Analysis III

Wintersemester 2006/07
Wintersemester 2010/11

Skript zur Analysis III
Wintersemester 2006/07
Wintersemester 2010/11
Prof. Dr. Daniel Grieser

Carl-von-Ossietzky-Universität Oldenburg
Institut für Mathematik
D - 26111 Oldenburg i. O.
daniel.grieser@uni-oldenburg.de

Die Homepage zur Vorlesung:
http://www.staff.uni-oldenburg.de/daniel.grieser/wwwlehre/06WS.analysis_3/
http://www.staff.uni-oldenburg.de/daniel.grieser/wwwlehre/10WS.analysis_3/

Bearbeitung: Uwe Batterham, Christina Delfs, Stefan Grahl, Stefan Hellbusch, Andreas Hettler, Roman Rathje, Jörg Sauter, Falko Schulz, Heike de Vries und Hero Wanders

Veröffentlicht zu den Creative-Commons-Bedingungen
http://creativecommons.org/licenses/by-nc-sa/2.5/deed.de

Zuletzt bearbeitet: **1. April 2011**

Vorwort zur Überarbeitung im Wintersemester 2010/11

Während der Vorlesung Analysis III im Wintersemester 2010/2011 habe ich folgende Kapitel des Skripts überarbeitet: Untermannigfaltigkeiten, Hauptsätze der Differential- und Integralrechnung, Laplace-Gleichung, Faltung, Fouriertransformation.

Die Abschnitte 5.3, 5.4 und 5.5 über Fourierreihen sind noch in der Rohfassung, wie sie von der Vorlesungsmitschrift abgetippt wurden. Hier ist noch einiges unvollständig und nicht ganz korrekt. Ich hoffe, diese Abschnitte bei der nächsten Analysis 3-Vorlesung zu überarbeiten!

Im Wintersemester 2010/2011 habe ich die Theorie des Lebesgue-Integrals grundlegend anders eingeführt als im Wintersemester 2006/2007: Damals hatte ich den Zugang gewählt, erst das Integral zu definieren und daraus abgeleitet das Maß von Teilmengen des $\mathbb{R}^n$. Dieses Mal habe ich zuerst den allgemeinen Begriff des Maßes eingeführt und dann das Integral für beliebige Maßräume definiert. Das Lebesgue-Maß und Lebesgue-Integral auf dem $\mathbb{R}^n$ ist dann ein Spezialfall. Der Vorteil dieses Zugangs liegt darin, dass der Begriff des Maßes, eines der ganz fundamentalen Konzepte der Mathematik, in den Vordergrund gestellt wird. Die Konstruktion des Integrals aus einem Maß ist dann sehr natürlich. Im Unterschied dazu ist die direkte Konstruktion des Lebesgue-Integrals – mit Hilfe der Idee der Vervollständigung des Integrals auf Treppenfunktionen bezüglich einer geeigneten Norm – zwar sehr effizient, wirkt aber für Studierende des dritten Semesters eher unmotiviert. In Prüfungen habe ich oft erlebt, dass die Studierenden dann zwar die Definition der L^1-Norm mit Hilfe von Treppenfunktionen, die hier benötigt wird, richtig wiedergeben konnten, aber wenig Verständnis für die grundlegenden Ideen hatten.

Ein weiterer Grund, den Zugang über die Maßtheorie zu wählen, liegt darin, dass dies auf die Wahrscheinlichkeitstheorie vorbereitet, wo der Maßbegriff auch für andere Räume als $\mathbb{R}^n$ benötigt wird; weiterhin wird in der Spektraltheorie selbstadjungierter Operatoren in der Funktionalanalysis der Maßbegriff benötigt. Schließlich wird auf diese Weise die Analogie von Summieren und Integrieren auf eine solide Basis gestellt.

Für den maßtheoretischen Zugang zur Integrationstheorie habe ich im Wintersemester 2010/2011 ein Kurzskript herausgegeben, das zum großen Teil von der Tutorin Katrin Tönjes aus Vorlesungsaufzeichnungen geschrieben wurde. Ich danke ihr hier noch einmal dafür. Dieses Kurzskript kann auf meiner Homepage heruntergeladen werden. Das erste Kapitel des vorliegenden Skripts habe ich nicht verändert, es enthält also weiterhin den kurzen Weg zum Integral, ohne Maßtheorie. Ab dem zweiten Kapitel folgte die Vorlesung 2010 weitgehend diesem Skript, wobei ich in der Vorlesung natürlich die Beweise des Satzes von Fubini und der Transformationsformel entsprechend der geänderten Definition des Integrals angepasst habe.

Oldenburg, den 31.3.2011

Daniel Grieser

Vorwort

Dieses Skript entsteht parallel zur Vorlesung unter Mitarbeit von Hörerinnen und Hörern der Vorlesung. Es wird abschnittsweise herausgegeben. Der Inhalt geht an einigen Stellen über den Vorlesungsstoff leicht hinaus. Keine Bange: In der Klausur wird nur der Vorlesungsstoff erwartet.

Die zusätzlichen Definitionen und Sätze werden mit Nummern größer als 100 versehen. Auf diese Weise kann die Nummerierung der Definitionen, Sätze usw. der Vorlesung beibehalten und trotzdem an späterer Stelle auf die neuen Dinge verwiesen werden.

Für Hinweise auf Druckfehler und Anregungen bin ich immer dankbar. Am besten schreiben Sie mir eine E-Mail (grieser@mathematik.uni-oldenburg.de). Am Semesterende wird das gesamte Skript überarbeitet werden.

Oldenburg, den 20.12.2006
Daniel Grieser

Einleitung

Nachdem wir uns in Analysis I mit der Analysis in einer Dimension und in Analysis II mit der Differentialrechnung in mehreren Dimensionen befasst haben, ist das Thema der Analysis III die Integration in mehreren Dimensionen. Die Hauptthemen dieser Analysis-III-Vorlesung sind:

- Wie berechnet man das Volumen von Mengen $A \subset \mathbb{R}^n$?
  Hier muss man sich zunächst fragen, was »Volumen« überhaupt bedeutet. Wir werden sogar sehen, dass man für beliebige Mengen gar nicht von einem Volumen in konsistenter Weise sprechen kann! Daher ist Sorgfalt geboten.

- Wie integriert man Funktionen $f\colon \mathbb{R}^n \to \mathbb{R}$?
  Auch hier müssen wir uns die Frage stellen, welche Funktionen überhaupt integrierbar sind. In Analysis I haben wir gesehen, wie man z.B. stetige Funktionen (allgemeiner Regelfunktionen) integriert. Diese sind für viele Zwecke zu speziell. Wir werden viel allgemeinere Funktionen, sogenannte Lebesgue-integrierbare Funktionen, integrieren können.

- Ändert man den Wert einer Funktion $f : [a,b] \to \mathbb{R}$ an einem Punkt, so ändert sich $\int_a^b f$ nicht. Ebenso, wenn man den Wert an endlich vielen Punkt ändert. Wie groß darf eine »Änderungsmenge« höchstens sein, damit das noch stimmt? Dies führt zum Begriff der *Nullmenge*.

- Wie berechnet man Längen von Kurven oder den Flächeninhalt einer Oberfläche?
  Hierzu müssen zunächst »Oberflächen« eingeführt werden. Dies führt zum Begriff der *Untermannigfaltigkeit des* $\mathbb{R}^n$. Allgemeiner werden wir untersuchen, wie man eine Funktion über eine Untermannigfaltigkeit integriert.

- Gibt es höherdimensionale Verallgemeinerungen des Hauptsatzes der Differential- und Integralrechnung?
  Dieser Satz besagt, dass
  $$\int_a^b f(x)\,dx = F(b) - F(a)$$
  für $f = F'$ gilt. Mit anderen Worten: Ist f die Ableitung von F, so kann das Integral von f über ein Intervall $[a,b]$ mittels der Werte von F am Rand des Intervalls ausgedrückt werden.

Ist nun $\Omega \subset \mathbb{R}^n$ eine offene Menge und bezeichnen wir ihren Rand mit $\partial\Omega$, so können wir fragen, ob eine Relation der Art

$$\int_\Omega f = \int_{\partial\Omega} F$$

gilt, falls f, F in geeigneter Beziehung zueinander stehen. In der Tat stimmt dies, es gibt sogar verschiedene Varianten hiervon, und diese sogenannten Integralsätze (von Stokes, Green, Gauss) sind für die höherdimensionale Analysis zentral.

Bei der Suche nach diesen sogenannten Integralsätzen werden wir natürlich die Notwendigkeit erkennen, nicht nur Funktionen, sondern auch andere Objekte (Vektorfelder, Differentialformen) zu integrieren.

Wie schon in Analysis II wird die Lineare Algebra durchgehend eine wichtige Rolle spielen. Ein einfaches Beispiel: Die Fläche eines Parallelogramms, das von zwei Vektoren a, b in der Ebene aufgespannt wird, ist gleich $|\det(a, b)|$. Dies ist die Grundlage für die Transformationsformel, die mehrdimensionale Verallgemeinerung der Substitutionsregel, die für die Berechnung vieler mehrdimensionaler Integrale unentbehrlich ist.

Inhaltsverzeichnis Analysis III

1 Definition des Lebesgue-Integrals **1**

Vorbemerkungen zum Volumen-/Integralbegriff . . . 1

Das Grundprinzip der Integration . . . 2

1.1 Treppenfunktionen . . . 3

1.2 Die L^1-Halbnorm . . . 6

1.3 Definition und einfache Eigenschaften des Lebesgue-Integrals . . . 12

1.4 Nullmengen, Nullfunktionen und »fast überall« . . . 15

1.5 Konvergenzsätze . . . 19

1.6 Das Maß von Mengen . . . 24

2 Berechnung von Integralen **31**

2.1 Der Satz von Fubini . . . 31

2.2 Die Transformationsformel . . . 35

2.3 Beweis der Transformationsformel . . . 37

2.4 Weitere Beispiele zur Integralberechnung . . . 46

3 Untermannigfaltigkeiten des $\mathbb{R}^N$ und deren Volumen **49**

3.1 Untermannigfaltigkeiten des $\mathbb{R}^N$ . . . 49

3.2 Der Tangentialraum einer Mannigfaltigkeit . . . 56

3.3 Das Volumen von Parallelotopen . . . 58

3.4 Volumen von und Integration über Untermannigfaltigkeiten des $\mathbb{R}^N$ . . . 60

4 Hauptsätze der Differential- und Integralrechnung in mehreren Dimensionen **69**

4.1 Kurvenintegrale und Gradientenfelder . . . 69

4.2 Mannigfaltigkeiten mit Rand . . . 80

4.3 Der Satz von Gauß . . . 83

5 Laplace-Gleichung, Faltung und Fourieranalysis **99**

5.1 Die Laplace-Gleichung . . . 99

5.2 Die Faltung im $\mathbb{R}^n$ und Anwendungen . . . 104

5.3 Grundbegriffe über Fourierreihen . . . 112

5.4 Konvergenz von Fourierreihen; der Satz von Fejér . . . 114

5.5 Fourierreihen und Orthogonalität . . . 118

5.6 Fouriertransformation . . . 123

1 Definition des Lebesgue-Integrals

Vorbemerkungen zum Volumen-/Integralbegriff

Bereits aus Analysis I sind wir damit vertraut, dass Flächenberechnung sehr eng mit Integration verwandt ist: Für $f\colon [a,b] \to \mathbb{R}$ ist

$$\int_a^b f(x)\,dx = \text{Fläche (mit Vorzeichen) unter dem Graphen von } f$$

(Man sollte natürlich geeignete Annahmen machen, damit das Integral überhaupt definiert ist, z. B. dass f eine Regelfunktion, etwa stückweise stetig, ist; auf Präzision soll es in dieser Einleitung aber nicht ankommen.)

Es wird nicht überraschen, dass etwas ähnliches auch für die Berechnung von Volumina geht: Ist $f : [a,b] \times [c,d] \to \mathbb{R}$, so ist der Graph von f eine Fläche im Raum, und dann ist

$$\int_c^d \int_a^b f(x,y)\,dx\,dy = \text{Volumen unter dem Graph von } f$$

(hierbei integriert man zunächst bzgl. x für jedes feste y, dann das Ergebnis bzgl. y).

Mengen »unter Graphen« nennt man **Ordinatenmengen**. Diese sind recht speziell, haben z. B. immer mindestens eine gerade Kante bzw. Seitenfläche.

Wie kann man nun das Volumen allgemeinerer Mengen berechnen? Zum Beispiel durch Zerlegung in Ordinatenmengen (zerlege etwa die Kreisscheibe in eine obere und eine untere Hälfte).

Anhand einfacher Beispiele überzeugt man sich leicht, dass eine solche Zerlegung sehr kompliziert werden kann.
Noch etwas schwieriger wird's bei Körpern: Selbst so einfache Körper wie ein Zylinder

$$Z = \{(x,y,z) :\ x^2 + y^2 \le 1,\ z \in [0,1]\}$$

bereiten eine Schwierigkeit: Zwar ist Z die Ordinatenmenge der Funktion

$$f : K \to \mathbb{R}\,, \quad (x,y) \mapsto 1\,, \quad K = \{(x,y) :\ x^2 + y^2 \le 1\}$$

doch der Definitionsbereich dieser Funktion ist kein Rechteck, sondern eine Kreisscheibe, das Rezept oben funktioniert also nicht. Ein einfacher Ausweg ist die Beobachtung, dass für jedes feste y die Menge der »erlaubten« Werte von x das Intervall $[-\sqrt{1-y^2}, \sqrt{1-y^2}]$ ist, daher sollte wohl

$$\text{Volumen von } Z = \int_{-1}^{1} \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} 1\ dx\,dy$$

sein. Wir werden hierfür abkürzend $\int_K f$ schreiben (das Integral von f über die Menge K).

Anstatt nun Integrale über verschiedenste Mengen $K \subset \mathbb{R}^n$ zu betrachten und dabei untersuchen zu müssen, für welche Arten von Mengen das überhaupt sinnvoll ist und für welche nicht, ist es konzeptuell einfacher, von Anfang an nur Integrale über den ganzen $\mathbb{R}^n$ zu betrachten. Ist dann $f : K \to \mathbb{R}$ mit $K \subset \mathbb{R}^n$, so kann man dies leicht auf den $\mathbb{R}^n$-Fall zurückführen: Setze

$$\tilde{f}(x) = \begin{cases} f(x) & \text{für } x \in K \\ 0 & \text{sonst,} \end{cases}$$

dann sollte offenbar $\int_K f = \int_{\mathbb{R}^n} \tilde{f}$ sein.

Das kleine Problem hierbei ist, dass $\tilde{f}$ im Allgemeinen unstetig sein wird, selbst wenn f stetig ist. Denn am Rand von K springt $\tilde{f}$ auf Null.

Also: Haben wir einen Integralbegriff für (möglicherweise unstetige) Funktionen f auf $\mathbb{R}^{n-1}$, so können wir damit das Volumen von Ordinatenmengen im $\mathbb{R}^n$ erklären.

Es geht noch besser: Ein Integralbegriff für Funktionen im $\mathbb{R}^n$ (statt $\mathbb{R}^{n-1}$!) liefert einen Volumenbegriff für (ziemlich) beliebige Mengen $A \subset \mathbb{R}^n$:

$$\text{Volumen von } A = \int_{\mathbb{R}^n} \chi_A$$

Hierbei verwenden wir folgende Schreibweise:

1.0.1 Definition

Sei $A \subset \mathbb{R}^n$. Die **charakteristische Funktion** von A ist

$$\chi_A(x) = \begin{cases} 1 & \text{für } x \in A \\ 0 & \text{für } x \notin A \end{cases}$$

(χ = chi).

Zusammenfassend: Wir gehen so vor, dass wir

- ▷ erst Integrale definieren und untersuchen;
- ▷ dann das Volumen (Maß) von Mengen.

Man kann auch umgekehrt vorgehen: Erst das Maß von Mengen definieren und untersuchen und dann daraus das Integral für Funktionen erklären. Das ist jedoch aufwändiger, und da wir ja in dieser Vorlesung auch noch manches andere als Integrationstheorie kennenlernen möchten, habe ich diesen Zugang gewählt.

Das Grundprinzip der Integration

Wie kann man einen Begriff des Integrals einer Funktion definieren, der der Vorstellung eines »Flächeninhalts unter einem Graphen« entspricht und mathematisch exakt ist?
Erinnerung: Integraldefinition Analysis I. Sei $f\colon [a,b] \to \mathbb{R}$. Zwei Schritte:

1. Falls f Treppenfunktion ist, dann ist $\int_a^b f(x)\,dx$ als Summe von Rechteckflächen definiert.
2. Falls f sich gleichmäßig durch Treppenfunktion approximieren lässt, $f = \operatorname{glm}\lim_{n\to\infty} t_n$, definiere
$$\int_a^b f(x)\,dx = \lim_{n\to\infty} \int_a^b t_n(x)\,dx \qquad (\text{»}\textbf{Regelintegral}\text{«}).$$

Funktionen, die sich gleichmäßig durch Treppenfunktionen approximieren lassen, heißen Regelfunktionen. Für fast alle praktischen Zwecke reichen diese aus, jedoch sind sie für eine »runde« mathematische Theorie ungeeignet, siehe unten.

Wir wollen dies in zwei Richtungen verbessern. Erstens wollen wir mehr Funktionen integrieren können, und zweitens wollen wir in mehreren Dimensionen integrieren. Also brauchen wir:

1. Was sind Treppenfunktionen im $\mathbb{R}^n$?
2. Approximiere allgemeine Funktionen f durch Treppenfunktionen, mit einem schwächerem Konvergenzbegriff als gleichmäßiger Konvergenz.

Was geht bei Regelfunktionen schief? Sie verhalten sich schlecht unter Grenzwertbildung. Wir hätten gerne folgendes (und werden es später für Lebesgue-integrierbare Funktionen beweisen):
Falls $f_k, f\colon [a,b] \to \mathbb{R}$ $(k \in \mathbb{N})$ und

- $f_k \xrightarrow{k\to\infty} f$ punktweise
- es existiert $C \in \mathbb{R}$ mit $\forall k\, \forall x\colon |f_k(x)| \le C$.

Dann gilt: $\displaystyle\int_a^b f_k(x)\,dx \xrightarrow{k\to\infty} \int_a^b f(x)\,dx$.
(Genauer: Sind alle f_k Lebesgue-integrierbar, so auch f, und diese Grenzwertbeziehung gilt.)
Für Regelfunktionen stimmt dies nicht!

Diese Eigenschaft ist zum Beispiel im Zusammenhang mit der Volumenberechnung nützlich: Wird etwa eine Menge A von innen durch Mengen $A_1, A_2, \ldots$ approximiert (d. h. gilt $A_1 \subset A_2 \subset \ldots$ und $A = \bigcup_k A_k$), so konvergiert die Funktionenfolge $(\chi_{A_k})_k$ punktweise gegen χ_A, aber nicht gleichmäßig.

1.1 Treppenfunktionen

1.1.1 Definition

Ein **Quader** im $\mathbb{R}^n$ ist eine Menge der Form

$$Q = I_1 \times I_2 \times \cdots \times I_n$$

wobei $I_1, \ldots, I_n \subset \mathbb{R}$ nichtleere, beschränkte Intervalle sind.

Beispiele:

- $n = 1$: Quader = Intervall.
- $n = 2$: Achsenparallele Rechtecke.

Bemerkung: Die Intervalle I_k können offen, abgeschlossen oder halboffen sein. Punkte sind auch erlaubt.

1.1.2 Definition

Das **Volumen** des Quaders $Q = I_1 \times I_2 \times \cdots \times I_n$ ist

$$\mathbf{vol}(Q) := (\text{Länge von } I_1) \cdots (\text{Länge von } I_n)$$

wobei Länge von (a,b) = Länge von $(a,b]$ = Länge von $[a,b)$ = Länge von $[a,b] = b - a$.

Bemerkung: Es ist $\mathrm{vol}(Q) = 0$ genau dann, wenn eines der I_k ein Punkt ist.

Bezeichnung auch $\mathbf{vol}_n(Q)$ statt $\mathrm{vol}(Q)$.

1.1.3 Definition

Eine **Treppenfunktion** auf $\mathbb{R}^n$ ist eine Funktion $\varphi\colon \mathbb{R}^n \to \mathbb{R}$ der Form

$$\varphi = \sum_{k=1}^{K} c_k \cdot \chi_{Q_k}$$

wobei $c_1, \ldots, c_k \in \mathbb{R}$ und $Q_1, \ldots, Q_k$ Quader sind.

Beispiele:

$n = 1$: $\chi_{[0,1)} + 2 \cdot \chi_{(\frac{1}{2},2)}$.

$n = 2$: »Skyline von Manhattan«.

1.1.4 Lemma

Jede Treppenfunktion $\varphi\colon \mathbb{R}^n \to \mathbb{R}$ hat eine Darstellung

$$\varphi = \sum_{k=1}^{K} c_k \cdot \chi_{Q_k}$$

wobei die Q_k paarweise disjunkt sind.

Beweis (Skizze): Sei $\varphi = \sum_{l=1}^{L} b_l \cdot \chi_{P_l}$. Man zeigt zunächst, dass sich $\bigcup_{l=1}^{L} P_l$ derart in paarweise disjunkte Quader Q_k zerlegen lässt, dass jedes P_l Vereinigung einiger dieser Teile ist. Dies ist anschaulich recht klar, ein exakter Beweis ist etwas langwierig (man zerschneidet $\bigcup_{l=1}^{L} P_l$ mittels sämtlicher Hyperebenen $\{x_i = a_i\}$, wobei a_i über die Endpunkte sämtlicher Intervalle variiert, die als i-ter Faktor in einem der P_l vorkommen, für $i = 1, \ldots, n$).

Es gilt nun: Für $A, B \in \mathbb{R}^n$ mit $A \cap B = \emptyset$ ist $\chi_{A \cup B} = \chi_A + \chi_B$. Damit folgt für jedes l:

$\chi_{P_l} = \sum_{k:\, Q_k \subset P_l} \chi_{Q_k}$. Man setze dies nun in $\varphi = \sum_{l=1}^{L} b_l \cdot \chi_{P_l}$ ein und fasse alle Terme mit demselben Q_k zusammen. □

1.1.5 Lemma

(a) Die Menge der Treppenfunktionen auf $\mathbb{R}^n$, $\mathcal{T}(\mathbb{R}^n, \mathbb{R})$, ist ein Unterraum des Vektorraums $\mathcal{F}(\mathbb{R}^n, \mathbb{R}) = \{\text{Alle Funktionen } \mathbb{R}^n \to \mathbb{R}\}$.

