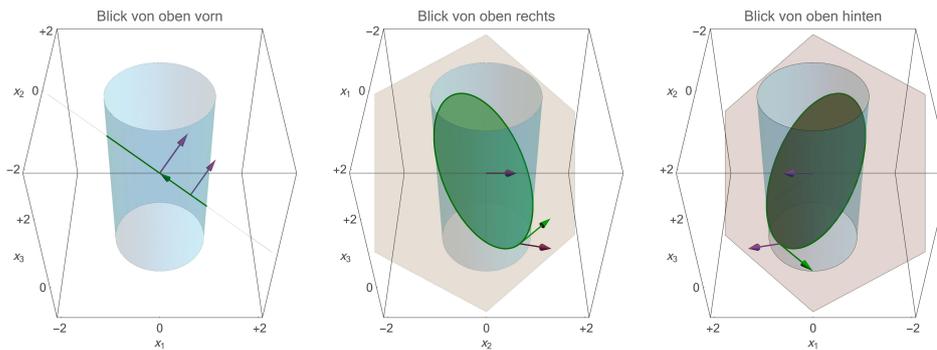


Abbildung Die Bilder zum eben gerechneten zweiten Beispiel. F ist die schiefliegende elliptische Scheibe (grün).

Ausblick: Elektrische und magnetische Felder* Die *Maxwellschen Gleichungen* modellieren die Gesamtheit aller elektromagnetischen Erscheinungen und können als

$$\begin{aligned} \operatorname{div} \mathbf{D}(t, \mathbf{x}) &= 4\pi \varrho(t, \mathbf{x}) \\ \operatorname{div} \mathbf{B}(t, \mathbf{x}) &= 0 \\ + c^{-1} \partial_t \mathbf{B}(t, \mathbf{x}) + \operatorname{rot} \mathbf{E}(t, \mathbf{x}) &= 0 \\ - c^{-1} \partial_t \mathbf{D}(t, \mathbf{x}) + \operatorname{rot} \mathbf{H}(t, \mathbf{x}) &= 4\pi c^{-1} \mathbf{j}(t, \mathbf{x}) \end{aligned}$$

geschrieben werden. Hierbei ist t die Zeit und $\mathbf{x} \in \mathbb{R}^3$ die Raumvariable, wobei sich die Differentialoperatoren div und rot nur auf diese \mathbf{x} -Koordinaten beziehen. Außerdem ist c die Lichtgeschwindigkeit und es gelten die punktweisen Identitäten

$$\mathbf{D}(t, \mathbf{x}) = \varepsilon \mathbf{E}(t, \mathbf{x}), \quad \mathbf{B}(t, \mathbf{x}) = \mu \mathbf{H}(t, \mathbf{x}),$$

wobei die physikalischen Konstanten ε bzw. μ die Dielektrizitätskonstante bzw. die magnetische Permeabilität genannt werden. Im Standardfall sind die Ladungsdichte ϱ sowie der Leitungsstrom \mathbf{j} bekannt und man möchte mit Hilfe der obigen Gesetze die elektrische Feldstärke \mathbf{E} , die magnetische Feldstärke \mathbf{H} , die dielektrische Verschiebung \mathbf{D} sowie die magnetische Flussdichte \mathbf{B} bestimmen, wobei noch geeignete Anfangs- und Randbedingungen gestellt werden können.

Bei den Maxwellschen Gleichungen handelt es sich aus mathematischer Sicht um *partielle Differentialgleichungen*, deren Theorie und Praxis wir in dieser Vorlesung nicht behandeln können. Wir wollen aber festhalten, dass sie wegen der Sätze von Gauß und Stokes äquivalent zu den Integralformeln

$$\begin{aligned} \int_{\partial G} \langle \mathbf{D}(t, \mathbf{x}), \boldsymbol{\nu}(\mathbf{x}) \rangle d\sigma &= 4\pi \int_G \varrho(t, \mathbf{x}) d\mathbf{x}, \\ \int_{\partial G} \langle \mathbf{B}(t, \mathbf{x}), \boldsymbol{\nu}(\mathbf{x}) \rangle d\sigma &= 0 \end{aligned}$$

sowie

$$\begin{aligned} \int_{\partial F} \langle \mathbf{E}(t, \mathbf{x}), \boldsymbol{\tau}(\mathbf{x}) \rangle ds &= -c^{-1} \int_F \langle \partial_t \mathbf{B}(t, \mathbf{x}), \boldsymbol{\nu}(\mathbf{x}) \rangle d\sigma, \\ \int_{\partial F} \langle \mathbf{H}(t, \mathbf{x}), \boldsymbol{\tau}(\mathbf{x}) \rangle ds &= +c^{-1} \int_F \langle \partial_t \mathbf{D}(t, \mathbf{x}), \boldsymbol{\nu}(\mathbf{x}) \rangle d\sigma + 4\pi c^{-1} \int_F \langle \mathbf{j}(t, \mathbf{x}), \boldsymbol{\nu}(\mathbf{x}) \rangle d\sigma \end{aligned}$$

sind, wobei G ein beliebiges *Testvolumen* mit Normalenvektor ν und F eine *Testfläche* in ∂G mit Tangentialvektor τ bezeichnet. Die integrale Form der Maxwell'schen Gleichungen ist aus physikalischer Sicht eigentlich fundamentaler und kann benutzt werden, um die differentielle Variante herzuleiten.