(b) Mit φ, ψ sind auch $|\varphi|$, $\max(\varphi, \psi)$ und $\min(\varphi, \psi)$ Treppenfunktionen.

Beweis:

(a) φ, ψ Treppe $\implies$ $\varphi + \psi$ Treppe. $\surd$

φ Treppe, $c \in \mathbb{R}$ $\implies$ $c\varphi$ Treppe. $\surd$

$\varphi(x) = 0$ für alle x ist Treppe. $\surd$

(b) $\left| \sum_{k=1}^{K} c_k \chi_{Q_k} \right| = \sum_{k=1}^{K} |c_k| \chi_{Q_k}$, falls Q_k paarweise disjunkt.

Denn es gilt: $x \in \mathbb{R}^n$ $\implies$ Es gibt höchstens ein $k = k_0$ mit $x \in Q_{k_0}$.

- Wenn es kein k_0 gibt $\implies$ Beide Seiten sind gleich Null.
- Sonst $\sum_{k=1}^{K} c_k \chi_{Q_k}(x) = c_{k_0}$ und $|c_{k_0}| = \sum_{k=1}^{K} |c_k| \chi_{Q_k}(x)$.

Also ist mit φ auch $|\varphi|$ eine Treppenfunktion. Die anderen Aussagen folgen aus

$$\max(\varphi, \psi) = (\varphi + \psi + |\varphi - \psi|)/2, \ \min(\varphi, \psi) = (\varphi + \psi - |\varphi - \psi|)/2.$$

(Warum stimmt das? Zunächst gilt für Zahlen $a, b \in \mathbb{R}$: $\min\{a, b\} = \frac{1}{2}[a + b - |a - b|]$, denn für $a \geq b$ ist die rechte Seite $\frac{1}{2}(a + b - (a - b)) = b$, und für $a \leq b$ ist sie analog gleich a, also in jedem Fall

das Minimum von a und b. Analog zeigt man $\max\{a,b\} = \frac{1}{2}[a+b+|a-b|]$. Damit gelten diese Formeln auch für Funktionen, da man ja nur beliebige Werte x einsetzen muss.) □

Wir kommen nun zur Definition des **Integrals für Treppenfunktionen**, dem ersten Schritt der Grundprinzipien der Integration.

1.1.6 Definition (+ Satz)

Für Treppenfunktionen $\varphi = \sum_{k=1}^{K} c_k \chi_{Q_k}$ auf $\mathbb{R}^n$ definiere

$$\int_{\mathbb{R}^n} \varphi(x)\, dx := \sum_{k=1}^{K} c_k \operatorname{vol}(Q_k)\,.$$

Dies ist wohldefiniert, d. h. falls $\varphi = \sum_{k=1}^{K} c_k \chi_{Q_k} = \sum_{l=1}^{L} c'_l \chi_{Q'_l}$, so $\sum_{k=1}^{K} c_k \operatorname{vol}(Q_k) = \sum_{l=1}^{L} c'_l \operatorname{vol}(Q'_l)$.

Schreibweise:

$$\int_{\mathbb{R}^n} \varphi(x)\, dx = \int \varphi(x)\, dx = \int \varphi\, dx = \int \varphi$$

In dieser Vorlesung gilt:

Alle Integrale sind bestimmte Integrale!

Also ist $\int \varphi$ eine Zahl, keine Funktion.

Beweis (der Wohldefiniertheit): Wähle Q''_m, $m = 1, \dots, M$,

- ▷ paarweise disjunkt,
- ▷ $\bigcup_{m=1}^{M} Q''_m = \bigcup_k Q_k \cup \bigcup_l Q'_l$,
- ▷ jedes Q_k und jedes Q'_l ist Vereinigung von einigen der Q''_m.

Es gilt

$$\chi_{Q_k} = \sum_{m:\, Q''_m \subset Q_k} \chi_{Q''_m}\,.$$

Also

$$\varphi = \sum_k c_k \chi_{Q_k} = \sum_k c_k \sum_{m:\, Q''_m \subset Q_k} \chi_{Q''_m} = \sum_m c''_m \chi_{Q''_m} \quad \text{mit} \quad c''_m = \sum_{k:\, Q''_m \subset Q_k} c_k\,.$$

Selbe Umformung:

$$\sum_k c_k \operatorname{vol}(Q_k) = \sum_m c''_m \operatorname{vol}(Q''_m) \qquad (*)$$

Bemerke: c''_m sind durch φ eindeutig bestimmt, denn für $x \in Q''_{m_0}$ ist

$$\varphi(x) = \sum_m c''_m \underbrace{\chi_{Q''_m}(x)}_{=\begin{cases} 0 & \text{falls } m \neq m_0 \\ 1 & \text{falls } m = m_0 \end{cases}} = c''_{m_0}$$

Wegen $(*)$ ist also $\sum c_k \operatorname{vol}(Q_k)$ durch φ allein bestimmt, also gleich $\sum c'_l \operatorname{vol}(Q'_l)$. □

1.1.7 Satz

Die Abbildung

$$\int : \mathcal{T}(\mathbb{R}^n, \mathbb{R}) \longrightarrow \mathbb{R}\,, \quad \varphi \longmapsto \int \varphi\, dx$$

ist

1. linear,
2. monoton, d. h. $\varphi \leq \psi \Rightarrow \int \varphi\, dx \leq \int \psi\, dx$,
3. es gilt $|\int \varphi\, dx| \leq \int |\varphi|\, dx$.

Beweis:

1. Linearität: Sollte Ihnen direkt aus der Definition klar sein.
2. Monotonie: Klar, wenn φ, ψ mit denselben paarweise disjunkten Quadern ausgedrückt:

$$\begin{aligned} & \varphi = \sum c_k \chi_{Q_k}\,, \quad \psi = \sum d_k \chi_{Q_k}\,, \quad \varphi \leq \psi \ \ (\text{d. h. } \varphi(x) \leq \psi(x) \text{ für alle } x) \\ \Longrightarrow \quad & c_k \leq d_k \\ \Longrightarrow \quad & \int \varphi\, dx \leq \int \psi\, dx \end{aligned}$$

   Nach dem Argument im Beweis der Wohldefiniertheit kann man zwei beliebige Treppenfunktionen mit Hilfe derselben paarweise disjunkten Quader schreiben!

3. $-|\varphi| \leq \varphi \leq |\varphi| \quad \Longrightarrow \quad -\int |\varphi| \leq \int \varphi \leq \int |\varphi|$ d. h. $\left|\int \varphi\right| \leq \int |\varphi|$. □

1.2 Die L^1-Halbnorm

Im Folgenden ist es nützlich, mit ∞ zu rechnen:

1.2.1 Definition (Rechnen mit ∞ in der Integrationstheorie)

$\overline{\mathbb{R}} := \mathbb{R} \cup \{\infty\}$.

- ▷ $\infty + c = c + \infty = \infty$ für alle $c \in \overline{\mathbb{R}}$.
- ▷ $\infty \cdot c = c \cdot \infty = \begin{cases} \infty & c \neq 0 \\ 0 & c = 0\,. \end{cases}$
- ▷ $|\infty| = \infty$.
- ▷ $c < \infty$ für alle $c \in \mathbb{R}$.

(Beachte, dass $\infty \cdot 0$ im Zusammenhang mit Grenzwerten nicht definiert war. Im Rahmen der Integration ist es sinnvoll, es gleich Null zu setzen. Vorstellung: Die Fläche einer Geraden in der Ebene ist Null, auch wenn sie unendlich lang ist.)

Wir kommen zum 2. Schritt in den Grundprinzipien der Integration. Frage:

In welchem Sinn sollen wir allgemeine Funktionen durch Treppen approximieren?

In Analysis I verwendeten wir die gleichmäßige Konvergenz.

Analysis II: Gleichmäßige Konvergenz $\Longleftrightarrow$ Konvergenz bezüglich der Supremumsnorm:
Mit $\|g\|_\infty = \sup_x |g(x)|$ ist

$$t_m \xrightarrow{m\to\infty} f \text{ gleichmäßig} \iff \|f - t_m\|_\infty \xrightarrow{m\to\infty} 0\,.$$

Problem: Leider sind manche Funktionen, die wir integrieren wollen, nicht gleichmäßig durch Treppenfunktionen approximierbar.
Idee: $t_m \to f$ soll bedeuten, dass das Volumen zwischen dem Graphen von t_m und von f gegen Null geht (alles für $m \to \infty$). Denn es kommt uns ja am Ende nur auf Integration (Volumenberechnung) an!

Problem: »Volumen« ist bisher nicht definiert. Wir drehen uns also im Kreis!
Idee: Wir haben zwar bisher keinen Begriff davon, was »Volumen« mathematisch bedeutet, aber wir können leicht mathematisch fassen, was es bedeuten soll, »ein Volumen kleiner-gleich ε zu haben«. Das werden wir auf $g = f - t_m$ anwenden.

Wie? Wir verstecken die Ordinatenmenge von $|g|$ unter Quadern und addieren deren Volumina. Ist die Summe $\leq \varepsilon$, sagen wir, dass $|g|$ »kleiner als ε bezüglich der L^1-Norm« ist, und verwenden das als Maß für die Güte der Approximation von t_m an f.

1.2.2 Definition

Sei $g\colon \mathbb{R}^n \longrightarrow \overline{\mathbb{R}}$. Eine **Hüllreihe** für g ist eine Reihe

$$\boldsymbol{\Phi} := \sum_{k=1}^{\infty} c_k \chi_{Q_k}$$

wobei $c_k \geq 0$, Q_k offene Quader für alle k, $|g(x)| \leq \Phi(x)$ für alle x.

Bemerkung: Da alle $c_k \chi_{Q_k} \geq 0$ sind, ist der Wert der Reihe $\sum_{k=1}^{\infty} c_k \chi_{Q_k}(x) \in \overline{\mathbb{R}}$ wohldefiniert für jedes x. Jedes g hat eine Hüllreihe, zum Beispiel $\Phi = \sum_{k=1}^{\infty} \chi_{(-k,k)^n}$, denn $\Phi(x) = \infty$ für alle x. Hierbei ist $(-k,k)^n = \underbrace{(-k,k) \times \cdots \times (-k,k)}_{n\text{-mal}}$. Zum Beispiel: $(-1,1) \times (-1,1) =$ Quadrat.

1.2.3 Definition

Sei $g\colon \mathbb{R}^n \longrightarrow \overline{\mathbb{R}}$. Die **$L^1$-Halbnorm** von g ist

$$\|g\|_1 := \inf\{I(\Phi)\colon \Phi \text{ ist Hüllreihe von } g\}$$

wobei $I(\Phi) := \sum_{k=1}^{\infty} c_k \operatorname{vol}(Q_k)$ für $\Phi = \sum_{k=1}^{\infty} c_k \chi_{Q_k}$.

1.2.4 Lemma

$\|\cdot\|_1\colon \mathcal{F}(\mathbb{R}^n, \overline{\mathbb{R}}) \longrightarrow \overline{\mathbb{R}}$ ist eine Halbnorm, d. h. $\forall g, h \in \mathcal{F}(\mathbb{R}^n, \overline{\mathbb{R}})$, $c \in \overline{\mathbb{R}}$:

(a) $0 \leq \|g\|_1 \leq \infty$

(b) $\|c \cdot g\|_1 = |c| \cdot \|g\|_1$

(c) $\|g + h\|_1 \leq \|g\|_1 + \|h\|_1$

Außerdem gilt:

(d) $|g| \leq |h| \Longrightarrow \|g\|_1 \leq \|h\|_1$

(e) $\|\sum_{i=1}^{\infty} g_i\|_1 \leq \sum_{i=1}^{\infty} \|g_i\|_1$, falls alle $g_i \geq 0$.

Bemerkung: Wieso lassen wir unendliche Reihen (statt endlicher Summen) bei der Definition von $\|\cdot\|_1$ zu? Weil sonst (e) nicht gelten würde. Teil (e) wird wesentlich sein für die Konvergenzsätze.

Beispiel: (Zentral zum Verständnis alles Folgenden!)
$A = \mathbb{Q} \cap [0,1]$, $g := \chi_A$, also $g(x) = 1$, falls x eine rationale Zahl zwischen Null und Eins ist, sonst $g(x) = 0$.

A ist abzählbar, $A = \{a_1, a_2, \dots\}$.

Setze $g_i := \chi_{\{a_i\}}$, dann ist $\|g_i\|_1 = 0$, denn $\Phi_\varepsilon = \chi_{(a_i-\varepsilon, a_i+\varepsilon)}$ ist eine Hüllreihe für g_i für jedes $\varepsilon > 0$, und $I(\Phi_\varepsilon) = 2\varepsilon$, also $\|g_i\|_1 = \inf\{I(\Phi) : \Phi \text{ ist Hüllreihe von } g_i\} = 0$.

Aus $g = \sum_{i=1}^{\infty} g_i$ und (e) folgt

$$\|g\|_1 \le \sum_{i=1}^{\infty} 0 = 0\,, \quad \text{also} \quad \|g\|_1 = 0\,.$$

Bemerkung: Hätte man nur endliche Summen als Hüllreihen zugelassen, käme im Beispiel nicht Null, sondern Eins heraus, denn für $K < \infty$ ist $g \le \Phi = \sum_{k=1}^{K} c_k \chi_{Q_k} \Longrightarrow I(\Phi) \ge 1$. (Übung!)

Beweis (von Lemma 1.2.4): (a) ist klar, (b) ist auch einfach, (c) ist ein Spezialfall von (e).
Beweis von (d): Jede Hüllreihe von h ist auch eine Hüllreihe von g, also ok.
Beweis von (e): Sei $\varepsilon > 0$. Zu jedem i wähle eine Hüllreihe Φ_i für g_i mit $I(\Phi_i) \le \|g_i\|_1 + \varepsilon 2^{-i}$. Dann ist $\Phi = \sum_{i=1}^{\infty} \Phi_i$ eine Hüllreihe für g mit

$$I(\Phi) = \sum_{i=1}^{\infty} I(\Phi_i) \le \sum_{i=1}^{\infty} (\|g_i\|_1 + \varepsilon 2^{-i}) = \varepsilon + \sum_{i=1}^{\infty} \|g_i\|\,.$$

(Genaugenommen muss man hier in der ersten Gleichheit eine Doppelreihe umordnen, da jedes $I(\Phi_i)$ durch eine Reihe gegeben ist. Da aber alle Terme ≥ 0 sind, ist das kein Problem.) Also gibt es zu jedem $\varepsilon > 0$ eine Hüllreihe $\Phi = \Phi_\varepsilon$ für g mit $I(\Phi) \le \varepsilon + \sum_{i=1}^{\infty} \|g_i\|$. Daraus folgt (e). □

Im Beispiel von oben sieht das so aus: Sei $\varepsilon > 0$. Wähle $\Phi_\varepsilon = \sum \chi_{\left(a_i - \frac{\varepsilon}{2^{i+1}}, a_i + \frac{\varepsilon}{2^{i+1}}\right)}$.

Dann:

$$I(\Phi_\varepsilon) = \sum 2\frac{\varepsilon}{2^{i+1}} = 2\varepsilon \underbrace{\sum_{i=1}^{\infty} \frac{1}{2^{i+1}}}_{=1/2} = \varepsilon$$

Φ_ε Hüllreihe von $g \Longrightarrow \|g\|_1 = 0$.

Bemerkung:
Unendliche Φ-Reihen $\rightsquigarrow$ Lebesgue-Theorie des Integrals.
Endliche Φ-Reihen $\rightsquigarrow$ Riemann-Theorie des Integrals.

Bemerkung: Falls es einen Begriff von Summe über überabzählbar viele Terme gäbe, so könnte damit zum Beispiel (e) nicht gelten:

$$[0,1] = \bigcup_{x \in [0,1]} \{x\}$$

$$\left[\left[\underbrace{\chi_{[0,1]}}_{\|\ \|_1 = 1} = \underbrace{\sum_{x \in [0,1]} \chi_{\{x\}}}_{\|\ \|_1 = 0} \right]\right] \text{ gilt nicht!}$$

Der Unterschied abzählbar $\leftrightarrow$ überabzählbar spielt also in der Integrationstheorie eine wesentliche Rolle: Unter abzählbar vielen Punkten kann sich keine positive Fläche verbergen, unter überabzählbar vielen aber doch! (Wir werden später aber sehen, dass es gewisse überabzählbare Mengen gibt, unter denen sich auch keine Fläche verbirgt.)

Der Satz von Heine-Borel

Im nächsten Abschnitt benötigen wir eine Eigenschaft kompakter Mengen, die ich in Analysis II nicht behandelt habe. Sie wird von jetzt an mehrfach gebraucht, daher soll sie hier diskutiert werden. Für den Moment begeben wir uns also wieder in die Welt der metrischen Räume. Wenn Ihnen das zu abstrakt ist, können Sie sich immer $X = \mathbb{R}^n$ mit der euklidischen Metrik vorstellen.

Zur Erinnerung: Eine Teilmenge A eines metrischen Raumes heißt kompakt, wenn jede Folge in A einen Häufungspunkt in A hat. Ist der metrische Raum der $\mathbb{R}^n$, so ist das dazu äquivalent, dass A abgeschlossen und beschränkt ist.

1.2.5 Definition

Sei (X, d) ein metrischer Raum und $A \subset X$. Eine **offene Überdeckung** von A ist eine Familie von offenen Teilmengen $(U_i)_{i\in I}$ von X mit

$$A \subset \bigcup_{i\in I} U_i$$

wobei I eine beliebige Indexmenge ist.

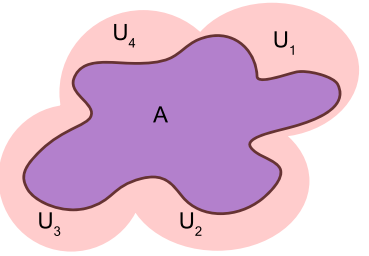


Abbildung 1.1. Die Familie $U_1, \ldots, U_4$ überdeckt A

1.2.6 Satz

Sei (X, d) metrischer Raum und $A \subset X$, dann sind äquivalent:

(1) A ist kompakt.

(2) Jede offene Überdeckung von A hat eine endliche Teilüberdeckung. Das heißt:
Für *jede* offene Überdeckung $(U_i)_{i\in I}$ von A exististieren endlich viele $i_1, \ldots, i_N \in I$, so dass $A \subset U_{i_1} \cup \ldots \cup U_{i_N}$.

Beispiele:

(1) (Nicht-Beispiel): Seien $A = (0, 1]$ und $U_i = (\frac{1}{i}, 2)$ mit $i \in \mathbb{N}$, dann gilt $A \subset \bigcup_{i\in\mathbb{N}} U_i$, denn für $x \in A$ gilt $0 < x \leq 1$, und aus dem Archimedischen Prinzip folgt, dass es ein $i \in \mathbb{N}$ gibt mit $\frac{1}{i} < x$, also $x \in U_i$. Allerdings überdecken beliebige endlich viele der U_i nicht ganz A, denn seien $i_1, \ldots, i_N$ beliebig und $m = \max\{i_1, \ldots, i_N\}$, dann ist offenbar $U_{i_1} \cup \ldots \cup U_{i_N} \subset (\frac{1}{m}, 2)$. Also folgt $\frac{1}{2m} \in A$, aber $\frac{1}{2m} \notin U_{i_1} \cup \ldots \cup U_{i_N}$.

(2) Seien $A = [0, 1]$ und $U_i = (\frac{1}{i}, 2)$ mit $i \in \mathbb{N}$. Damit 0 auch überdeckt wird brauchen wir eine weitere offene Menge U_0 mit $0 \in U_0$. Dann existiert wegen der Offenheit von U_0 ein $\varepsilon > 0$, so dass gilt $(-\varepsilon, \varepsilon) \subset U_0$. Wir wählen i_0 mit $\frac{1}{i_0} < \varepsilon$. Somit bildet $U_0 \cup U_{i_0}$ eine endliche offene Teilüberdeckung von A.

Bemerkung: In Beispiel (2) ist *nicht* die Kompaktheit von $[0,1]$ bewiesen. Für diese müssten wir mit einer *beliebigen* offenen Überdeckung anfangen.

Beweis (des Satzes): Wir beweisen zuerst die Richtung (2) $\Rightarrow$ (1), nehmen also an, dass (2) gilt, und zeigen, dass A kompakt ist. Angenommen, A ist nicht kompakt. Dann gibt es eine Folge $(x_k)_{k\in\mathbb{N}}$ in A ohne einen Häufungspunkt in A. Das heißt, für jedes $x \in A$ existiert ein ε_x, so dass $K_{\varepsilon_x}(x)$ nur endlich viele Folgenglieder enthält.

Wir setzen $U_x = K_{\varepsilon_x}(x)$; dann sind alle U_x offen. Offenbar ist $A \subset \bigcup\limits_{x\in A} U_x$, denn für $x \in A$ gilt $x \in K_{\varepsilon_x}(x)$, also auch $x \in \bigcup\limits_{y\in A} U_y$.

Da wir Teil (2) des Satzes voraussetzen, gibt es $y_1, \ldots, y_N \in A$, so dass $A \subset U_{y_1} \cup \ldots \cup U_{y_N}$.
Per Konstruktion enthält jedes U_{y_i} nur endlich viele Folgenglieder, also auch $U_{y_1} \cup \ldots \cup U_{y_N}$ und damit auch A. Dies steht im Widerspruch dazu, dass die komplette unendliche Folge in A liegt.
Also ist (2) $\Rightarrow$ (1) gezeigt.

Für den Beweis von (1) $\Rightarrow$ (2) zeigen wir zuerst eine Zwischenbehautung:
Zwischenbehauptung: Sei A eine kompakte Menge, dann gibt es zu jedem $\varepsilon > 0$ endlich viele ε-Kugeln, die A überdecken.

Beweis (der Zwischenbehauptung): A sei kompakt. Angenommen, es gibt ein $\varepsilon > 0$, so dass A nicht von endlichen vielen ε-Kugeln überdeckt werden kann.

Wir wählen $x_1 \in A$ beliebig, dann ist $A \not\subset K_\varepsilon(x_1)$ per Annahme. Also existiert $x_2 \in A \setminus K_\varepsilon(x_1)$. Per Annahme überdecken $K_\varepsilon(x_1)$ und $K_\varepsilon(x_2)$ nicht ganz A, also existiert
$x_3 \in A \setminus (K_\varepsilon(x_1) \cup K_\varepsilon(x_2))$. Analog erhält man
$x_4 \in A \setminus (K_\varepsilon(x_1) \cup \ldots \cup K_\varepsilon(x_3))$
$\vdots$

Dies ergibt eine Folge $(x_k)_{k\in\mathbb{N}}$ in A, wobei $d(x_k, x_l) \geq \varepsilon$ für alle $k > l$. Daher kann $(x_k)_{k\in\mathbb{N}}$ keinen Häufungspunkt haben und somit kann A nicht kompakt sein. Also haben wir die Zwischenbehauptung bewiesen. Der Beweis zeigt auch, dass es egal ist, ob wir dabei offene oder abgeschlossene Kugeln betrachten, und dass die Mittelpunkte der Kugeln in A gewählt werden können. □

Wir führen jetzt den Beweis von (1) $\Rightarrow$ (2) im Satz. Sei A kompakt und $(U_i)_{i\in I}$ eine Überdeckung von A. Angenommen, es gibt *keine* endliche Teilüberdeckung zu $(U_i)_{i\in I}$.
(1) Wir überdecken A mit endlich vielen abgeschlossenen 1-Kugeln (alle auftretenden Kugeln sollen Mittelpunkte in A haben), dann wird der Schnitt jeder dieser 1-Kugeln mit A von $(U_i)_{i\in I}$ überdeckt.

Also gibt es unter diesen mindestens eine 1-Kugel K_1, so dass $A_1 = K_1 \cap A$ keine endliche Teilüberdeckung besitzt (sonst hätte A eine endliche Teilüberdeckung).

Da A kompakt ist, ist es abgeschlossen, und da K_1 abgeschlossen ist, ist auch A_1 abgeschlossen. Da $A_1 \subset A$ und abgeschlossene Teilmengen kompakter Mengen kompakt sind, ist A_1 kompakt.
(2) Wir überdecken A_1 nun mit endlich vielen abgeschlossenen $\frac{1}{2}$-Kugeln.

Wie vorher muss es nun darunter mindestens eine $\frac{1}{2}$-Kugel K_2 geben, so dass $A_2 = K_2 \cap A_1$ keine endliche Teilüberdeckung hat.
(3), (4), ... Auf dieselbe Weise erhalten wir $A_3, A_4, \ldots$. Diese haben also folgende Eigenschaften:
$A \supset A_1 \supset A_2 \supset \ldots$; A_k ist in in einer Kugel K_k vom Radius $1/k$ enthalten; keines der A_k besitzt eine Überdeckung durch endlich viele der Mengen U_i.

Sei x_k der Mittelpunkt der Kugel K_k. Offenbar ist $(x_k)_{k\in\mathbb{N}}$ eine Folge in A. Da A kompakt folgt, dass eine konvergente Teilfolge $x_{k'}$ existiert, $x_{k'} \to x \in A$. Also existiert ein $i \in I$ mit $x \in U_i$.

Da U_i offen ist, existiert ein ε, so dass $K_\varepsilon(x) \subset U_i$. Wegen $x_{k'} \to x$ existiert ein k' mit $d(x, x_{k'}) < \frac{\varepsilon}{2}$ und $\frac{1}{k'} < \frac{\varepsilon}{2}$, also $K_{k'} \subset K_\varepsilon(x)$. Daraus folgt $A_{k'} \subset K_{k'} \subset K_\varepsilon(x) \subset U_i$.
Also wird $A_{k'}$ von der einen Menge U_i überdeckt, im Widerspruch zur Konstruktion der A_k. □

Die L^1-Halbnorm von Treppenfunktionen

Zentral für alles folgende ist der folgende unscheinbare Sachverhalt:

1.2.7 Lemma

Sei φ eine Treppenfunktion. Dann gilt $\|\varphi\|_1 = \int |\varphi|\, dx$.

Beweis: Wegen $\|\varphi\|_1 = \| \, |\varphi| \, \|_1$ können wir o. B. d. A. annehmen, dass $\varphi \geq 0$.
1. Schritt: Es ist ziemlich klar, dass $\rangle \leq \langle$ gilt, da man φ als Hüllreihe für sich selbst nehmen kann und für Treppenfunktionen $\varphi \geq 0$ offenbar $I(\varphi) = \int \varphi$ gilt.

Kleines Detail: Für Hüllreihen war angenommen worden, dass sie mit offenen Quadern gebildet werden. Das ist aber unwesentlich: Wenn $\varphi = \sum_{k=1}^{N} a_k \chi_{P_k}$ mit beliebigen (evtl. nicht offenen) Quadern P_k und $a_k \geq 0$, dann wähle zu $\varepsilon > 0$ offene Q_k mit $\operatorname{vol}(Q_k) \leq \operatorname{vol}(P_k) + \varepsilon$. Setze $\Phi = \sum a_k Q_k$, dann gilt $\|\varphi\|_1 \leq I(\Phi) = \sum a_k \operatorname{vol}(Q_k) \leq \sum a_k \operatorname{vol}(P_k) + \sum a_k \varepsilon = \int \varphi + \varepsilon\Big(\sum_{k=1}^{K} a_k\Big)$. Da dies für alle $\varepsilon > 0$ gilt, folgt $\|\varphi\|_1 \leq \int \varphi$.

Es bleibt, die Ungleichung $\|\varphi\|_1 \geq \int \varphi$ zu zeigen. Dies ist der eigentliche Kern des Lemmas.

Vorbemerkung: Für *endliche* Hüllreihen $\Phi = \sum_{k=1}^{K} c_k \chi_{Q_k}$ (also Treppenfunktionen) wissen wir:
Aus $\varphi \leq \Phi$ folgt $\int \varphi \leq \int \Phi = I(\Phi)$ (Monotonie des Integrals, Satz 1.1.7).
Das heißt: $\int \varphi \leq \inf\{I(\Phi) : \Phi \text{ ist endliche Hüllreihe von } \varphi\}$
Also ist das einzige Problem der Übergang von endlich nach abzählbar.

2. Schritt: Wir betrachten zunächst den Fall $\varphi = \chi_Q$, wobei Q ein abgeschlossener Quader ist. Angenommen, $\Phi = \sum_{k=1}^{\infty} c_k \chi_{Q_k}$ ist eine Hüllreihe für χ_Q, d. h. $\chi_Q \leq \Phi$.

Sei $\varepsilon > 0$. Für jedes $x \in Q$ gilt $1 \leq \sum_{k=1}^{\infty} c_k \chi_{Q_k}(x)$. Nach Definition von Reihen existiert also ein N_x, so dass $1 - \varepsilon \leq \sum_{k=1}^{N_x} c_k \chi_{Q_k}(x)$.

Da alle Q_k offen sind und der Schnitt endlich vieler offener Mengen wieder offen ist, existiert eine offene Umgebung U_x von x, die in $Q_1, \ldots, Q_{N_x}$ enthalten ist. Dann ist $\sum_{k=1}^{N_x} c_k \chi_{Q_k}$ auf U_x konstant, also folgt $1 - \varepsilon \leq \sum_{k=1}^{N_x} c_k \chi_{Q_k}$ auf U_x. Offenbar ist $(U_x)_{x \in Q}$ eine offene Überdeckung von Q. Außerdem ist Q kompakt, da abgeschlossen und beschränkt und in $\mathbb{R}^n$. Also folgt die Existenz einer endlichen Teilüberdeckung $U_{x_1}, \ldots U_{x_m}$. Wähle $N = \max_{i=1,\ldots,M} N_{x_i}$, dann folgt $1 - \varepsilon \leq \sum_{k=1}^{N} c_k \chi_{Q_k}(x)$ für alle $x \in Q$. Da dies eine endliche Summe ist, folgt (wie in der Vorbemerkung), dass $(1-\varepsilon)\operatorname{vol}(Q) \leq \sum_{k=1}^{N} c_k \operatorname{vol}(Q_k) \leq \sum_{k=1}^{\infty} c_k \operatorname{vol}(Q_k) = I(\Phi)$.

Wir haben gezeigt, dass $(1-\varepsilon)\int \chi_Q \leq I(\Phi)$ für alle $\varepsilon > 0$, also $\int \chi_Q \leq I(\Phi)$. Da dies für jede Hüllreihe Φ von χ_Q gilt, haben wir $\int \chi_Q \leq \|\chi_Q\|_1$ bewiesen.

3. Schritt: Wir müssen nun $\int \varphi \leq \|\varphi\|_1$ für beliebige Treppenfunktionen $\varphi \geq 0$ zeigen. Man kann Schritt für Schritt argumentieren, es ist aber etwas umständlich, da man mit offenen/abgeschlossenen Quadern aufpassen muss. Hier ist ein hübscher Trick:

Wähle einen abgeschlossenen Quader A, der so groß ist, dass $\varphi = 0$ außerhalb von A ist. Setze $M = \max \varphi$ und $\psi = M\chi_A - \varphi$. Dann ist $\psi \geq 0$.
Wende Schritt 1 auf ψ und Schritt 2 auf $M\chi_A$ an, dann folgt

$$\int \psi \geq \|\psi\|_1$$
$$\int M\chi_A \leq \|M\chi_A\|_1 .$$

Subtrahiert man die Ungleichungen (das geht wegen der verschiedenen Richtungen), folgt

$$\int \varphi = \int (M\chi_A - \psi) \leq \|M\chi_A\|_1 - \|\psi\|_1 \leq \|\varphi\|_1 ,$$

wobei am Schluss nur eine Umstellung der Dreiecksungleichung für $M\chi_A = \varphi + \psi$ verwendet wurde. Dies war zu zeigen. □

Bemerkung: Der schwierigere Teil der Behauptung, $\int |\varphi| \leq \|\varphi\|_1$, sagt im Wesentlichen folgendes: Überdeckt man einen Quader Q mit abzählbar vielen Quadern Q_i, so ist die Summe der Volumina der Q_i mindestens gleich dem Volumen von Q. (Man sieht die Relation mit der Aussage des Lemmas am einfachsten, wenn man das Lemma auf $\chi_{[0,1]}$ anwendet und den Graphen von $\chi_{[0,1]}$ und Φ zeichnet.)

Das war ziemlich viel Aufwand für ein scheinbar so offensichtliches Lemma! Dass wir immerhin den Satz von Heine-Borel gebraucht haben, legt nahe, dass es gar nicht so offensichtlich ist. Bei Unendlichkeiten kann eben leicht etwas schief gehen, daher muss man genau argumentieren.

1.3 Definition und einfache Eigenschaften des Lebesgue-Integrals

1.3.1 Definition

Eine Funktion $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ heißt (Lebesgue-) integrierbar, falls es Treppenfunktionen $\varphi_1, \varphi_2, \ldots$ gibt mit $\varphi_k \to f$ bezüglich $\|\cdot\|_1$, das heißt $\|f - \varphi_k\|_1 \xrightarrow{k\to\infty} 0$.
In diesem Fall setze $\int \boldsymbol{f\,dx} \coloneqq \lim\limits_{k\to\infty} \int \varphi_k\,dx$.

Bemerkung: Es bleiben die Fragen:

- ▷ Ist das wohldefiniert?
- ▷ Sind vielleicht alle Funktionen integrierbar?
- ▷ Wie berechnet man $\int f\,dx$?

Schreibweise: $\boldsymbol{\mathcal{L}^1} = \mathcal{L}^1(\mathbb{R}^n) = \{\text{Integrierbare Funktionen auf } \mathbb{R}^n\}$.

Die Frage, ob $\int f\,dx$ wohldefiniert (d. h. unabhängig von der Wahl der Folge φ_k ist, sowie einige weitere wichtige Eigenschaften des Integrals folgen aus einfachen Überlegungen über Vektorräume und Halbnormen. Die Definition von $\mathcal{L}^1$ ist nämlich ein Beispiel eines sehr allgemeinen Konstruktionsprinzips:

$\mathcal{L}^1 =$ der Abschluss von $\mathcal{T}(\mathbb{R}^n)$ in $\mathcal{F}(\mathbb{R}^n, \overline{\mathbb{R}})$ bezüglich $\|\cdot\|_1$

Wir betrachten zunächst eine analoge Situation für beliebige Vektorräume. Ist V ein Vektorraum und $\|\cdot\| : V \to [0,\infty]$ eine Halbnorm auf V, so hat man genau wie bei einem normierten Vektorraum die topologischen Grundbegriffe wie offene, abgeschlossene Mengen, Abschluss, Inneres, Stetigkeit etc. Konkret ist zum Beispiel der Abschluss einer Menge $A \subset V$ definiert durch

$$\overline{A} = \{f \in V : \underset{\text{Folge } (\varphi_k) \text{ in } A}{\exists} \text{ mit } \|f - \varphi_k\| \xrightarrow{k\to\infty} 0\} .$$

Ein **lineares Funktional** auf V ist eine lineare Abbildung $l : V \to \mathbb{R}$. l heißt **beschränkt**, falls es eine Zahl C gibt mit

$$|l(f)| \leq C\|f\| \quad \text{für alle } f \in V\,.$$

(Vorsicht: Dies ist etwas anders als die Definition einer beschränkten Funktion. Auf der rechten Seite steht nicht C, sondern $C\|f\|$.)

Man kann leicht zeigen, dass ein lineares Funktional genau dann beschränkt ist, wenn es stetig ist (Übung).

1.3.2 Lemma (Lineare Algebra)

Seien V ein Vektorraum, $\|\cdot\|$ eine Halbnorm auf V, $V_e := \{v \in V : \|v\| < \infty\}$ und $W \subset V_e$ ein Untervektorraum.

(a) Für den Abschluss von W gilt:
- $\overline{W}$ ist ein Untervektorraum
- $\overline{W}$ ist abgeschlossen
- $\overline{W} \subset V_e$

(b) Ist $l : W \to \mathbb{R}$ ein beschränktes lineares Funktional, dann hat l eine eindeutige Fortsetzung als beschränktes lineare Funktional $\bar{l} : \overline{W} \to \mathbb{R}$, mit derselben Konstante C wie l. Es gilt

$$\bar{l}(f) = \lim_{k\to\infty} l(\varphi_k)\,, \text{ falls } \varphi_k \to f\,.$$

Beweis:

(a)
- Seien $\varphi_k \to f$ und $\psi_k \to g$. Dann folgt $\varphi_k + \psi_k \to f + g$. Also folgt aus $f, g \in \overline{W}$, dass auch $f + g \in \overline{W}$.

  Analog folgt aus $f \in \overline{W}$ und $c \in \mathbb{R}$, dass $c \in \overline{W}$.
- Wie bei normierten Vektorräumen ist der Abschluss einer Menge abgeschlossen.
- Seien $\overline{W} \subset V_e$ und $\varphi_k \to f$, dann existiert ein k, so dass $\|f - \varphi_k\| < 1$.
  Also folgt $\|f\| \leq \|\varphi_k\| + \|f - \varphi_k\| < \infty$.

(b) Übung. □

In unserem Kontext wenden wir dies an mit:

$V := \mathcal{F}(\mathbb{R}^n, \overline{\mathbb{R}}) = \{\text{alle Funktionen } \mathbb{R}^n \to \overline{\mathbb{R}}\}$

$\|\cdot\| = \|\cdot\|_1 : V \to [0, \infty]$ die L^1-Halbnorm

$W := \mathcal{T}(\mathbb{R}^n) = \{\text{Treppenfunktionen auf } \mathbb{R}^n\}$

$l : W \to \mathbb{R}$, $\varphi \mapsto \int \varphi$ ist ein beschränktes lineares Funktional, wobei $|l(\varphi)| \leq \|\varphi\|_1$ gilt, also die Beschränktheitskonstante gleich eins ist.

Offenbar gilt $W \subset V_e$, also ist das Lemma anwendbar, und man erhält, da per Definition $\mathcal{L}^1(\mathbb{R}^n) = \overline{W}$ ist:

1.3.3 Satz

(a) $\mathcal{L}^1$ ist ein Vektorraum.

(b) $\int : \mathcal{L}^1 \to \mathbb{R}$ ist wohldefiniert und linear.

(c) $|\int f| \leq \|f\|_1$.

(d) $\int : \mathcal{L}^1 \to R$ ist eine stetige Abbildung.

(e) $\mathcal{L}^1$ ist abgeschlossen, das heißt: Sind $f_k \in \mathcal{L}^1$ für alle k, $f \in \mathcal{F}(\mathbb{R}^n, \overline{\mathbb{R}})$ und gilt $f_k \xrightarrow{k\to\infty} f$ bezüglich $\|\cdot\|_1$, dann folgt $f \in \mathcal{L}^1$.

Wir brauchen noch ein paar weitere einfache Eigenschaften des Integrals:

1.3.4 Lemma

Seien $f, g \in \mathcal{L}^1$.

(a) $f \leq g \Rightarrow \int f \leq \int g$.

(b) $\max\{f,g\}$, $\min\{f,g\}$, $|f| \in \mathcal{L}^1$.

(c) $f \geq 0 \Rightarrow \int f = \|f\|_1$.

Beweis:

(c) Seien φ_k Treppenfunktionen mit $\varphi_k \to f$. O. B. d. A. sei $\varphi \geq 0$. (Denn $||\varphi_k| - f| \leq |\varphi_k - f|$, also $\| |\varphi_k| - f\|_1 \leq \|\varphi_k - f\|_1$; konvergiert also (φ_k) gegen f, so auch $(|\varphi_k|)$.)

Schreibe $f = \varphi_k + (f - \varphi_k)$, dann folgt $\|f\|_1 \leq \|\varphi_k\|_1 + \|f - \varphi_k\|_1$. Nach Lemma 1.2.7 gilt $\|\varphi_k\|_1 = \int \varphi_k$

$$\|f\|_1 \leq \lim_{k\to\infty} \|\varphi_k\|_1 + \lim_{k\to\infty} \|f - \varphi_k\|_1 = \lim_{k\to\infty} \int \varphi_k + 0 = \int f.$$

Wir haben $\|f\|_1 \leq \int f$ gezeigt. Die Umkehrung, $\int f \leq \|f\|_1$, ist Satz 1.3.3 (c).

(a) Setze $h = g - f \geq 0$. Dann folgt $\int h = \|h\|_1 \geq 0$ und damit $\int g \leq \int f$.

(b) Wie im Beweis von Lemma 1.1.5 ist

$$\min\{f,g\} = \frac{1}{2}[f + g - |f - g|], \quad \max\{f,g\} = \frac{1}{2}[f + g + |f - g|].$$

Also genügt es zu zeigen, dass aus $f \in \mathcal{L}^1$ auch $|f| \in \mathcal{L}^1$ folgt.

Es ist $||f| - |\varphi_k|| \leq |f - \varphi_k|$, also folgt $\| |f| - |\varphi_k| \|_1 \leq \|f - \varphi_k\|_1$.

Das heißt, wenn $\varphi_k \to f$ dann $|\varphi_k| \to |f|$ bezüglich $\|\cdot\|_1$. □

Bisher haben wir nur über ganz $\mathbb{R}^n$ integriert. Das lässt sich leicht erweitern.

1.3.5 Definition

Sei $A \subset \mathbb{R}^n$ und $f : A \to \overline{\mathbb{R}}$. Setze $\overline{f}(x) = \begin{cases} f(x) & x \in A \\ 0 & x \notin A. \end{cases}$

f heißt **über A integrierbar**, falls $\overline{f}$ integrierbar ist. In diesem Fall setzen wir

$$\int_A f\,dx := \int_{\mathbb{R}^n} \overline{f}\,dx.$$

Regel- und Lebesgue-Integral, I

Wie verhalten sich der alte Integralbegriff (aus Analysis I) und der neue zueinander?

1.3.6 Satz

Seien $a, b \in \mathbb{R}$ mit $a \leq b$.

(a) Jede Regelfunktion f auf $[a,b]$ ist (Lebesgue-) integrierbar und $\int_a^b f\,dx = \int_{[a,b]} f\,dx$.

(b) Es gibt (Lebegue-) integrierbare Funktionen, die keine Regelfunktionen sind.

Bemerkung: Die verschiedenen Integralbegriffe unterscheiden sich nur darin, wie groß die Menge der integrierbaren Funktionen ist:
$\{\text{Regelfunktionen}\} \subset \{\text{Rieman-int.bare Funktionen}\} \subset \{\text{Lebesgue-int.bare Funktionen}\} \subset \{\text{Alle Funktionen}\}$
Allerdings stimmt dies nur für »eigentliche« Regelintegrale. Gewisse uneigentlich regelintegrierbare Funktionen sind nicht Lebesgue-integrierbar. Dazu später mehr.

Beweis:

(a) Per Definition ist f Regelfunktion genau dann, wenn Treppenfunktionen φ_k existieren mit $\varphi_k \to f$ gleichmäßig. Das heißt $\|f - \varphi_k\|_\infty = \sup_{x\in[a,b]} |f(x) - \varphi_k(x)| \to 0$.

Hilfsbehauptung: Für beliebige beschränkte Funktionen $g : [a,b] \to \mathbb{R}$ gilt: $\|g\|_1 \leq \|g\|_\infty \cdot (b-a)$.

Beweis: Es ist $|g(x)| \leq \|g\|_\infty$ auf $[a,b]$, das heißt $|g| \leq \|g\|_\infty \cdot \chi_{[a,b]}$. Daraus folgt $\|g\|_1 \leq \|\|g\|_\infty \cdot \chi_{[a,b]}\|_1 = \|g\|_\infty \cdot \|\chi_{[a,b]}\|_1 = \|g\|_\infty (b-a)$. Somit ist die Hilfsbehauptung bewiesen.

Daher gilt $\|f - \varphi_k\|_1 \leq \|f - \varphi_k\|_\infty \cdot (b-a)$ und somit $\|f - \varphi_k\|_1 \to 0$. Also ist $f \in \mathcal{L}^1$.

Da für Treppenfunktionen das Regel- und das Lebesgue-Integral per Definition übereinstimmen, ergibt sich für das Integral: $\int_{[a,b]} f = \lim_{k\to\infty} \int_{[a,b]} \varphi_k = \lim_{k\to\infty} \int_a^b \varphi_k = \int_a^b f$.

(b) Wir geben ein Gegenbeispiel.

Sei $A = \mathbb{Q} \cap [0,1]$ und $f = \chi_A$. Offenbar ist f keine Regelfunktion, da überall unstetig (Regelfunktionen haben höchstens abzählbar viele Unstetigkeitsstellen). Allerdings ist f integrierbar, denn wir sahen bereits, dass $\|f\|_1 = 0$, also gilt mit $\varphi_k = 0$ für alle k, dass $\|f - \varphi_k\| = 0$ für alle k, also ist f Lebesgue-integrierbar.

Für das Integral von $f = \chi_A$ folgt $\int f = \lim \int \varphi_k = \lim 0 = 0$.
Zum Vergleich: Sei $B = [0,1] \setminus A$ und $g = \chi_B = \chi_{[0,1]} - \chi_A$, dann ist $\int g = \int \chi_{[0,1]} - \int \chi_A = 1 - 0 = 1$.

1.4 Nullmengen, Nullfunktionen und »fast überall«

Wir behandeln nun die Frage, an wieviel Punkten man eine Funktion höchstens abändern darf, damit sich ihr Integral nicht ändert.

1.4.1 Definition

▷ Eine **Nullmenge** ist eine Menge $N \subset \mathbb{R}^n$ mit $\|\chi_N\|_1 = 0$.

▷ Eine **Nullfunktion** ist eine Funktion $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ mit $\|f\|_1 = 0$.

1.4.2 Lemma

(a) Jede Teilmenge einer Nullmenge ist eine Nullmenge.

(b) Jede abzählbare Vereinigung von Nullmengen ist eine Nullmenge.

Beweis:

(a) Sei N Nullmenge und $N' \subset N$. Daraus folgt $0 \leq \chi_{N'}(x) \leq \chi_N(x)$ für alle x. Also $\|\chi_{N'}\|_1 \leq \|\chi_N\|_1 = 0$. Somit folgt $\|\chi_{N'}\| = 0$.

(b) Seien $N_1, N_2, \ldots$ Nullmengen und $N = \bigcup_{k=1} N_k$, dann ist offenbar $\chi_N \leq \sum_{k=1}^{\infty} \chi_{N_k}$. Also folgt

$$\|\chi_N\| = \|\sum_{k=1}^{\infty} \chi_{N_k}\|_1 \leq \sum_{k=1}^{\infty} \|\chi_{N_k}\|_1 = \sum_{k=1}^{\infty} 0 = 0. \qquad \square$$

Beispiele:

▷ Jede abzählbare Menge ist eine Nullmenge, denn $\{p\}$ mit $p \in \mathbb{R}$ ist Nullmenge.

▷ Sei $H = \{x \in \mathbb{R}^n : x_1 = 0\}$. H ist eine Nullmenge, denn

$$H = \bigcup_{k_1,\ldots,k_n \in \mathbb{Z}} \{(x_1, \ldots, x_n) : x_1 = 0, x_2 \in [k_2, k_1 + 1), x_3 \in [k_3, k_3 + 1), \ldots, x_n \in [k_n, k_n + 1)\}.$$

Das Volumen jedes dieser Quader ist Null, denn es ist gleich $0 \cdot 1 \cdot \ldots \cdot 1 = 0$.

▷ $[0, 1] \subset \mathbb{R}$ ist keine Nullmenge, denn $\|\chi_{[0,1]}\|_1 = 1$.

1.4.3 Satz

Für Mengen $N \subset \mathbb{R}^n$ sind äquivalent:

(1) Für alle $\varepsilon > 0$ gibt es Quader $Q_1, Q_2, \ldots$ mit

(a) $\sum_{i=1}^{\infty} \text{vol}(Q_i) < \varepsilon$

(b) $N \subset \bigcup_{i=1}^{\infty} Q_i$

(2) N ist Nullmenge.

Beweis: Aus (1) folgt (2): Sei $\varepsilon > 0$. Finde Q_i wie in (1). Aus $N \subset \bigcup_{i=1}^{\infty} Q_i$ folgt $||\chi_N||_1 \leq ||\sum_i \chi_{Q_i}||_1 \leq \sum_i ||\chi_{Q_i}||_1 = \sum_i \text{vol}(Q_i) < \varepsilon$. Also $||\chi_N||_1 \leq \varepsilon$ für alle $\varepsilon > 0$. Somit folgt $\|\chi_N\|_1 = 0$.

Aus (2) folgt (1): Sei N Nullmenge, also $\|\chi_N\|_1 = 0$. Nach der Definition der L^1-Norm gibt es dann zu jedem $\varepsilon > 0$ Quader Q_i und Zahlen $c_i \geq 0$, $i \in \mathbb{N}$, mit $\chi_N \leq \Phi := \sum_{i=1}^{\infty} c_i \chi_{Q_i}$, $\sum_{i=1}^{\infty} c_i \,\text{vol}(Q_i) < \varepsilon$.

Diese Q_i funktionieren leider nicht für den Beweis. Sie müssen zwar N überdecken, aber aus $\sum_{i=1}^{\infty} c_i \,\text{vol}(Q_i) < \varepsilon$ folgt nicht $\sum_{i=1}^{\infty} \text{vol}(Q_i) < \varepsilon$, da die c_i sehr klein sein können.

Was tun?

Erste Idee: Falls nur endlich viele Q_i vorkommen, argumentiere wie folgt: Die Menge $D = \{x : \Phi(x) \geq 1\}$ ist eine endliche Vereinigung disjunkter Quader Q_j' (siehe das Lemma unten). Außerdem ist $N \subset D$. Die Quader Q_j' überdecken also N. Nach Definition von D ist $\chi_D \leq \Phi$, also ist $\sum_j \text{vol}(Q_j') = \|\chi_D\|_1 \leq \|\Phi\|_1 < \varepsilon$.

Zweite Idee: Falls unendlich viele Q_i vorkommen, führen wir die Situation auf den endlichen Fall wie folgt zurück: Sei $\Phi_k = \sum_{i=1}^{k} c_i \chi_{Q_i}$ und $D_k = \{x : \Phi_k \geq 1/2\}$. Offensichtlich gilt $D_1 \subset D_2 \subset \ldots$, und aus $\sum_{i=1}^{\infty} c_i \chi_{Q_i}(x) \geq 1$ für $x \in N$ und der Definition der Konvergenz einer Reihe folgt $\cup_{k=1}^{\infty} D_k \supset N$.

Also $N \subset (D_1 \setminus D_0) \cup (D_2 \setminus D_1) \cup (D_3 \setminus D_2) \cup \ldots$ (mit $D_0 := \emptyset$).

Da die Φ_k Treppenfunktionen sind, ist jedes $D_k \setminus D_{k-1}$ eine disjunkte Vereinigung $\bigcup_j Q'_{k,j}$ endlich vieler Quader (Argument ähnlich wie im folgenden Lemma).

Die abzählbar vielen Quader $Q'_{i,j}$ (über alle i, j) überdecken N. Weiterhin ist nach Definition von D_k: $\chi_{D_k} \leq 2\Phi_k$, also $\sum_{i=1}^{k} \sum_j \mathrm{vol}(Q'_{i,j}) = \|\chi_{D_k}\|_1 \leq 2\|\Phi_k\|_1 \leq 2\varepsilon$ für alle k und damit auch $\sum_{i=1}^{\infty} \sum_j \mathrm{vol}(Q'_{i,j}) \leq 2\varepsilon$. □

1.4.4 Lemma

Sei ϕ eine Treppenfunktion. Dann lassen sich die Mengen $\{\phi > c\}$, $\{\phi \geq c\}$, $c \in \mathbb{R}$, als disjunkte Vereinigungen endlich vieler Quader schreiben.

Beweis: $\{\phi > c\}$ ist die Menge der Punkte, wo die Treppenfunktion $\psi = \max\{\phi - c, 0\}$ strikt positiv ist. Schreibe $\psi = \sum_i c_i \chi_{Q_i}$ mit endlich vielen, disjunkten Q_i. Dann ist $\{\phi > c\}$ die Vereinigung der Q_i mit $c_i > 0$.

Da eine Treppenfunktion nur endlich viele Werte annimmt, ist $\{\phi \geq c\} = \{\phi > c - \varepsilon\}$ für genügend kleines $\varepsilon > 0$, damit folgt die Behauptung auch für $\{\phi \geq c\}$. □

Beispiel: Der Graph einer stetigen Funktion $f : [a, b] \to \mathbb{R}$ ist eine Nullmenge in $\mathbb{R}^2$. (Übung)

Frage: Gibt es überabzählbare Nullmengen in $\mathbb{R}$? Ja!

Beispiel: Die **Cantor-Menge** $C = \left\{x \in \mathbb{R} : x = \sum_{i=1}^{\infty} a_i 3^{-i}, \text{ alle } a_i \in \{0, 2\}\right\}$ ist eine überabzählbare Nullmenge (Übung).

Wir untersuchen nun Nullfunktionen. Sie lassen sich leicht mittels Nullmengen charakterisieren:

1.4.5 Lemma

$f : \mathbb{R}^n \to \overline{\mathbb{R}}$ ist eine Nullfunktion $\iff$ $A = \{x : f(x) \neq 0\}$ ist eine Nullmenge.

Beweis: Zuerst die leichtere Richtung ›⇐‹: Sei A Nullmenge. Wegen $|f(x)| \leq \infty$ für $x \in A$ und $f(x) = 0$ für $x \notin A$ ist $|f| \leq \infty \cdot \chi_A = \sum_{i=1}^{\infty} \chi_A$, also $\|f\|_1 \leq \sum_{i=1}^{\infty} \|\chi_A\|_1 = \sum_{i=1}^{\infty} 0 = 0$.
Zeige nun ›⇒‹: Idee: Ist $|f(x)| \geq 1$ für alle $x \in A$, so folgt $|f| \geq \chi_A$, also $0 = \|f\|_1 \geq \|\chi_A\|_1$, also ist A Nullmenge.

Was tun für allgemeines f? Zerlege in abzählbar viele analoge Fälle: Sei $A_k = \left\{x : |f(x)| > \frac{1}{k}\right\}$, dann folgt $|f| > \frac{1}{k}\chi_{A_k} \Rightarrow 0 = \|f\|_1 \geq \frac{1}{k}\|\chi_{A_k}\|_1$, also ist A_k Nullmenge. Damit ist auch $A = \bigcup_{k=1}^{\infty} A_k$ Nullmenge. □

Wie steht's mit dem Wert ∞?

1.4.6 Lemma

Sei $f : \mathbb{R}^n \to \overline{\mathbb{R}}$. Falls $\|f\|_1 < \infty$, so ist $U = \{x : f(x) = \infty\}$ eine Nullmenge.

Beweis: Für alle $k \in \mathbb{N}$ gilt $|f| \geq k\chi_U$, also $\|f\|_1 \geq k\|\chi_U\|_1$. Wäre $\|\chi_U\|_1 > 0$, so folgte $\|f\|_1 = \infty$. □

Wir können nun die eingangs gestellte Frage beantworten.

1.4.7 Satz

Seien $f, g : \mathbb{R}^n \to \overline{\mathbb{R}}$. Falls $\{x : f(x) \neq g(x)\}$ eine Nullmenge ist, und falls f integrierbar ist, so ist auch g integrierbar, und es gilt $\int f\,dx = \int g\,dx$.

Offenbar ist dies im Allgemeinen nicht richtig, falls $A = \{f \neq g\}$ keine Nullmenge ist. (Wähle $f = 0$, $g = \chi_A$.)

Beweis: Nach Lemma 1.4.5 ist $\|g - f\|_1 = 0$.
Nach Definition gilt: f integrierbar $\Leftrightarrow$ Es existieren Treppenfunktionen φ_k mit $\|f - \varphi_k\|_1 \xrightarrow{k\to\infty} 0$.
Dann folgt $\|g - \varphi_k\|_1 \leq \|g - f\|_1 + \|f - \varphi_k\|_1 = \|f - \varphi_k\|_1 \to 0$, also $\|g - \varphi_k\|_1 \to 0$. Somit ist g integrierbar und $\int g\,dx = \lim\limits_{k\to\infty} \int \varphi_k\,dx = \int f\,dx$. □

Dies motiviert die folgende Sprechweise.

1.4.8 Definition

Eine Ausssage $A(x)$ über Punkte $x \in \mathbb{R}^n$ gilt **fast überall (f. ü.)**, falls die Menge $\{x : A(x)$ gilt nicht$\}$ eine Nullmenge ist. Analog spricht man von **fast nie** und **für fast alle** x.

Beispiele:

(1) $f = 0$ fast überall $\Leftrightarrow$ $\int |f| = 0$.

(2) Ist $f = g$ fast überall, f integrierbar, so ist g integrierbar und $\int f = \int g$.

Bemerkung: In der Integrationstheorie kommt es meist nur auf die Werte einer Funktion fast überall an. Selbst wenn eine Funktion nur *fast überall definiert* ist, ist es sinnvoll, davon zu sprechen, ob sie integrierbar ist (und was ihr Integral ist). Man setzt einfach die Werte auf der fehlenden Nullmenge beliebig fest (z. B. gleich Null). Nach Satz 1.4.7 kommt es nicht darauf an, *wie* man diese Werte wählt.

Die Idee, zwei Objekte als »im Wesentlichen gleich« anzusehen, setzt man mathematisch mit dem Begriff der Äquivalenzrelation um. In unserem Kontext setzen wir also

$$f \equiv g :\Leftrightarrow f(x) = g(x) \quad \text{fast überall.}$$

Man zeigt leicht, dass $\equiv$ eine Äquivalenzrelation ist (Übung).

1.4.9 Definition

Sei $\boldsymbol{L^1(\mathbb{R}^n)} := \mathcal{L}^1(\mathbb{R}^n)/\equiv$, also die Menge der Äquivalenzklassen. Die Elemente von $L^1(\mathbb{R}^n)$ werden mit $[f] = \{g : g \equiv f\}$ bezeichnet.

Mit anderen Worten, $L^1(\mathbb{R}^n)$ ist der Quotientenraum $\mathcal{L}^1(\mathbb{R}^n)/\mathcal{N}$, $\mathcal{N} = \{$ Nullfunktionen auf $\mathbb{R}^n\}$, im Sinne der linearen Algebra. Also ist $L^1(\mathbb{R}^n)$ ein Vektorraum (mit $[f] + [g] = [f + g]$ etc.).

Aus Satz 1.4.7 folgt sofort, dass $\int : L^1(\mathbb{R}^n) \to \mathbb{R}$ wohldefiniert ist, und aus Lemma 1.4.5, dass $\|\cdot\|_1$ eine Norm (nicht bloß eine Halbnorm) auf $L^1(\mathbb{R}^n)$ definiert, die sogenannte $\boldsymbol{L^1}$**-Norm**.

1.5 Konvergenzsätze

Dies ist der Höhepunkt der Lebesgue-Theorie.
Frage: Gilt für Funktionenfolgen (f_m)

$$\int \lim_{m\to\infty} f_m(x)\,dx = \lim_{m\to\infty} \int f_m(x)\,dx \;? \qquad (*)$$

Etwas präzisere Frage: Angenommen, alle f_m sind integrierbar, und der Grenzwert $\lim_{m\to\infty} f_m(x)$ existiert für jedes x. Existiert dann notwendigerweise der Grenzwert rechts und gilt die Gleichheit?
Beachte: Wir betrachten hier die *punktweise Konvergenz* der Funktionenfolge.
Antwort: Nein.

Beispiel: Sei $f_m(x) = m\,\chi_{(0,1/m)}$. Dann ist $\lim_{m\to\infty} f_m(x) = 0$ für jedes x, also ist die linke Seite von $(*)$ gleich 0. Aber $\int f_m(x)\,dx = 1$ für jedes m, also ist die rechte Seite gleich 1.

Wir brauchen also eine zusätzliche Bedingung an die Folge (f_k). Wir werden zwei solche Bedingungen kennenlernen. Zunächst ein etwas speziellerer Fall.

1.5.1 Satz (Satz von der monotonen, integralbeschränkten Konvergenz, Satz von Beppo Levi)
Seien $f_1, f_2, \ldots : \mathbb{R}^n \to \mathbb{R}$ integrierbar und $f_1 \le f_2 \le \ldots$. Angenommen, es gibt ein $L \in \mathbb{R}$ mit $\int f_m\,dx \le L$ für alle m. Dann ist $f := \lim_{m\to\infty} f_m$ integrierbar und $\int f\,dx = \lim_{m\to\infty} \int f_m\,dx$.
Eine analoge Aussage gilt für monoton fallende Funktionenfolgen.

Bemerkung: $f_1 \le f_2 \le \ldots$ bedeutet $f_1(x) \le f_2(x) \le \ldots$ für jedes x. Daher existiert $f(x) := \lim_{m\to\infty} f_m(x) \in \overline{\mathbb{R}}$. Dies ist gleich ∞, falls $(f_m(x))_m$ unbeschränkt ist, sonst existiert der endliche Grenzwert.

Beweis: Idee: Konvergenz von $(\int f_m)$ kann mittels der L^1-Halbnorm formuliert werden. Diese verhält sich gut bzgl. unendlicher Reihen (Lemma 1.2.4(e)). Daher sollte man die Konvergenz der Folge auf die Konvergenz einer Reihe zurückführen.
Setze $g_m = f_{m+1} - f_m$, $m \in \mathbb{N}$. Wir haben nun:

- ▷ $\sum_{m=1}^{\infty} \|g_m\|_1$ konvergiert. Denn wegen $g_m \ge 0$ ist $\|g_m\|_1 = \int g_m$, also
  $\sum_{m=1}^{M} \|g_m\|_1 = \sum_{m=1}^{M} \int g_m = \sum_{m=1}^{M} (\int f_{m+1} - \int f_m) = \int f_{M+1} - \int f_1 \le L - \int f_1$
  (Linearität des Integrals und Teleskopsumme). Eine Reihe mit nicht-negativen Gliedern, deren Partialsummen beschränkt sind, konvergiert.
- ▷ $f - f_N = \sum_{m=N}^{\infty} g_m$. Denn $\sum_{m=N}^{M} g_m = f_{M+1} - f_N \xrightarrow{M\to\infty} f - f_N$.
- ▷ $\|f - f_N\|_1 \xrightarrow{N\to\infty} 0$. Denn $\|f - f_N\|_1 \le \sum_{m=N}^{\infty} \|g_m\|_1$, und für eine konvergente Reihe $s = \sum_{m=1}^{\infty} a_m$ mit Partialsummen $s_N = \sum_{m=1}^{N} a_m$ ist $\sum_{m=N}^{\infty} a_m = s - s_{N-1} \xrightarrow{N\to\infty} 0$.

Da $\mathcal{L}^1$ abgeschlossen und $\int$ stetig bzgl. $\|\cdot\|_1$ ist (Satz 1.3.3), folgt daraus die Behauptung. □

Eine erste Anwendung ist folgende:

1.5.2 Satz (Integration durch Ausschöpfung)
Seien $A \subset \mathbb{R}^n$, $A_1, A_2, \ldots \subset \mathbb{R}^n$ mit $A_1 \subset A_2 \subset A_3 \subset \ldots$, $A = \bigcup_{i=1}^{\infty} A_i$. Sei $f : A \to \overline{\mathbb{R}}$ über jedes A_m integrierbar. Falls die Folge $\left(\int_{A_m} |f|\right)_{m\in\mathbb{N}}$ beschränkt ist, so ist f auch über A integrierbar und es gilt

$$\int_A f = \lim_{m\to\infty} \int_{A_m} f\,.$$

Beweis: Wir nehmen zunächst $f \geq 0$ an. Aus den Bedingungen an die A_m folgt, dass $\chi_{A_m}(x) \xrightarrow{m\to\infty} \chi_A(x)$ für alle x, und $\chi_{A_1} \leq \chi_{A_2} \leq \ldots$.
Setze $f_m = f \cdot \chi_{A_m}$. Dann folgt $f_m \to f$ und wegen $f \geq 0$ ist $f_1 \leq f_2 \leq \ldots$. Außerdem ist $\left(\int f_m\right)_m$ beschränkt nach Vorraussetzung. Nach Satz 1.5.1 ist f ist integrierbar und $\int f_m \to \int_A f$.
Sei nun f beliebig. Setze $f_+ = \max\{f, 0\}$ und $f_- = \min\{f, 0\}$. Nach Lemma 1.3.4(b) sind f_+, f_- über A_m integrierbar. Es gilt $f = f_+ + f_-$ und $|f| = f_+ - f_-$ und damit $0 \leq f_+ \leq |f|$ und $0 \leq -f_- \leq |f|$. Daher sind die Folgen $\left(\int_{A_m} f_\pm\right)_m$ beschränkt. Wende nun den ersten Schritt auf f_+ und $-f_-$ an, dann folgt $\int_{A_m} f_+ \to \int_A f_+$ und $\int_{A_m}(-f_-) \to \int_A (-f_-)$ und durch Subtraktion $\int_{A_m} f \to \int_A f$. □

Bemerkung: Es gilt auch umgekehrt: Ist f über A integrierbar, dann ist die Folge $\left(\int_{A_m} |f|\right)_m$ beschränkt. Denn mit f ist auch $|f|$ über A integrierbar, und $\int_{A_m} |f| \leq \int_A |f|$.

Beispiele:

(1) $\int e^{-|x|}\,dx$. Hier ist $A = \mathbb{R}$. Sei $A_m = [-m, m]$. Dann ist nach Satz 1.3.6

$$\int_{A_m} e^{-|x|}\,dx = \int_{-m}^{m} e^{-|x|}\,dx = 2(1 - e^{-m}) \to 2\,.$$

Also $e^{-|x|} \in \mathcal{L}^1(\mathbb{R})$ und $\int e^{-|x|}\,dx = 2$.

(2) $\int_0^1 \frac{1}{x}\,dx$. Hier ist $A = [0, 1]$. (Dass $1/x$ bei $x = 0$ nicht definiert ist, stört uns ja nicht!)
Setze $A_m = \left[\frac{1}{m}, 1\right]$. Dann ist: $\int_{A_m} \frac{1}{x}\,dx = \int_{\frac{1}{m}}^1 \frac{1}{x}\,dx = (\log|x|)\big|_{1/m}^1 = \log(1) - \log\left(\frac{1}{m}\right) = \log(m) \to \infty$.
Somit ist $\frac{1}{x}$ nicht über $[0, 1]$ integrierbar. Ähnlich prüft man nach:

$$\int_0^1 \frac{1}{x^a}\,dx \text{ existiert genau dann, wenn } a < 1.$$

Wir kommen nun zum wichtigsten Konvergenzsatz.

1.5.3 Satz (Satz von der majorisierten Konvergenz, Satz von Lebesgue)
Seien $f_1, f_2, \ldots : \mathbb{R}^n \to \overline{\mathbb{R}}$. Angenommen, $f(x) := \lim_{m\to\infty} f_m(x)$ existiert für fast alle x, und es gibt eine integrierbare Funktion F, so dass gilt:

$$|f_m(x)| \leq F(x)$$

für alle m, x. Dann ist f integrierbar und $\int f = \lim_{m\to\infty} \int f_m$.

F heißt integrierbare **Majorante** für die Folge $(f_m)_{m\in\mathbb{N}}$.

Beispiele:

(1) Ist Q ein Quader und sind die $f_m : Q \to \mathbb{R}$ gleichmäßig beschränkt (das heißt, es existiert ein L mit $|f_m(x)| \leq L$ für alle $m \in \mathbb{N}, x \in Q$), dann folgt

$$\int_Q \lim_{m\to\infty} f_m(x)\,dx = \lim_{m\to\infty} \int_Q f_m(x)\,dx\,,$$

falls der punktweise Grenzwert links existiert. Denn $F = L\chi_Q$ ist eine integrierbare Majorante für $f_m\chi_Q$.

(2) Die Folge $f_m = m\chi_{(0,1/m)}$ vom Anfang dieses Kapitels kann nach dem Satz keine integrierbare Majorante besitzen. Überzeugen Sie sich davon direkt! Im Wesentlichen läuft das auf die Tatsache $\frac{1}{x} \notin \mathcal{L}^1((0,1))$ hinaus.

Beweis: Idee: Führe das auf den monotonen Fall zurück. Dies geht mit einem ähnlichen Trick wie bei unserem Beweis des Satzes von Bolzano-Weierstraß:
Setze $g_m = \sup\{f_m, f_{m+1}, \ldots\}$. Dann ist $g_1 \geq g_2 \geq \ldots$ und $g_m \xrightarrow{m\to\infty} f$.
Warum ist g_m integrierbar? Es gilt $g_m = \lim\limits_{l\to\infty} h_{m,l}$ mit $h_{m,l} = \max\{f_m, \ldots, f_{m+l}\}$. Da die f_m integrierbar sind, sind alle $h_{m,l}$ integrierbar (wiederholte Anwendung von Lemma 1.3.4b). Weiterhin ist $h_{m,1} \leq h_{m,2} \leq \ldots$, und die Folge $(\int h_{m,l})_l$ ist beschränkt, da $|h_{m,l}| \leq \max\{|f_m|, \ldots, |f_{m+l}|\} \leq F$, also $|\int h_{m,l}| \leq \int |h_{m,l}| \leq \int F$ gilt.
Nach dem Satz über die monotone Konvergenz ist also g_m integrierbar für jedes m, und $\int g_m = \lim\limits_{l\to\infty} \int h_{m,l}$.
Wegen $|\int h_{m,l}| \leq \int F$ ist die Folge $(\int g_m)_m$ beschränkt. Wenden wir den Satz über die monotone Konvergenz auf die monoton fallende Folge $(g_m)_m$ an, folgt

$$f \text{ ist integrierbar und } \int f = \lim_{m\to\infty} \int g_m .$$

Wir hätten aber gerne rechts $\int f_m$ stehen, nicht $\int g_m$. Wie hängen diese zusammen? Offenbar ist $g_m \geq f_m$, aber das reicht natürlich nicht für unsere Zwecke.

Lösung des Dilemmas: Wir wiederholen das ganze Argument »von der anderen Richtung«, also mit $\tilde{g}_m = \inf\{f_m, f_{m+1}, \ldots\}$. Wir erhalten $\int f = \lim\limits_{m\to\infty} \int \tilde{g}_m$ und $\tilde{g}_m \leq f_m$.
Schließlich haben wir $\tilde{g}_m \leq f_m \leq g_m$, also $\int \tilde{g}_m \leq \int f_m \leq \int g_m$, und da die linke und rechte Seite gegen $\int f$ konvergieren, tut es auch die Mitte. □

Als Anwendung zeigen wir, dass sehr viele der praktisch vorkommenden Funktionen integrierbar sind.

1.5.4 Satz

Sei $Q \subset \mathbb{R}^n$ ein Quader und $f : Q \to \mathbb{R}$ beschränkt und fast überall stetig. Dann ist f über Q integrierbar.

Später werden wir sehen, dass der Satz auch für viel allgemeinere Mengen Q gilt (siehe Satz 1.6.9).

Beweis: Für jedes $m \in \mathbb{R}$ zerlege Q in n^m Teilquader, $Q = \bigcup\limits_{j=1}^{n^m} Q_{m,j}$, durch Unterteilung jeder Kante in m gleiche Teile. Setze nun

$$f_m(x) = \begin{cases} \inf\limits_{y\in Q_{m,j}} f(y) & \text{falls } x \text{ im Innern eines } Q_{m,j} \text{ liegt} \\ 0 & \text{sonst.} \end{cases}$$

f_m ist eine Treppenfunktion, und die Folge $(f_m)_m$ ist gleichmäßig beschränkt. Wir zeigen, dass $f_m \to f$ fast überall.
Die Menge der Punkte, die auf dem Rand eines der $Q_{m,j}$ liegen (m, j beliebig), ist eine Nullmenge, ebenso die Menge der Unstetigkeitsstellen von x. Also genügt es zu zeigen, dass $f_m(x) \xrightarrow{m\to\infty} f(x)$ für alle anderen $x \in Q$ gilt.

Sei $\varepsilon > 0$. Da f in x stetig ist, gibt es ein $\delta > 0$ mit $|f(y) - f(x)| < \varepsilon$ für $y \in K_\delta(x)$, wobei $K_\delta(x)$ die Kugel bzgl. der Maximumsnorm auf $\mathbb{R}^n$ ist. Für m genügend groß (größer als δ^{-1}-mal die größte Kantenlänge von Q) enthält $K_\delta(x)$ den Quader $Q_{m,j}$, der x enthält, also folgt $|f_m(x) - f(x)| < \varepsilon$ für diese m. Also folgt $f_m(x) \xrightarrow{m\to\infty} f(x)$. □

Vertauschen von Differentiation und Integration

Mit Hilfe des Satzes über die majorisierte Konvergenz lässt sich ein einfacher Beweis des folgenden wichtigen Satzes geben.

1.5.5 Satz (Vertauschen von Ableitung und Integral)

Sei $I \subset \mathbb{R}$ ein Intervall, $U \subset \mathbb{R}^n$ und $F : I \times U \to \mathbb{R}$ eine Funktion, die folgende Bedingungen erfüllt:

- $t \mapsto F(t,x)$ ist differenzierbar für jedes $x \in U$
- $F(t,\cdot)$ ist für jedes t integrierbar über U
- Es gibt eine integrierbare Funktion M auf U (sogenannte **Majorante**) mit

$$\left|\frac{\partial F}{\partial t}(t,x)\right| \leq M(x) \quad \forall t \in I, x \in U$$

Dann gilt für jedes $t \in I$

$$\frac{d}{dt}\int\limits_U F(t,x)\,dx = \int\limits_U \frac{\partial}{\partial t}F(t,x)\,dx$$

Genauer: $t \mapsto \int\limits_U F(t,x)\,dx$ ist differenzierbar und $x \mapsto \frac{\partial}{\partial t}F(t,x)$ ist integrierbar für jedes t, und die Formel gilt.

Meist hat man folgende einfachere Situation (manchmal aber auch nicht, siehe die Integralformel für Lösungen der Laplace-Gleichung):

1.5.6 Korollar

Sei $I \subset \mathbb{R}$ ein Intervall, $U \subset \mathbb{R}^n$ offen und $F : I \times U \to \mathbb{R}$ eine C^1-Funktion. Sei $K \subset U$ kompakt. Dann gilt für jedes $t \in I$

$$\frac{d}{dt}\int\limits_K F(t,x)\,dx = \int\limits_K \frac{\partial}{\partial t}F(t,x)\,dx$$

Beweis (von Satz 1.5.5): Sei

$$f_h(t,x) = \frac{F(t+h,x) - F(t,x)}{h}$$

für $t \in I, x \in U$ und $h \neq 0$ genügend klein. Dann ist $\lim\limits_{h\to 0} f_h = \partial F/\partial t$ punktweise und mit $I(t) = \int\limits_U F(t,x)\,dx$ ist $\frac{I(t+h)-I(t)}{h} = \int\limits_U f_h(t,x)\,dx$. Es ist also nur zu zeigen, dass $\lim\limits_{h\to 0}\int\limits_U f_h(t,x)\,dx = \int\limits_U \lim\limits_{h\to 0} f_h(t,x)\,dx$ gilt. Dies folgt aus dem Satz über die majorisierte Konvergenz, falls es eine integrierbare Majorante für die Funktionen f_h gibt. Nach dem Mittelwertsatz ist $f_h(t,x) = \frac{\partial F}{\partial t}(\tau,x)$ für ein $\tau \in (t, t+h)$, und dies ist im Betrag durch $M(x)$ majorisiert. M ist also die gesuchte Majorante. □

Beweis (des Korollars): Sei $t \in I$ und I' eine kompakte Umgebung von t in I. Da $\partial F/\partial t$ stetig ist, ist es auf der kompakten Menge $I' \times K$ im Betrag durch eine Konstante C beschränkt, d.h. $M(x) = C$ ist eine Majorante wie im Satz gefordert, auf $I' \times K$. Da K kompakt ist, ist M über K integrierbar. Man kann also den Satz (mit K statt U und I' statt I) anwenden. □

Regel- und Lebesgue-Integral, II

Wir sahen bereits, dass eine Regelfunktion auf einem abgeschlossenen Intervall Lebesgue-integrierbar ist. Wie steht es mit uneigentlich integrierbaren Regelfunktionen?

1.5.7 Satz

Sei $a < b$ in $\mathbb{R} \cup \{\pm\infty\}$, $f : (a,b) \to \mathbb{R}$ Regelfunktion.
f ist auf (a,b) Lebesgue-integrierbar $\Leftrightarrow$ Das uneigentliche Integral von $|f|$ über (a,b) existiert.
In diesem Fall ist

$$\int_{(a,b)} f(x)\,dx = \int_a^b f(x)\,dx\,.$$

Zur Erinnerung: Das uneigentliche Integral von f über (a,b) ist so definiert:
Falls für beliebige Folgen $a_1 \geq a_2 \geq \ldots \to a$ und $b_1 \leq b_2 \leq \ldots \to b$ reeller Zahlen der Grenzwert $\lim\limits_{k\to\infty} \int_{a_k}^{b_k} f(x)\,dx$ existiert, so heißt f über (a,b) uneigentlich integrierbar. Der Grenzwert hängt dann nicht von der Wahl dieser Folgen ab und wird mit $\int_a^b f(x)\,dx$ bezeichnet.

Die Bedingungen an die Folgen (a_k), (b_k) bedeuten gerade, dass $I_k = [a_k, b_k]$ eine Ausschöpfung von $I = (a,b)$ ist, d. h., dass $I_1 \subset I_2 \subset \cdots$, $\bigcup_{j=1}^{\infty} I_j = I$.

Außerdem sei daran erinnert, dass aus der Existenz des uneigentlichen Integrals von $|f|$ die Existenz des uneigentlichen Integrals von f folgt (analog zu Reihen: konvergiert eine Reihe absolut, so konvergiert sie).

Beweis: Sei $I = (a,b)$, und I_k bezeichne Intervalle wie unter ›Erinnerung‹ oben.
›⇒‹: Ist $f \in \mathcal{L}^1(I)$, so konvergiert nach Satz 1.5.2 $\int_{I_k} f \to \int_I f$ für eine beliebige Ausschöpfung (I_k) von I. Nach Voraussetzung ist $f_{|I_k}$ Regelfunktion und nach Satz 1.3.6 ist $\int_{I_k} f = \int_{a_k}^{b_k} f$, da I_k abgeschlossen ist. Also ist f uneigentlich über I integrierbar und es gilt $\int_I f = \int_a^b f$.

Weiterhin gilt nach Lemma 1.3.4 $f \in \mathcal{L}^1(I) \Rightarrow |f| \in \mathcal{L}^1(I)$, also ist nach demselben Argument auch $|f|$ uneigentlich über I integrierbar.
›⇐‹: Da das uneigentliche Integral von $|f|$ und damit von f über I existiert, existiert der Grenzwert der Folge $\int_{I_k} f = \int_{a_k}^{b_k} f$, also ist diese Folge beschränkt, also ist nach Satz 1.5.2 $f \in \mathcal{L}^1(I)$. □

Beispiele: Das wichtigeste Beispiel haben wir bereits kennengelernt: $f(x) = x^{-a}$ ist über $(0,1)$ genau dann uneigentlich integrierbar (und damit Lebesgue-integrierbar), wenn $a < 1$.

Ein Beispiel, wo die Betragsstriche im Satz eine Rolle spielen, ist folgendes:
Aus Analysis I wissen wir: $\int_{-\infty}^{\infty} \frac{\sin(x)}{x}\,dx$ existiert als uneigentliches Integral, aber $\int_{-\infty}^{\infty} |\frac{\sin(x)}{x}|\,dx$ nicht. Also ist $\frac{\sin(x)}{x}$ nicht Lebesgue-integrierbar.

Regel- und Lebesgue-Integral, II

Wir sahen bereits, dass eine Regelfunktion auf einem abgeschlossenen Intervall Lebesgue-integrierbar ist. Wie steht es mit uneigentlich integrierbaren Regelfunktionen?

1.5.8 Satz

Sei $a < b$ in $\mathbb{R} \cup \{\pm\infty\}$, $f : (a,b) \to \mathbb{R}$ Regelfunktion.
f ist auf (a,b) Lebesgue-integrierbar $\Leftrightarrow$ Das uneigentliche Integral von $|f|$ über (a,b) existiert.
In diesem Fall ist

$$\int_{(a,b)} f(x)\,dx = \int_a^b f(x)\,dx\,.$$

Zur Erinnerung: Das uneigentliche Integral von f über (a,b) ist so definiert:
Falls für beliebige Folgen $a_1 \geq a_2 \geq \ldots \to a$ und $b_1 \leq b_2 \leq \ldots \to b$ reeller Zahlen der Grenzwert $\lim\limits_{k\to\infty} \int_{a_k}^{b_k} f(x)\,dx$ existiert, so heißt f über (a,b) uneigentlich integrierbar. Der Grenzwert hängt dann nicht von der Wahl dieser Folgen ab und wird mit $\int_a^b f(x)\,dx$ bezeichnet.

Die Bedingungen an die Folgen (a_k), (b_k) bedeuten gerade, dass $I_k = [a_k, b_k]$ eine Ausschöpfung von $I = (a,b)$ ist, d. h., dass $I_1 \subset I_2 \subset \cdots$, $\bigcup_{j=1}^{\infty} I_j = I$.

Außerdem sei daran erinnert, dass aus der Existenz des uneigentlichen Integrals von $|f|$ die Existenz des uneigentlichen Integrals von f folgt (analog zu Reihen: konvergiert eine Reihe absolut, so konvergiert sie).

Beweis: Sei $I = (a,b)$, und I_k bezeichne Intervalle wie unter ›Erinnerung‹ oben.
›⇒‹: Ist $f \in \mathcal{L}^1(I)$, so konvergiert nach Satz 1.5.2 $\int_{I_k} f \to \int_I f$ für eine beliebige Ausschöpfung (I_k) von I. Nach Voraussetzung ist $f_{|I_k}$ Regelfunktion und nach Satz 1.3.6 ist $\int_{I_k} f = \int_{a_k}^{b_k} f$, da I_k abgeschlossen ist. Also ist f uneigentlich über I integrierbar und es gilt $\int_I f = \int_a^b f$.

Weiterhin gilt nach Lemma 1.3.4 $f \in \mathcal{L}^1(I) \Rightarrow |f| \in \mathcal{L}^1(I)$, also ist nach demselben Argument auch $|f|$ uneigentlich über I integrierbar.
›⇐‹: Da das uneigentliche Integral von $|f|$ und damit von f über I existiert, existiert der Grenzwert der Folge $\int_{I_k} f = \int_{a_k}^{b_k} f$, also ist diese Folge beschränkt, also ist nach Satz 1.5.2 $f \in \mathcal{L}^1(I)$. □

Beispiele: Das wichtigeste Beispiel haben wir bereits kennengelernt: $f(x) = x^{-a}$ ist über $(0,1)$ genau dann uneigentlich integrierbar (und damit Lebesgue-integrierbar), wenn $a < 1$.

Ein Beispiel, wo die Betragsstriche im Satz eine Rolle spielen, ist folgendes:
Aus Analysis I wissen wir: $\int_{-\infty}^{\infty} \frac{\sin(x)}{x}\,dx$ existiert als uneigentliches Integral, aber $\int_{-\infty}^{\infty} |\frac{\sin(x)}{x}|\,dx$ nicht. Also ist $\frac{\sin(x)}{x}$ nicht Lebesgue-integrierbar.

1.6 Das Maß von Mengen

Wie schon anfangs angekündigt, kann man mit Hilfe von Integralen leicht einen Volumenbegriff für Mengen $A \subset \mathbb{R}^n$ definieren. Die Idee ist, $\mathrm{vol}(A) = \int \chi_A\,dx$ zu setzen. da wir aber auch Mengen mit unendlichem Volumen betrachten wollen (z. B. den ganzen $\mathbb{R}^n$), definieren wir:

1.6.1 Definition

Sei $Q_R = [-R,R]^n$ für $R > 0$.

(1) $A \subset \mathbb{R}^n$ heißt **(Lebesgue-) messbar**, falls $\chi_{A\cap Q_R} \in \mathcal{L}^1(\mathbb{R}^n)$ für alle $R > 0$.

(2) Ist A messbar, so heißt $\mathrm{vol}(A) := \lim\limits_{R\to\infty} \int \chi_{A\cap Q_R}\,dx \in [0,\infty]$ das **Volumen** (oder **Lebesgue-Maß**) von A.

Manchmal schreibt man auch vol_n statt vol. Für $n = 1$ spricht man meist von Länge und für $n = 2$ von Fläche.

Bemerkung: Der Grenzwert in (2) existiert, da $\int \chi_{A\cap Q_R}$ bezüglich R monoton wächst. Man sieht leicht, dass er gleich $\|\chi_A\|_1$ ist.

Falls A beschränkt ist, dann gilt: A messbar $\Leftrightarrow$ $\chi_A \in \mathcal{L}^1$, und $\mathrm{vol}(A) = \int \chi_A\,dx$.
Allgemeiner gilt: $\mathrm{vol}(A) < \infty \Leftrightarrow \chi_A \in \mathcal{L}^1(\mathbb{R}^n)$. Dies folgt aus dem folgenden Lemma.

Folgende Beobachtung ist im Folgenden wesentlich:

1.6.102 Lemma

Seien $B_1 \subset B_2 \subset \cdots$ Teilmengen von $\mathbb{R}^n$ und $A = \bigcup_{k=1}^{\infty} B_k$, dann konvergiert die Folge (χ_{B_k}) punktweise gegen χ_A und ist von χ_A dominiert.

Ist zusätzlich die Folge $(\mathrm{vol}(B_k))$ beschränkt, so ist $\chi_A \in \mathcal{L}^1$ und $\mathrm{vol}(B_k) \xrightarrow{k\to\infty} \mathrm{vol}\, A$.

Beweis: Der erste Teil ist klar. Der zweite folgt aus dem Satz über die monotone, integralbeschränkte Konvergenz. □

Wir stellen nun einige grundlegende Eigenschaften messbarer Mengen und der Abbildung vol zusammen.

1.6.2 Satz

Die Menge $\mathcal{M}$ der messbaren Teilmengen von $\mathbb{R}^n$ ist eine **σ-Algebra** auf $\mathbb{R}^n$, d. h.

(a) $\mathbb{R}^n \in \mathcal{M}$

(b) $A \in \mathcal{M} \Rightarrow A^c := \mathbb{R}^n \setminus A \in \mathcal{M}$

(c) $A_1, A_2, \cdots \in \mathcal{M} \Rightarrow \bigcup_{k=1}^{\infty} A_k \in \mathcal{M}$

Beweis:

(a) $\chi_{\mathbb{R}^n \cap Q} = \chi_Q \in \mathcal{L}^1$, weil χ_Q Treppenfunktion.

(b) Sei $R > 0$. Da $A^c \cap Q_R$, $A \cap Q_R$ eine Zerlegung von Q_R bilden, gilt $\chi_{A^c \cap Q_R} = \chi_{Q_R} - \chi_{A \cap Q_R} \in \mathcal{L}^1$.

(c) Sei $A = \bigcup_{k=1}^{\infty} A_k$, dann ist $A \cap Q_R = \bigcup_{k=1}^{\infty} (A_k \cap Q_R)$. Wendet man Lemma 1.6.102 auf die Mengen $B_k = \bigcup_{i=1}^{k} (A_i \cap Q_R)$ an, so folgt $\chi_{A \cap Q_R} \in \mathcal{L}^1$; denn alle $B_k \subset Q_R$, also ist die Folge $(\mathrm{vol}(B_k))$ beschränkt. □

1.6.3 Satz

$\mathrm{vol} : \mathcal{M} \to [0, \infty]$ ist ein **Maß**, d. h.

(a) $\mathrm{vol}(\emptyset) = 0$

(b) $A_1, A_2, \cdots \in \mathcal{M}$ paarweise disjunkt $\Rightarrow \mathrm{vol}\Big(\bigcup_{k=1}^{\infty} A_k\Big) = \sum_{k=1}^{\infty} \mathrm{vol}(A_k)$ (**σ-Additivität**)

Beweis:

(a) $\int 0 = 0$.

(b) Sei $A = \bigcup_{i=1}^{\infty} A_i$ und $B_k = \bigcup_{i=1}^{k} A_i$. Da die A_i paarweise disjunkt sind, ist $\chi_{B_k} = \sum_{i=1}^{k} \chi_{A_i}$, also nach Integration $\mathrm{vol}(B_k) = \sum_{i=1}^{k} \mathrm{vol}(A_i)$. Falls die Folge $(\mathrm{vol}(B_k))$ beschränkt ist, folgt die Behauptung nun aus Lemma 1.6.102, sonst folgt offenbar $\mathrm{vol}(A) = \infty$ und damit ebenfalls die Behauptung. □

Bemerkung: Aus diesem Satz lassen sich nun einige »offensichtliche« Dinge leicht ablesen, z. B.

$$A \subset B \text{ messbar} \implies \mathrm{vol}(A) \le \mathrm{vol}(B)$$

(Beweis: Wende (b) auf $A_1 = A$, $A_2 = B \setminus A$, $A_3 = A_4 = \ldots = \emptyset$ an.)

Im folgenden Satz verwenden wir für $A \subset \mathbb{R}^n$, $y \in \mathbb{R}^n$ die Notation

$$A + y := \{x + y : x \in A\}$$

(Verschiebung von A um y).

1.6.4 Satz

(a) $\mathcal{M}$ und vol sind **translationsinvariant**, d. h.: Ist $A \in \mathcal{M}$ und $y \in \mathbb{R}^n$, so ist auch $A + y \in \mathcal{M}$ und $\mathrm{vol}(A + y) = \mathrm{vol}(A)$.

(b) vol ist **normiert**, d. h. für alle Quader Q ist $\mathrm{vol}(Q) =$ die alte Definition von $\mathrm{vol}(Q)$.

Beweis: (a) Beweisskizze für A beschränkt:
Es gilt $\chi_{A+y}(x) = \chi_A(x - y)$, denn $x \in A + y \Leftrightarrow x - y \in A$. Also folgt die Behauptung aus der **Translationsinvarianz des Lebesgue-Integrals**:

$$f \in \mathcal{L}^1(\mathbb{R}^n),\ y \in \mathbb{R}^n \ \Rightarrow\ f(\cdot - y) \in \mathcal{L}^1(\mathbb{R}^n) \text{ und } \int f(x - y)\, dx = \int f(x)\, dx\,.$$

Beweis hiervon: Dies gilt zunächst für charakteristische Funktionen von Quadern, da das Volumen von Quadern mittels der Seitenlängen, also translationsinvariant definiert ist, dann wegen Linearität für Treppenfunktionen, und dann mittels Approximation für beliebige $f \in \mathcal{L}^1$.
(b) Klar. □

Bemerkung: vol ist auch rotationsinvariant. Dies werden wir später beweisen. Genau wie die Translationsinvarianz ist dies Spezialfall einer allgemeinen Substitutionsregel (der Transformationsformel).

Existenz nicht messbarer Mengen

Warum brauchen wir den Begriff ›messbare Menge‹? Hat nicht jede (zumindest jede beschränkte) Menge ein Volumen? Man kann sie doch einfach in einen Eimer Wasser werfen und beobachten, wie stark der Wasserspiegel steigt!

Die Antwort ist leider: Nein! Der mathematische Mengenbegriff ist so allgemein, dass er auch sehr »nichtphysikalische« Objekte zulässt und daher diese Intuition nicht zulässig ist. Dies wollen wir nun zeigen.

Da jeder »vernünftige« Volumenbegriff wohl translationsinvariant und normiert sein sollte, präzisieren wir die Frage wie folgt:
Maßproblem: Gibt es eine Abbildung $\mu : \mathcal{P}(\mathbb{R}^n) \to [0, \infty]$, die ein translationsinvariantes und normiertes Maß ist?

1.6.5 Satz

Das Maßproblem ist unlösbar für alle $n \geq 1$.

Beweis: Für $n = 1$: Auf $\mathbb{R}$ definiere eine Äquivalenzrelation durch: $x \sim y :\Leftrightarrow x - y \in \mathbb{Q}$. Offenbar schneidet jede Äquivalenzklasse das Intervall $[0, 1]$. Man bilde eine Menge Y, indem man aus jeder Äquivalenzklasse ein Element $y \in [0, 1]$ nimmt.

Wir zeigen, dass die Menge Y zu Problemen führt.
Per Definition gilt:

(a) $x, y \in Y, x \neq y \Rightarrow x - y \notin \mathbb{Q}$.

(b) $x \in \mathbb{R} \Rightarrow \exists y \in Y$ mit $x - y \in \mathbb{Q}$.

Seien $q_1, q_2, \cdots$ die rationalen Zahlen in $[-1,1]$, und setze $X = \bigcup_{i=1}^{\infty}(Y + q_i)$.
Behauptung: $[0,1] \subset X \subset [-1,2]$.

> Beweis: Ist $x \in [0,1]$, so wähle ein $y \in Y$ mit $x - y \in \mathbb{Q}$. Wegen $x, y \in [0,1]$ ist $x - y \in [-1,1]$, also $x - y = q_i$ für ein i und daher $x = y + q_i \in Y + q_i$, also $x \in X$. Dies beweist die linke Inklusion. Die rechte folgt sofort aus $Y \subset [0,1]$, $q_i \in [-1,1]\ \forall i$.

Wir nehmen nun an, das Maßproblem wäre lösbar durch ein Maß μ. Wegen der Normierung ist $\mu[0,1] = 1, \mu([-1,2]) = 3$, also $1 \leq \mu(X) \leq 3$ (vergleiche die Bemerkung nach Satz 1.6.3).

Andererseits ist die Vereinigung $X = \bigcup_{i=1}^{\infty}(Y + q_i)$ disjunkt, denn aus $z \in (Y + q_i) \cap (Y + q_j)$ folgt $z - q_i \in Y, z - q_j \in Y$, und da sich diese um die rationale Zahl $q_i - q_j$ unterscheiden, müssen sie gleich sein, also $q_i = q_j$.

Wegen der Translationsinvarianz ist $\mu(Y + q_i) = \mu(Y)$, also $\mu(X) = \sum_{i=1}^{\infty} \mu(Y + q_i) = \sum_{i=1}^{\infty} \mu(X)$, und daher muss

entweder $\mu(X) = 0$ sein (falls $\mu(Y) = 0$),
oder $\mu(X) = \infty$ (falls $\mu(Y) > 0$).

Dies ist ein Widerspruch zu $1 \leq \mu(X) \leq 3$.
Die Behauptung für $n > 1$ folgt leicht aus der für $n = 1$. □

Bemerkung: Der Beweis zeigt auch: Y ist eine nicht Lebesgue-messbare Menge. Und daher: χ_Y ist nicht Lebesgue-integrierbar.

Man könnte nun auf die Idee kommen, dass das Problem nur daher rührt, dass wir von μ verlangen, dass es additiv unter abzählbaren Summen ist (Teil (b) von Satz 1.6.3). Als Mindestanforderung an einen vernünftigen Volumenbegriff darf man (neben Translationsinvarianz und Normiertheit) aber wohl die **endliche Additivität** fordern:

$$A_1, \ldots, A_k \text{ paarweise diskunkt } \Rightarrow \text{ vol}(A_1 \cup \ldots \cup A_k) = \text{vol}(A_1) + \ldots + \text{vol}(A_k)$$

(Diese eingeschränkte Version des Maßproblems nennt man auch **Inhaltsproblem**.)

Doch auch hier gibt es Probleme (falls man auch noch die Rotationsinvarianz fordert), allerdings erst ab Dimension 3.

1.6.6 Satz (Das Banach-Tarski-Paradoxon)

Man kann eine Vollkugel im $\mathbb{R}^3$ so in sechs Teile zerlegen, dass sich diese Teile nach geeignetem Verschieben und Rotieren zu zwei neuen Vollkugeln *mit demselben Radius* zusammensetzen lassen.

Formal: Es gibt Mengen $A_1, \ldots, A_6, B_1, \ldots, B_6 \subset \mathbb{R}^3$ mit

- ▷ A_i paarweise disjunkt.
- ▷ $A_1 \cup \cdots \cup A_6 = K_1(0)$ (Einheitskugel).
- ▷ B_i geht aus A_i durch eine Bewegung (Translation und Rotation) hervor, für $i = 1 \ldots 6$.
- ▷ $B_1 \cup B_2 \cup B_3 = K_1(0)$, disjunkte Vereinigung.
- ▷ $B_4 \cup B_5 \cup B_6 = K_1(0)$, disjunkte Vereinigung.

Der Beweis wird an dieser Stelle nicht durchgeführt. Er besteht in einer Variation des Themas ›$\mathbb{N} = \{\text{gerade Zahlen}\} \cup \{\text{ungerade Zahlen}\}$, und trotzdem gibt es genau so viele gerade Zahlen/ungerade

Zahlen wie natürliche Zahlen‹. Hinzu kommt etwas Gruppentheorie (die Nicht-Auflösbarkeit der Gruppe der Rotationen im $\mathbb{R}^3$ ist hier wesentlich, daher funktioniert das Analogon nicht in $\mathbb{R}$ oder $\mathbb{R}^2$).

Intuitiv geht das natürlich nicht. Denn die Intuition verlangt, was wir von einem Maß verlangen. Mathematisch bedeutet dies, dass die $A_1, \dots A_6$ nicht alle messbar sein können, denn andernfalls würde

(1) $\mathrm{vol}(A_1) + \cdots + \mathrm{vol}(A_6) = \mathrm{vol}(K_1(0))$,

(2) $\mathrm{vol}(A_i) = \mathrm{vol}(B_i)$,

(3) $\mathrm{vol}(B_1) + \cdots + \mathrm{vol}(B_3) = \mathrm{vol}(K_1(0))$,

(4) $\mathrm{vol}(B_4) + \cdots + \mathrm{vol}(B_6) = \mathrm{vol}(K_1(0))$

gelten, und aus $(3)+(4)\text{-}(1)$ der Widerspruch $0 = \mathrm{vol}(K_1(0))$ folgen.

Es gibt jedoch auch gute Nachrichten: Alle »vernünftigen» Mengen sind messbar! »Vernünftig« heißt: Diejenigen, die man in endlich vielen Schritten mittels (abzählbaren) Vereinigungen, Schnitten und Komplementbildungen aus offenen und abgeschlossenen Mengen erhalten kann. Da Schnitte mittels Komplementbildung auf Vereinigungen zurückgeführt werden können, genügt es wegen der Maßeigenschaften (Satz 1.6.3) hierzu, folgendes nachzuprüfen:

1.6.7 Satz

Offene und abgeschlossene Mengen in $\mathbb{R}^n$ sind messbar.

Beweis:

▷ Sei $U \subset \mathbb{R}^n$ offen. Es gibt abzählbar viele rationale Quader im $\mathbb{R}^n$, d. h. solche deren Seitenlängen und Mittelpunktskoordinaten rational sind. Seien $Q_1, Q_2, \dots$ diejenigen rationalen Quader, die in U enthalten sind. Offenbar gilt $U = \bigcup_{i=1}^{\infty} Q_i$ (denn ist $p \in U$ beliebig, so existiert $\varepsilon > 0$ mit $K_\varepsilon(p) \subset U$; wählt man darin einen rationalen Punkt genügend nahe bei p, so passt auch noch ein rationaler Quader um diesen Mittelpunkt in $K_\varepsilon(p)$, der auch p enthält).

Alle Q_i sind messbar, also ist auch U messbar.

▷ Sei $A \in \mathbb{R}^n$ abgeschlossen, dann ist $\mathbb{R}^n \setminus A$ offen.

Wie oben gezeigt ist $\mathbb{R}^n \setminus A$ dann messbar, also ist auch A messbar. □

Messbare Funktionen

Wir wollen kurz den Begriff der messbaren Funktion einführen. Die Idee ist, dass eine Funktion messbar heißen soll, wenn sie »schön« ist, nicht allzu wild hin- und herspringt. Dies wird nicht vom Begriff der Integrierbarkeit abgedeckt, denn selbst eine stetige Funktion $f : \mathbb{R}^n \to \mathbb{R}$, die für $|x|$ nicht schnell genug gegen Null geht, ist nicht integrierbar; auch Funktionen, die in drastischer Weise unbeschränkt sind, sind nicht integrierbar, z. B. $f(x) = 1/x^2$ auf $\mathbb{R}$.

1.6.8 Definition (Messbare Funktion)

Eine Funktion $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ heißt **messbar**, wenn gilt:
Für alle Intervalle $I \subset \overline{\mathbb{R}}$ ist $f^{-1}(I)$ messbar.

Hier sind einige wichtige (meist leicht zu zeigende) Eigenschaften messbarer Funktionen, ohne Beweis:

- $f \in \mathcal{L}^1 \Rightarrow f$ ist messbar (aber nicht umgekehrt!).

  Zum Beispiel ist $f(x) = 1$ für alle x messbar, aber nicht integrierbar.

- Genauer gilt: f messbar $\Longleftrightarrow$ Für alle $R > 0$ ist die »abgeschnittene« Funktion

$$f_R(x) = \begin{cases} f(x) & \text{falls } x \in Q_R\,,\ |f(x)| \le R \\ 0 & \text{falls } x \notin Q_R \\ R & \text{falls } f(x) > R \\ -R & \text{falls } f(x) < -R \end{cases}$$

  integrierbar.

- Ist $F : \mathbb{R}^k \to \mathbb{R}$ stetig und sind $f_1, \ldots, f_k$ messbar, so ist $F(f_1, \ldots, f_k)$ messbar. Insbesondere sind mit f, g auch $f + g$, fg, $\max\{f, g\}$, $\min\{f, g\}$ und $|f|$ messbar.

- Eine fast überall stetige Funktion ist messbar.

- **Integrierbarkeitskriterium**: Ist f messbar und g integrierbar und $|f| \le g$, so ist f integrierbar.

  Dies ist nützlich, da Messbarkeit meist leicht nachzuprüfen ist; dann muss man für die Integrierbarkeit nur noch prüfen, ob die Funktion »nicht zu groß«, d. h. von einer integrierbaren Funktion dominiert ist.

- $A \subset \mathbb{R}^n$ messbar $\Leftrightarrow$ χ_A messbar.

- $f : \mathbb{R}^n \to \mathbb{R}$, $f \ge 0$ messbar $\Leftrightarrow$ Die Ordinatenmenge $\{(x, y) : 0 \le y \le f(x)\} \subset \mathbb{R}^{n+1}$ ist messbar.

Besonders soll folgendes hervorgehoben werden:

1.6.9 Satz

Ist $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ integrierbar und $A \subset \mathbb{R}^n$ messbar, so ist f über A integrierbar.

Beweis: f, χ_A messbar $\Rightarrow f\chi_A$ messbar. $|f\chi_A| \le |f| \Rightarrow f\chi_A$ integrierbar. □

Zusammen mit Satz 1.5.4 erhält man so einen reichhaltigen Vorrat an Funktionen von denen man weiß, dass sie integrierbar sind.

Bemerkung: Bei den Details über messbare Funktionen muss man genau aufpassen. Es ist zum Beispiel nicht richtig, dass für jede messbare Funktion f und jede messbare Menge A das Urbild $f^{-1}(A)$ messbar ist. Dies gilt nicht einmal für alle stetigen Funktionen (Beispiel: z. B. Kaballo, Analysis III, 14.10.c). Es stimmt aber, dass $f^{-1}(A)$ für alle *Borel*-Mengen messbar ist (die Klasse der **Borel-Mengen** ist die kleinste $\sigma-$Algebra, die alle offenen Mengen enthält; sie ist strikt in der Klasse der Lebesgue-messbaren Mengen enthalten; jede Lebesgue-messbare Menge kann jedoch durch eine Nullmenge so abgeändert werden, dass sie Borel wird). Es stimmt auch *nicht*, dass die Komposition messbarer Funktionen messbar ist.

Noch einmal:

Vorstellung: Eine Funktion ist messbar, wenn sie nicht allzu wild hin- und herspringt. Sie ist integrierbar, wenn sie messbar ist und ihr Absolutbetrag im Mittel nicht allzu groß ist.

2 Berechnung von Integralen

2.1 Der Satz von Fubini

Bisher können wir im Wesentlichen nur solche Funktionen integrieren, die wir schon in Analysis I integrieren konnten: Funktionen einer Variablen.

Frage: Wie berechnet man mehrdimensionale Integrale?
Antwort: Durch wiederholte eindimensionale Integration.

Schreibe $\mathbb{R}^n = \mathbb{R}^p \times \mathbb{R}^q$ (mit $p+q=n$) mit der Variable $(x,y) : x \in \mathbb{R}^p,\ y \in \mathbb{R}^q$.
Zum Beispiel: $\mathbb{R}^3 = \mathbb{R}^2 \times \mathbb{R}^1$.

2.1.1 Satz (Fubini)

Sei $f \in \mathcal{L}^1(\mathbb{R}^n)$. Dann gilt:

$$\int_{\mathbb{R}^n} f(x,y)\, d(x,y) = \int_{\mathbb{R}^q} \Big(\int_{\mathbb{R}^p} f(x,y)\, dx \Big)\, dy \qquad (*)$$

Genauer:

(1) Die Funktion $f(\cdot, y) : \mathbb{R}^p \to \bar{\mathbb{R}}, x \mapsto f(x,y)$ ist für fast alle $y \in \mathbb{R}^q$ über $\mathbb{R}^p$ integrierbar.

(2) Die Funktion $F(y) = \displaystyle\int_{\mathbb{R}^p} f(x,y)\, dx$ ist über $\mathbb{R}^q$ integrierbar.

(3) $(*)$ gilt.

Notation: Statt $d(x,y)$ schreibt man daher meist $\boldsymbol{dx\, dy}$.

Bemerkung: F ist nur fast überall definiert. Das macht aber nichts!

Beweis: Idee (etwa in $\mathbb{R}^2$): Da die Fläche eines Rechtecks gleich dem Produkt der Seitenlängen ist, gilt die Aussage für charakteristische Funktionen von Rechtecken. Indem man davon Linearkombinationen bildet, erhält man die Aussage für Treppenfunktionen. Die Formel gilt dann auch für allgemeine Funktionen, da man diese durch Treppen approximieren kann. Genauer:

(1) Für $f = \chi_Q$, wobei Q ein Quader in $\mathbb{R}^n$ ist, d. h. $Q = \underbrace{I_1 \times \cdots \times I_p}_{Q' \subset \mathbb{R}^p} \times \underbrace{I_{p+1} \cdots \times I_n}_{Q'' \subset \mathbb{R}^q} = Q' \times Q''$:
Es gilt: $\mathrm{vol}_n(Q) = \prod_{i=1}^{n} \text{Länge}(I_i) = \mathrm{vol}_p(Q') \cdot \mathrm{vol}_q(Q'')$, also
$\displaystyle\int_{\mathbb{R}^n} \chi_Q(x,y)\, d(x,y) = \int_{\mathbb{R}^p} \chi_{Q'}(x)\, dx \cdot \int_{\mathbb{R}^q} \chi_{Q''}(y)\, dy = \int_{\mathbb{R}^q} \Big(\int_{\mathbb{R}^p} \chi_{Q'}(x) \chi_{Q''}(y)\, dx \Big)\, dy$, und wegen
$\chi_Q(x,y) = \chi_{Q'}(x)\chi_{Q''}(y)$ ist das genau $(*)$.

(2) Da beide Seiten von $(*)$ bezüglich f linear sind, folgt die Behauptung für Treppenfunktionen.

(3) Es bleibt zu zeigen, dass die Behauptung unter Approximation bezüglich der $\|\cdot\|_1$-Norm richtig bleibt.

Wir führen erst eine abkürzende Notation ein: Für eine Funktion g auf $\mathbb{R}^p \times \mathbb{R}^q$ schreibe
$|g|_x : \mathbb{R}^q \to \bar{\mathbb{R}}$, $|g|_x(y) = \|g(\cdot, y)\|_{1,\mathbb{R}^p}$ (L^1-Halbnorm bezüglich x, für fixiertes $y \in \mathbb{R}^q$).
Für $G : \mathbb{R}^q \to \bar{\mathbb{R}}$ schreibe $|G|_y := \|G\|_{1,\mathbb{R}^q}$ (die »übliche« L^1-Halbnorm von G).

2.1.2 Lemma

$$\big|\, |g|_x \,\big|_y \le \|g\|_{1,\mathbb{R}^n}.$$

Beweis: Übung. □

Wir setzen nun den Beweis des Satzes von Fubini fort.

Sei $f \in \mathcal{L}^1(\mathbb{R}^n)$. Dann gibt es eine Folge von Treppenfunktionen $\{t_k\}_k$ mit $\|f - t_k\|_1 \xrightarrow{k\to\infty} 0$.
Ohne Beschränkung der Allgemeinheit kann man annehmen, dass $\|f - t_k\|_1 \le 2^{-k}$. (Gehe zu einer Teilfolge der t_k über.)
Setze $h(y) := \sum_{k=1}^{\infty} |f - t_k|_x(y)$ für $y \in \mathbb{R}^q$.
Dann ist $|h|_y \le \sum_{k=1}^{\infty} \big|\, |f - t_k|_x \,\big|_y \le \sum_{k=1}^{\infty} \|f - t_k\|_1 < \infty$ (nach dem Lemma).
Nach Lemma 1.4.6 folgt: $h(y) < \infty$ für fast alle $y \in \mathbb{R}^q$, also $|f - t_k|_x(y) \xrightarrow{k\to\infty} 0$ für fast alle $y \in \mathbb{R}^q$.
Für diese y ist $|f(\cdot, y) - t_k(\cdot, y)|_x \to 0$, und da $t_k(\cdot, y)$ Treppenfunktionen auf $\mathbb{R}^p$ sind, ist $f(\cdot, y) : \mathbb{R}^p \to \bar{\mathbb{R}}$ integrierbar über $\mathbb{R}^p$. Damit ist die Behauptung (1) bewiesen.

Setze $F(y) := \int_{\mathbb{R}^p} f(x,y)\,dx$, $T_k(y) := \int_{\mathbb{R}^p} t_k(x,y)\,dx$. Dann ist
$F(y) - T_k(y) = \int_{\mathbb{R}^p} (f(x,y) - t_k(x,y))\,dx$, also $|F(y) - T_k(y)| \le \int_{\mathbb{R}^p} |f(x,y) - t_k(x,y)|\,dx = |f - t_k|_x(y)$
und daher $|F - T_k|_y \le \big|\, |f - t_k|_x \,\big|_y \le \|f - t_k\|_1 \xrightarrow{k\to\infty} 0$ (nach dem Lemma).
Da die T_k Treppenfunktionen auf $\mathbb{R}^q$ sind und $|F - T_k|_y \xrightarrow{k\to\infty}$ gilt, folgt:
F ist integrierbar über $\mathbb{R}^q$ und $\int_{\mathbb{R}^q} F(y)\,dy = \lim_{k\to\infty} \int_{\mathbb{R}^q} T_k(y)\,dy$. Somit ist (2) bewiesen.

Schließlich ist

$$\begin{aligned}\int_{\mathbb{R}^n} f(x,y)\,d(x,y) &= \lim_{k\to\infty} \int_{\mathbb{R}^n} t_k(x,y)\,d(x,y) = \lim_{k\to\infty} \int_{\mathbb{R}^q} \underbrace{\Big(\int_{\mathbb{R}^p} t_k(x,y)\,dx\Big)}_{T_k(y)}\,dy \\ &= \lim_{k\to\infty} \int_{\mathbb{R}^q} T_k(y)\,dy = \int_{\mathbb{R}^q} F(y)\,dy \\ &= \int_{\mathbb{R}^q} \Big(\int_{\mathbb{R}^p} f(x,y)\,dx\Big)\,dy.\end{aligned}$$

Somit ist (3) bewiesen. □

Beispiel: $f(x,y) = e^{-|x|-|y|} \in \mathcal{L}^1(\mathbb{R}^2)$. (Woher wissen wir das eigentlich? Am einfachsten mit dem Satz von Tonelli, siehe unten.)
Nach dem Satz von Fubini (mit $p = q = 1$) folgt:

$$\begin{aligned}\int_{\mathbb{R}^2} e^{-|x|-|y|}\,dx\,dy &= \int_{\mathbb{R}} \Big(\int_{\mathbb{R}} e^{-|x|-|y|}\,dx\Big)\,dy = \int_{\mathbb{R}} e^{-|y|}\Big(\int_{\mathbb{R}} e^{-|x|}\,dx\Big)\,dy \\ &= \int_{\mathbb{R}} e^{-|y|}\Big(2\int_0^{\infty} e^{-x}\,dx\Big)\,dy = \int_{\mathbb{R}} e^{-|y|}(-2e^{-x}|_0^{\infty})\,dy = 2\underbrace{\int_{\mathbb{R}} e^{-|y|}\,dy}_{=2} = 4\end{aligned}$$

Wendet man Fubini mehrfach an, erhält man:

2.1.3 Korollar

Für $f \in \mathcal{L}^1(\mathbb{R}^n)$ und beliebige Permutationen $(i_1, \dots, i_n)$ von $(1, \dots, n)$ gilt:

$$\int_{\mathbb{R}^n} f(x)\,dx = \underbrace{\int_{\mathbb{R}} \cdots \int_{\mathbb{R}}}_{n-mal} f(x_1, \dots, x_n)\,dx_1 \cdots dx_n$$
$$= \underbrace{\int_{\mathbb{R}} \cdots \int_{\mathbb{R}}}_{n-mal} f(x_1, \dots, x_n)\,dx_{i_1} \cdots dx_{i_n} \qquad (**)$$

Dabei ist $\int_{\mathbb{R}} \cdots \int_{\mathbb{R}} f(x_1, \dots, x_n)\,dx_1 \cdots dx_n$ so zu lesen, dass man erst über x_1 integriert (für feste $x_2, \dots, x_n$), dann das Resultat über x_2 (für feste $x_3, \dots, x_n$) etc. Nach dem Korollar ist die Reihenfolge aber sowieso egal, falls $f \in \mathcal{L}(\mathbb{R}^n)$ ist.

In der Praxis weiß man oft nicht vorher, ob $f \in \mathcal{L}^1(\mathbb{R}^n)$ ist. Das würde man gerne gerade durch die »Fubini-Rechnung« herausfinden (vgl. das Beispiel oben). Das ist in folgendem Sinne möglich:

2.1.4 Satz (Tonelli)

Sei f eine lokal integrierbare Funktion auf $\mathbb{R}^n$ (d. h. f ist integrierbar über jeden Würfel $[-k,k]^n$). Dann gilt:

$f \in \mathcal{L}^1(\mathbb{R}^n) \Leftrightarrow$ Mindestens eins der iterierten Integrale $(**)$ existiert für $|f|$.

Die Voraussetzung ist zum Beispiel für stetige Funktionen auf $\mathbb{R}^n$ immer erfüllt.

Beweis: $\Rightarrow$ ist gerade das Korollar zu Fubini, da mit f auch $|f|$ in $\mathcal{L}^1(\mathbb{R}^n)$ ist.
Beweis des interessanteren Teils $\Leftarrow$, der Einfachheit halber für $n = 2$ und $i_1 = 1$, $i_2 = 2$: Setze $f_k = |f| \cdot \chi_{[-k,k]^n}$. Nach Voraussetzung ist f_k integrierbar, also kann man den Satz von Fubini auf f_k anwenden: $\int_{\mathbb{R}^2} f_k(x)\,dx = \int_{\mathbb{R}} (\int_{\mathbb{R}} f_k(x_1, x_2)\,dx_1)\,dx_2$. Wegen $f_k \le |f|$ ist dies $\le \int_{\mathbb{R}} (\int_{\mathbb{R}} |f(x_1, x_2)|\,dx_1)\,dx_2$. Also ist die Folge $(\int_{\mathbb{R}^2} f_k)_k$ beschränkt. Da offenbar $(f_k)_k$ monoton wachsend gegen $|f|$ konvergiert, folgt nach dem Satz über die monotone, integralbeschränkte Konvergenz, dass $|f| \in \mathcal{L}^1(\mathbb{R}^n)$, also auch $f \in \mathcal{L}^1(\mathbb{R}^n)$. □

Bemerkung: Es gilt *nicht*, dass die Existenz eines iterierten Integrals für f (statt $|f|$) schon die Integrierbarkeit von f impliziert.

Man kann sogar Beispiele $f : \mathbb{R}^2 \to \mathbb{R}$ angeben, wo beide iterierte Integrale $\int_{\mathbb{R}} (\int_{\mathbb{R}} f(x_1, x_2)\,dx_1)\,dx_2$, $\int_{\mathbb{R}} (\int_{\mathbb{R}} f(x_1, x_2)\,dx_2)\,dx_1$ existieren, aber verschiedene Werte haben (insbesondere kann nach Fubini f dann nicht in $\mathcal{L}^1(\mathbb{R}^2)$ sein).

Für die Volumenberechnung erhält man daraus ein wichtiges Prinzip:

2.1.5 Korollar

Sei $A \subset \mathbb{R}^n$ messbar. Schreibe $\mathbb{R}^n = \mathbb{R}^p \times \mathbb{R}^q$ mit $p + q = n$ und setze
$A_y := \{x \in \mathbb{R}^p : (x, y) \in A\}$.
Dann gilt

$$\mathrm{vol}_n(A) = \int_{\mathbb{R}^q} \mathrm{vol}_p(A_y)\,dy\,.$$

Zum Beispiel für $p = 2$, $q = 1$: Um das Volumen eines Körpers $A \subset \mathbb{R}^3$ zu berechnen, berechnet man zunächst die Flächen der Schnitte von A mit horizonalen Ebenen und integriert die so erhaltene Funktion von y (der »Höhe« der Ebene).

Beweis: $\mathrm{vol}_n(A) = \int_{\mathbb{R}^n} \chi_A(x,y)\,dx\,dy = \int_{\mathbb{R}^q} \Big(\underbrace{\int_{\mathbb{R}^p} \chi_A(x,y)\,dx}_{\mathrm{vol}_p(A_y)} \Big)\,dy = \int_{\mathbb{R}^p} \mathrm{vol}_p(A_y)\,dy.$ □

Beispiele:

(1) Fläche einer Kreisscheibe im $\mathbb{R}^2$:

$K = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \le r^2\}$. $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$.

Für $y \in [-r,r]$ ist $K_y = [-\sqrt{r^2-y^2}, \sqrt{r^2-y^2}]$, für $y \notin [-r,r]$ ist $K_y = \emptyset$. Also

$\mathrm{vol}_2(K) = \int_{[-r,r]} \mathrm{vol}_1(K_y)\,dy = \int_{-r}^{r} 2\sqrt{r^2-y^2}\,dy = 2(\frac{y}{2}\sqrt{r^2-y^2} + \frac{r^2}{2}\arcsin(\frac{y}{r}))|_{-r}^{r} = 2(\frac{r^2}{2}\arcsin(1) + \frac{r^2}{2}\arcsin(1)) = 2r^2\arcsin(1) = \pi r^2$.

(Man kann dies auch mit einer viel einfacheren Rechnung herleiten, mittels Polarkoordinaten. Wir kommen darauf bei der Transformationsformel zurück.)

(2) Volumen einer Kugel im $\mathbb{R}^3$:

$K = \{(x,y,z) : x^2 + y^2 + z^2 \le r^2\}$. $\mathbb{R}^3 = \mathbb{R}^2 \times \mathbb{R}$.

K_z ist die Kreisscheibe mit Radius $\sqrt{r^2-z^2}$, falls $z \in [-r,r]$, also $\mathrm{vol}_2(K_z) = \pi(r^2-z^2)$.

Also $\mathrm{vol}_3(K) = \int_{-r}^{r} \mathrm{vol}_2(K_z)\,dz = \int_{-r}^{r} \pi(r^2-z^2)\,dz = \pi(r^2 z - \frac{z^3}{3})|_{-r}^{r} = \frac{4}{3}\pi r^3$.

(3) Volumen eines Zylinders:

Sei $B \subset \mathbb{R}^{n-1}$ eine messbare Menge und $h > 0$ eine Zahl. Unter einem Zylinder mit Basis B und Höhe h versteht man die Menge $Z := B \times [0,h] \subset \mathbb{R}^n$.

Hier ist $Z_y = B$ (mit $\mathbb{R}^n = \mathbb{R}^{n-1} \times \mathbb{R}$), also $\mathrm{vol}_n(Z) = \int_0^h \mathrm{vol}_{\mathbb{R}^{n-1}}(Z_y)\,dy = \mathrm{vol}_{n-1}(B) \cdot h$.

Aus Korollar 2.1.5 folgt sofort das **Prinzip von Cavalieri**:

Sind $A, B \subset \mathbb{R}^n$ messbar und ist $\mathrm{vol}_p(A_y) = \mathrm{vol}_p(B_y)$ für alle $y \in \mathbb{R}^q$, so folgt $\mathrm{vol}_n(A) = \mathrm{vol}_n(B)$.

Dies wurde zur Volumenberechnung genutzt, als es die Integralrechnung noch nicht gab.

Beispiel: Volumen eines Kegels: Betrachte für $r > 0$

$\mathrm{Ke} = \{(x,y,z) : 0 \le z \le r,\ x^2 + y^2 \le z^2\}$ (Kegel),

$Z = \{(x,y,z) : 0 \le z \le r,\ x^2 + y^2 \le r^2\}$ (Zylinder),

$A = \{(x,y,z) : x^2 + y^2 + z^2 \le r^2, z \ge 0\}$ (Halbkugel),

$B = Z \setminus \mathrm{Ke}$ (der zwischen Kegel und Zylinder eingeschlossene Körper).

Dann haben die horizontalen Schnitte von A und B die gleiche Fläche:

$\mathrm{vol}_2(A_z) = \pi(r^2-z^2) = \mathrm{vol}_2(B_z)$

Also folgt $\mathrm{vol}_3(A) = \mathrm{vol}_3(B) = \mathrm{vol}_3(Z) - \mathrm{vol}_3(\mathrm{Ke})$.

Kennt man Volumen von Kugel und Zylinder, kann man also das Volumen des Kegels berechnen:

$\mathrm{vol}_3(\mathrm{Ke}) = \pi r^2 \cdot r - \frac{2}{3}\pi r^3 = \frac{1}{3}\pi r^3$.

(Wir könnten das natürlich auch direkt ausrechnen.)

2.2 Die Transformationsformel

Frage: Gibt es eine mehrdimensionale Version der Substitutionsregel?
Erinnerung an die Substitutionsregel für bestimmte Integrale: Sei $a < b \in \mathbb{R}$, T eine stetig differenzierbare Funktion auf $[a,b]$ mit $T(a) = \alpha$, $T(b) = \beta$.
Durch die Substitution $y = T(x)$, $\frac{dy}{dx} = T'(x) \;\Rightarrow\; dy = T'(x)\,dx$ erhält man für $f : [\alpha,\beta] \to \mathbb{R}$:

$$\int_a^b f(T(x))\,T'(x)\,dx = \int_\alpha^\beta f(y)\,dy$$

Die höherdimensionale Formel sieht so ähnlich aus:

2.2.1 Satz (Transformationsformel)

Seien $U, V \subset \mathbb{R}$ offen und $T : U \to V$ ein Diffeomorphismus, d. h. T ist bijektiv und T, T^{-1} sind stetig differenzierbar. Dann gilt

$$\int_U f(T(x))\,|\det DT_{|x}|\,dx = \int_{T(U)} f(y)\,dy \qquad (*)$$

Genauer: f ist genau dann über $V = T(U)$ integrierbar, wenn $(f \circ T)|\det DT|$ über U integrierbar ist, und dann gilt $(*)$.

Erinnerung: Das Differential von T in x ist eine lineare Abbildung $DT_{|x} : \mathbb{R}^n \to \mathbb{R}^n$, die bzgl. der Standardbasis von $\mathbb{R}^n$ durch folgende Matrix gegeben ist:

$$DT_{|x} = \begin{pmatrix} \frac{\partial T_1}{\partial x_1}(x) & \cdots & \frac{\partial T_1}{\partial x_n}(x) \\ \vdots & & \vdots \\ \frac{\partial T_n}{\partial x_1}(x) & \cdots & \frac{\partial T_n}{\partial x_n}(x) \end{pmatrix}$$

Wie merkt man sich $(*)$?
Oft ist es einfacher, $y(x)$ statt $T(x)$ zu schreiben. Für die Matrix $Dy = \left(\frac{\partial y_i}{\partial x_j}\right)_{i,j=1,\to n}$ schreibt man auch $\frac{\partial(y_1,\ldots,y_n)}{\partial(x_1,\ldots,x_n)}$ (natürlich wird alles an einem Punkt x ausgewertet!). Dann wird $(*)$ zu

$$\int_V f(y)\,dy_1 \ldots dy_n = \int_U f(y(x)) \left|\det \frac{\partial(y_1,\ldots,y_n)}{\partial(x_1,\ldots,x_n)}\right| dx_1 \ldots dx_n$$

Man nennt T auch einen **Koordinatenwechsel**.

Polarkoordinaten

Was passiert bei der Substitution $x = r\cos\varphi$, $y = r\sin\varphi$?
Als Abbildung ist das

$$T : (r,\varphi) \mapsto (x,y) = (r\cos(\varphi), r\sin(\varphi))$$
$$(0,\infty) \times (0,2\pi) \to \mathbb{R}^2 \setminus \{(x,0) | x \geq 0\}$$

Dann folgt

$$DT = \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \varphi} \end{pmatrix} = \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix}$$

$$\begin{aligned} \det(DT) &= \cos\varphi \cdot r\cos\varphi - (-r\sin\varphi)\cdot\sin\varphi \\ &= r(\cos^2\varphi + \sin^2\varphi) = r \end{aligned}$$

Also

$$\int_{\mathbb{R}^2} f(x,y)\,dx\,dy = \int_0^\infty \int_0^{2\pi} f(r\cos\varphi, r\sin\varphi)\cdot r\;\; d\varphi\,dr$$

Genau genommen steht links das Integral über $\mathbb{R}^2$ minus die positive x-Achse. Da diese aber eine Nullmenge ist, kann man das Integral auch über ganz $\mathbb{R}^2$ nehmen.

Warum haben wir $r = 0$ und $\varphi = 2\pi$ bei der Definition herausgenommen? Damit T zu einem Diffeomorphismus wird: Mit $r = 0$ wäre T nicht injektiv, da $T(0,\varphi) = (0,0)$ für alle φ, und mit $\varphi = 2\pi$ wäre die Umkehrung von T nicht stetig (und der Definitionsbereich von T nicht offen).
Aber wegen des Nullmengenarguments ist das am Ende irrelevant.

Beispiele:

(1) Fläche des Kreises vom Radius R, $K_R = \{(x,y) : x^2 + y^2 \le R^2\}$.
In Polarkoordinaten ist K_R durch $r \in [0,R]$, φ beliebig gegeben, also ist

$$\mathrm{vol}_2(K_R) = \int_0^R \int_0^{2\pi} r\,dr\,d\varphi = 2\pi \int_0^R r\,dr = \pi R^2\,.$$

(Wenn Sie's etwas formaler mögen: $\chi_{K_R}(r\cos\varphi, r\sin\varphi) = \chi_{[0,R]}(r)$, also $\mathrm{vol}_2(K) = \int_{\mathbb{R}^2} \chi_K(x,y)\,dx\,dy = \int_0^\infty \int_0^{2\pi} \chi_{[0,R]}(r) r\,d\varphi\,dr = \int_0^R \int_0^{2\pi} r\,dr\,d\varphi$ etc.)

(2) Sei $f(x,y) = e^{-x^2-y^2} = e^{-(x^2+y^2)} = e^{-r^2}$, also $f(r\cos\varphi, r\sin\varphi) = e^{-r^2}$.
Aus

$$\begin{aligned}\int_{-\infty}^\infty \int_{-\infty}^\infty e^{-x^2-y^2}\,dx\,dy &= \int_0^\infty \int_0^{2\pi} e^{-r^2} r\,d\varphi\,dr \\ &= 2\pi \int_0^\infty r e^{-r^2}\,dr \\ &= 2\pi \cdot \frac{1}{2} \int_0^\infty e^{-s}\,ds \\ &= \pi(-e^{-s})|_0^\infty = \pi\end{aligned}$$

und andererseits (Fubini)

$$\begin{aligned}\int_{-\infty}^\infty \int_{-\infty}^\infty e^{-x^2-y^2}\,dx\,dy &= \int_{\mathbb{R}} \int_{\mathbb{R}} e^{-x^2} e^{-y^2}\,dx\,dy \\ &= \int_{\mathbb{R}} e^{-y^2} \left(\int_{\mathbb{R}} e^{-x^2}\,dx \right) dy \\ &= \int_{\mathbb{R}} e^{-x^2}\,dx \cdot \int_{\mathbb{R}} e^{-y^2}\,dy \\ &= \left(\int_{\mathbb{R}} e^{-x^2}\,dx \right)^2\end{aligned}$$

folgt als hübsches Nebenprodukt:

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

Verständnisfrage 1: Warum stehen in der Transformationsformel Betragsstriche? In der Substitutionsregel stehen doch keine!
Betrachten wir also wieder $n = 1$, also $U = (a,b)$. Sei $T(a) = \alpha$, $T(b) = \beta$.
T Diffeomorphismus $\Rightarrow T'(x) = \det DT \neq 0$ für alle x.
Also gilt $T'(x) > 0$ für alle $x \in (a,b)$ oder $T'(x) < 0$ für alle $x \in (a,b)$.

1. Fall: $T'(x) > 0$ für alle $x \in (a,b)$. Dann ist T streng monoton wachsend, also $\alpha < \beta$, also $V = (\alpha, \beta)$. Mit Satz 1.3.6 (sagen wir, f sei stetig auf $[a,b]$) und der Transformationsformel folgt
$\int_a^b f(T(x))T'(x)\,dx = \int_{(a,b)} f(T(x))|T'(x)|\,dx = \int_{(\alpha,\beta)} f(y)\,dy = \int_\alpha^\beta f(y)\,dy$.

2. Fall: $T'(x) < 0$ für alle $x \in (a,b)$. Dann ist T streng monoton fallend, also $\alpha > \beta$, also $V = (\beta, \alpha)$ (!!!). Nun ist $\int_a^b f(T(x))T'(x)\,dx = -\int_{(a,b)} f(T(x))|T'(x)|\,dx = -\int_{(\beta,\alpha)} f(y)\,dy = -\int_\beta^\alpha f(y)\,dy = \int_\alpha^\beta f(y)\,dy$.
In jedem Fall erhalten wir die gewohnte Substitutionsregel.

> *Die Betragsstriche um die Determinante gleichen aus, dass die Integralgrenzen bei negativen $T'(x)$ ausgetauscht werden.*

Verständnisfrage 2: Warum steht in der Transformationsformel die Determinante? Eine erste Idee bekommt man aus folgendem einfachen Beispiel:

Beispiel: Sei $U = V = \mathbb{R}^n$, $T(x) = 2x$. Also $DT = \begin{pmatrix} 2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 2 \end{pmatrix}$, $\det DT = 2^n$.

Die Transformationsformel ergibt:

$$\int f(2x)\,dx = \frac{1}{2^n} \int f(y)\,dy$$

Aus Analysis I kennen wir das bereits für $n = 1$. Warum steht hier 2^n, nicht 2? Weil man die Analysis-I-Substitution bezüglich jeder Variablen $x_1, \ldots, x_n$ separat macht:

$$\begin{aligned} \int f(2x)\,dx &= \int f(2x_1, \ldots 2x_n)\,dx_1 \cdots dx_n \\ &= \int \cdots \int f(y_1, \ldots y_n) \frac{dy_1}{2} \cdots \frac{dy_n}{2} \quad = \frac{1}{2^n} \int f(y)\,dy\,. \end{aligned}$$

Wendet man das auf charakteristische Funktionen an, erhält man die wichtige Einsicht:

> Verdoppelt man alle Abmessungen eines dreidimensionalen Körpers, dann verachtfacht sich das Volumen.

2.3 Beweis der Transformationsformel

Wir beweisen nun die Transformationsformel

$$\int_U (f \circ T)\, \det DT = \int_{T(U)} f\,.$$

Der Beweis ist sehr lehrreich. Zum einen lässt sich die Idee leicht erklären, zum anderen ist die Durchführung dieser Idee nicht ganz einfach, und wir werden hierzu vieles verwenden, was Sie in Analysis I-III und in der Linearen Algebra kennengelernt haben.

Woher kommt die Determinante?

Als die Determinante in der linearen Algebra eingeführt wurde, haben Sie sich vielleicht gefragt: Warum? Wie kommt man drauf?

Der Ursprung der Determinante liegt in der Volumen- (bzw. Flächen-)berechnung. Zu n Vektoren $a_1, \ldots, a_n \in \mathbb{R}^n$ bezeichne

$$P(a_1, \ldots, a_n) = \left\{ \sum_{i=1}^n t_i a_i : \; t_i \in [0,1] \;\text{ für alle }\; i \right\}$$

das von $a_1, \ldots, a_n$ aufgespannte **Parallelotop**.

2.3.1 Satz

(a) Das von $a_1, \dots, a_n$ aufgespannte Parallelotop hat das Volumen

$$\mathrm{vol}(P(a_1, \dots, a_n)) = |\det(a_1, \dots, a_n)|$$

(b) Sei $T : \mathbb{R}^n \to \mathbb{R}^n$ linear. Dann gilt für alle messbaren Mengen $M \subset \mathbb{R}^n$:

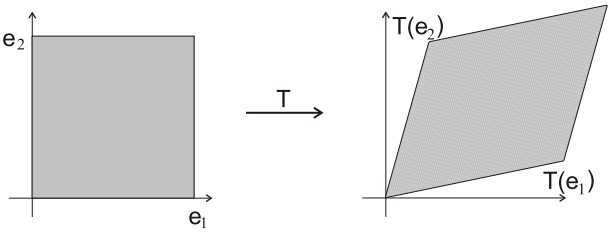
$$\mathrm{vol}(T(M)) = |\det T| \cdot \mathrm{vol}(M)$$

(c) Sei $T : \mathbb{R}^n \to \mathbb{R}^n$ linear und invertierbar. Dann gilt für alle $f \in \mathcal{L}^1(\mathbb{R}^n)$

$$\int f = |\det T| \int (f \circ T)$$

Genauer: Sei f messbar; dann ist $f \in \mathcal{L}^1 \Leftrightarrow f \circ T \in \mathcal{L}^1$, und die Formel gilt.

(c) ist ein Spezialfall der Transformationsformel, da für lineare Abbildungen $T : \mathbb{R}^n \to \mathbb{R}^n$ gilt: $DT_{|x} = T$ für alle x.



Beweis: Zunächst beobachten wir, dass (a) aus (b) folgt. Denn zu gegebenen $a_1, \dots, a_n \in \mathbb{R}^n$ sei T die eindeutig bestimmte lineare Abbildung mit $T(e_i) = a_i$, $i = 1, \dots, n$. Dann wird T durch die Matrix $(a_1, \dots, a_n)$ dargestellt (a_i sind als Spaltenvektoren geschrieben). Wendet man (b) auf den Einheitswürfel an, $M = [0,1]^n = \left\{ \sum_{i=1}^{n} t_i e_i : \ t_i \in [0,1] \ \text{ f. a. } i \right\}$, so ist offenbar $T(M) = P(a_1, \dots, a_n)$, und wegen $\mathrm{vol}(M) = 1$ folgt (a).

(b) ist ein Spezialfall von (c), falls T invertierbar ist: Es gilt

$$\chi_{T(M)}(T(x)) = \begin{cases} 1 & \text{falls } T(x) \in T(M) \\ 0 & \text{sonst.} \end{cases}$$

Da T invertierbar ist, gilt $T(x) \in T(M) \Leftrightarrow x \in M$, also folgt $\chi_{T(M)} \circ T = \chi_M$. Wendet man (c) auf $f = \chi_{T(M)}$ an, folgt $\mathrm{vol}(T(M)) = \int \chi_{T(M)} = |\det T| \int \chi_{T(M)} \circ T = |\det T| \int \chi_M = |\det T| \, \mathrm{vol}(M)$, also (b). (Siehe die Bemerkung unten zum Fall $\mathrm{vol}(M) = \infty$.)

Beweis von (b), wenn T nicht invertierbar ist: $T(M)$ ist in einer Hyperebene enthalten, da $\dim \mathrm{Bild}(T) \leq n - 1$, also Nullmenge, also sind beide Seiten von (b) gleich null.

Beweis von (c): Aus der linearen Algebra ist bekannt: Jede lineare Abbildung ist Komposition von mehreren Abbildungen der folgenden Typen:

$$A : \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \longrightarrow \begin{pmatrix} ax_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \ a \in \mathbb{R} \qquad B : \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \longrightarrow \begin{pmatrix} x_1 + x_2 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \qquad C_{ij} : \begin{pmatrix} x_1 \\ \vdots \\ x_i \\ \vdots \\ x_j \\ \vdots \\ x_n \end{pmatrix} \longrightarrow \begin{pmatrix} x_1 \\ \vdots \\ x_j \\ \vdots \\ x_i \\ \vdots \\ x_n \end{pmatrix} \ \text{f. a. } 1 \leqq i, j \leqq n$$

(Das ist nichts anderes als das Gaußsche Eliminationsverfahren: Multiplizieren einer Zeile mit einem Faktor, addieren einer Zeile zu einer anderen, Vertauschen von Zeilen; mit Hilfe geeigneter Vertauschungen kann man auch andere als die erste Zeile mit einem Faktor multiplizieren.)

Es genügt, (c) für $T = A, B, C_{ij}$ nachzuprüfen, denn:

Behauptung: Falls (c) für lineare Abbildungen T_1 und T_2 gilt, so gilt (c) auch für $T_1 \circ T_2$.

Beweis: Schreibe $T = T_1 \circ T_2$. Dann

$$\begin{aligned}\int f &= |\det T_1| \cdot \int f \circ T_1 = |\det T_1|\big(|\det T_2| \cdot \int (f \circ T_1) \circ T_2\big) \\ &= |\det(T_1 \circ T_2)| \cdot \int f \circ (T_1 \circ T_2) \\ &= |\det(T_1 \circ T_2)| \cdot \int f \circ T\end{aligned}$$

□

Nachprüfen von (c) für $T = A$: Nach Fubini ist

$$\int f(y)\, dy = \int \dots \int f(y_1, \dots y_n)\, dy_1 \dots dy_n\,.$$

1-dimensionale Variablensubstitution $y_1 = ax_1$ im y_1-Integral ergibt mit $dy_1 = a\, dx_1$:

$$\int f(y_1, \dots, y_n)\, dy_1 = |a| \int f(a\, x_1, y_2, \dots y_n)\, dx_1$$

Integration über $y_2, \dots, y_n$ ergibt $\int f = |a| \cdot \int f \circ A$. Das ist (c), da als Matrix

$$A = \begin{pmatrix} a & & 0 \\ & 1 & \\ & \ddots & \\ 0 & & 1 \end{pmatrix},$$

also $\det A = a$ gilt.

Nachprüfen von (c) für $T = B$: Zunächst gilt für beliebige $y_2, \dots, y_n$, dass

$$\int f(x_1 + y_2, y_2, \dots, y_n)\, dx_1 = \int f(y_1, y_2, \dots, y_n)\, dy_1$$

(Substitution $y_1 = x_1 + y_2$). Integriert man beide Seiten über $y_2, \dots, y_n$, folgt $\int f = \int f \circ B$. Das ist (c), weil

$$B = \begin{pmatrix} 1 & 1 & & 0 \\ & 1 & & \\ & & \ddots & \\ 0 & & & 1 \end{pmatrix},$$

also $\det B = 1$ gilt.

Nachprüfen von (c) für $T = C$: Vertauschen der Integrationsreihenfolge ändert das Integral nicht, also gilt $\int f = \int f \circ C_{ij}$ für alle i, j. Dies ist (c), weil

$$C = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 0 & 1 & & \\ & & 1 & 0 & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix},$$

also $|\det C_{ij}| = |-1| = 1$ gilt. □

Bemerkung: (Dies sollte eigentlich schon im Kapitel über messbare Funktionen stehen!) Ist $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ eine messbare Funktion mit $f \geq 0$, so kann $\int f$ immer definiert werden, wobei auch

$$\int f = \infty$$

möglich ist. Genauer:

Sei $K_1 \subset K_2 \subset \ldots \mathbb{R}^n$ eine Ausschöpfung durch kompakte Mengen und $f_N = \chi_{K_N} \min\{f, N\}$ (man schneidet f nach »oben« und in allen x-Richtungen ab), dann definiere

$$\int f := \lim_{N\to\infty} \int f_N \,.$$

Da f_N messbar und durch die integrierbare Funktion $N\chi_{K_N}$ beschränkt ist, ist $f_N \in \mathcal{L}^1(\mathbb{R}^n)$, also ist $\int f_N$ definiert. Außerdem konvergiert offenbar f_N für $N \to \infty$ punktweise und monoton wachsend gegen f. Mit dem Satz über die monotone, integralbeschränkte Konvergenz folgt leicht:

- $\int f \in [0, \infty]$ ist wohldefiniert, insbesondere unabhängig von der Wahl der K_N.
- $\int f < \infty \iff f \in \mathcal{L}^1(\mathbb{R}^n)$.
- Für beliebige messbare Mengen $M \subset \mathbb{R}^n$ gilt $\mathrm{vol}(M) = \int \chi_M$.

Dies ist ähnlich zur Situation bei Reihen: Sind alle $a_i \geq 0$, so ist $\sum_{i=1}^{\infty} a_i \in [0, \infty]$ immer definiert; die Reihe konvergiert genau dann, wenn dies endlich ist.

Beachte: Hierbei ist es wesentlich, dass $f \geq 0$ ist!

Aus Satz 2.3.1 folgt sofort:

2.3.2 Korollar

Das Lebesgue-Integral und Lebesgue-Maß sind invariant unter Euklidischen Bewegungen, d. h. Verknüpfungen von Translationen, Spiegelungen und Rotationen. Das heißt, ist T eine Euklidische Bewegung, so gilt für messbare Mengen M und integrierbare Funktionen f:

$$\mathrm{vol}(T(M)) = \mathrm{vol}(M)\,, \quad \int f \circ T = \int f$$

Bemerkung: Zur Erinnerung etwas **Allgemeinwissen aus Geometrie und linearer Algebra**:
Eine lineare Abbildung $T : \mathbb{R}^n \to \mathbb{R}^n$ (allgemeiner: auf einem Vektorraum mit Skalarprodukt) heißt **orthogonal**, wenn sie Längen erhält, d. h. $\|T(v)\| = \|v\|$ für alle v gilt. Dies ist äquivalent dazu, dass $T^t T = I$ (I ist die Identität, T^t ist die Transponierte von T). Daraus folgt sofort $\det T = \pm 1$. Gilt $\det T = 1$, so heißt T **speziell orthogonal** oder **Rotation**. In R^2 und R^3 sind das wirklich die Rotationen im üblichen Sinn.

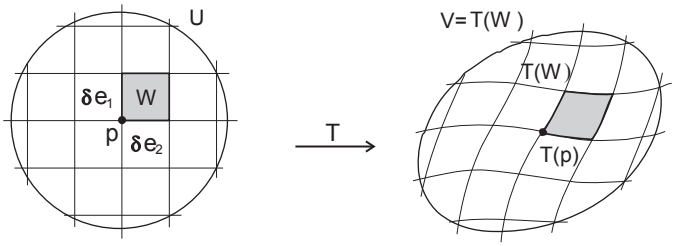
Eine orthogonale lineare Abbildung heißt **Spiegelung**, wenn es zwei orthogonale Unterräume U, V mit $\mathbb{R}^n = U \oplus V$ gibt, so dass $T_{|U} = I_U$ (die Identität auf U) und $T_{|V} = -I_V$. U ist dann die Spiegelungsebene (bzw. -gerade bzw. -punkt bzw....), V ihr orthogonales Komplement. Äquivalent dazu ist, dass $T^2 = I$ ist (setze dann $U = \ker(T - I), V = \ker(T + I)$).

Ein (einfacher) wichtiger Satz: Jede orthogonale lineare Abbildung kann als Produkt einer Spiegelung und einer Rotation geschrieben werden.

Noch ein wichtiger Satz, der viel schwieriger zu zeigen ist: Jede Abbildung $T : \mathbb{R}^n \to \mathbb{R}^n$, die Abstände erhält (also $\|T(x) - T(y)\| = \|x - y\|$ für alle $x, y \in \mathbb{R}^n$), kann als Produkt einer Translation und einer orthogonalen linearen Abbildung geschrieben werden.
(Beachte: Hierbei wird nicht vorausgesetzt, dass T (affin) linear ist; es wird gefolgert!)

Beweis (des Korollars): Wir wissen bereits die Translationsinvarianz (Satz 1.6.4). Rotationen und Spiegelungen sind lineare Abbildungen mit $|\det T| = 1$, also folgt die Behauptung aus Satz 2.3.1. □

Die Hauptidee des Beweises der allgemeinen Translationsformel



Erinnerung: Bedeutung des Differentials $DT_{|p}$:

$$W = \text{ kleiner Würfel bei } p \text{ mit Seiten } \delta e_1, \ldots, \delta e_n \ \ (\delta > 0 \text{ klein})$$
$$\Rightarrow \quad T(W) \approx \text{ Spat bei } T(p) \text{ mit Seiten } DT_{|p}(\delta e_1), \ldots, DT_{|p}(\delta e_n)$$

(denn $DT_{|p}$ ist die lineare Approximation an T nahe p, und $DT_{|p}$ bildet W exakt auf diesen Spat ab – bis auf eine Verschiebung).
Damit ist, mit $A = DT_{|p}$:

$$\begin{aligned}\operatorname{vol}(T(W)) &\approx \operatorname{vol}\left(\text{Spat mit Seiten } A(\delta e_1), \ldots, A(\delta e_n)\right)\\ &= |\det A| \cdot \operatorname{vol}\left(\text{Würfel mit Seiten } \delta e_1, \ldots, \delta e_n\right)\\ &= |\det A| \cdot \operatorname{vol}(W)\end{aligned}$$

Zerlege nun U in viele kleine achsenparallele Würfel:

$$U = \bigcup_i W_i\,. \quad \text{Wähle} \qquad p_i \in W_i\,. \quad \text{Dann ist}$$
$$T(U) = \bigcup_i T(W_i)\,, \qquad T(p_i) \in T(W_i)\,.$$

Dann folgt

$$\begin{aligned}\int_{T(U)} f = \sum_i \int_{T(W_i)} f &\approx \sum_i f(T(p_i)) \cdot \operatorname{vol}(TW_i)\\ &\approx \sum_i f(T(p_i)) \cdot \left|\det DT_{|p_i}\right| \cdot \operatorname{vol}(W_i)\\ &\approx \sum \int_{W_i} (f \circ T) \cdot |\det DT|\\ &= \int_U (f \circ T) \cdot |\det T|\,.\end{aligned}$$

Hierbei stammt

das erste und dritte $\approx$ von der Approximation von f bzw. $(f \circ T) \cdot |\det DT|$ durch Treppen,

das zweite $\approx$ von der Approximation von T mittels $DT_{|p_i}$ nahe p_i.

Problem: $\approx$ gilt nur für »kleine« Würfel.

- $\rightsquigarrow$ viele Summanden
- $\rightsquigarrow$ Fehler addieren sich auf
- $\rightsquigarrow$ sie könnten die Formel zerstören

Um zu sehen, dass das nicht passiert, müssen wir sehr genau argumentieren. Der Schlüssel liegt im Wesentlichen darin, dass der *relative* Fehler bei der approximativen Berechnung von $\operatorname{vol}(T(W))$ für kleine W beliebig klein wird, nicht der absolute Fehler.

Beweis der Transformationsformel

Vorbereitung zum Beweis: In diesem Satz bezeichne $|x| := \max\{|x_1|, \dots |x_n|\}$ für $x \in \mathbb{R}^n$ die Maximumsnorm. Dann ist die »Kugel« $\{x : |x| \le r\} = \{x : |x_i| \le r \text{ f. a. } i\}$ ein Würfel mit Kantenlänge $2r$.

Für $A : \mathbb{R}^n \longrightarrow \mathbb{R}^n$, $A = (a_{ij})$, setze $\|A\| := \max_i \underbrace{\sum_{j=1}^{n} |a_{ij}|}_{i\text{-te Zeilensumme über die Beträge}}$. Dann gilt:

- ▷ $\|\cdot\|$ ist eine Norm auf $\mathbb{R}^{n\times n}$, der Menge der $n \times n$-Matrizen (Übung).
- ▷ $|Av| \le \|A\| \cdot |v|$ für alle $v \in \mathbb{R}^n$. Denn

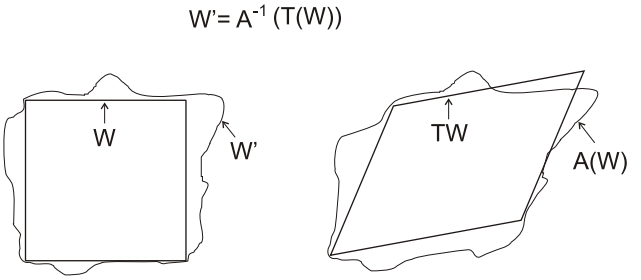
$$\begin{aligned} |(Av)_i| = |\sum_j a_{ij} v_j| &\le \sum_j |a_{ij}| \cdot |v_j| \\ &\le \Big(\sum_j |a_{ij}|\Big) \max_j |v_j| \le \|A\| \cdot |v| \,, \end{aligned}$$

  also $|Av| = \max_i |(Av)_i| \le \|A\| \cdot |v|$.

Hilfssatz 1: Sei $K \subset U$ kompakt, $\varepsilon > 0$. Dann existiert $\delta > 0$ mit:
Für alle Würfel $W \subset K$ mit Kantenlänge $< \delta$ ist

$$(1+\varepsilon)^{-n} |\det A| \operatorname{vol}(W) \le \operatorname{vol}(T(W)) \le (1+\varepsilon)^n |\det A| \operatorname{vol}(W) \,.$$

(Dies ist die Präzisierung der Aussage über den relativen Fehler im Abschnitt ›Hauptidee..‹)



Beweis:
1. Schritt: Sei $p \in U$. Wir beweisen die Aussage für Würfel mit Mittelpunkt p.
Sei o. B. d. A. $p = 0$, $T(p) = 0$. Setze $A = DT_{|p}$.
Ziel: Zeige, dass $T(W)$ »nahe« $A(W)$ ist. Setze $S = A^{-1} \circ T$, $W' = S(W)$.
Dann $T(W) = A(S(W)) = A(W')$, es ist also zu zeigen, dass W' »nahe« W ist.
Genauer: Da A linear ist, gilt $\operatorname{vol}(T(W)) = \operatorname{vol}(A(W')) = |\det A| \cdot \operatorname{vol}(W')$, es ist also zu zeigen:

$$(1+\varepsilon)^{-n} \operatorname{vol}(W) \le \operatorname{vol}(S(W)) \le (1+\varepsilon)^n \operatorname{vol} W \qquad (*)$$

Bemerkung: $DS_{|0} = DA^{-1} \circ DT_{|0} = A^{-1} \circ A = I$, also ist $(*)$ ein Spezialfall des Hilfssatzes mit $A = I$.
Erinnerung: Schrankensatz: Für $x \in W$ ist

$$|T(x) - T(0)| \le \Big(\max_{q\in W} \|DT_{|q}\|\Big) \cdot |x - 0|$$

Das wäre eigentlich in Analysis II dran gewesen, da ich es dort nicht gemacht habe, hier der einfache Beweis (derselbe Beweis funktioniert mit der Euklidischen Norm für Vektoren und der in Analysis II definierten Matrixnorm für $DT_{|q}$).

Beweis: Es gilt

$$T(x) - T(0) = \int_0^1 \underbrace{\left[\frac{d}{dt}\, T(tx)\right]}_{DT_{|tx}(x)}\, dt$$

(setze $h(t) = T(tx)$ und verwende $h(1) - h(0) = \int_0^1 h'(t)dt$) und $\frac{d}{dt}T(tx) = DT_{|tx}(x)$ nach der Kettenregel, also

$$|T(x) - T(0)| \le \int_0^1 |DT_{|tx}(x)|\, dt \le \int_0^1 \|DT_{|tx}\| \cdot |x|\, dt \le \left(\max_{q\in W} \|DT_{|q}\|\right) \cdot |x|$$

□

Wir zeigen nun, dass $(*)$ gilt, falls die Seitenlänge von W kleiner als δ ist, für ein gewisses $\delta > 0$.
Schreibe $W_r = \{x : |x| \le r\}$.

Beweis der rechten Ungleichung von $(*)$:
Für $x \in W_r$ ist $|S(x)| \underbrace{\le}_{\text{Schrankensatz}} \left(\max_{W_r} \|DS\|\right) \cdot |x|$. Es gilt:

- ▷ $DS_{|0} = I$, also $\|DS_{|0}\| = 1$.
- ▷ $x \mapsto \|DS_{|x}\|$ ist stetig.

Also gibt es zu $\varepsilon > 0$ ein $\delta > 0$, so dass für $2r < \delta$ gilt: $\max_{W_r} \|DS\| \le 1 + \varepsilon$.
Damit ist für $x \in W_r$: $|S(x)| \le (1+\varepsilon)|x| \le (1+\varepsilon)r$.
Das bedeutet $S(W_r) \subset W_{(1+\varepsilon)r} \;\Rightarrow\; \mathrm{vol}(S(W_r)) \le \mathrm{vol}(W_{(1+\varepsilon)r}) = (1+\varepsilon)^n \mathrm{vol}(W_r)$, was zu zeigen war.

Beweis der linken Ungleichung von $(*)$:
Trick: Argumentiere analog mit der inversen Abbildung S^{-1}!
S^{-1} ist auch C^1, und $D(S^{-1})_{|0} = (DS_{|0})^{-1} = I^{-1} = I$. Für r genügend klein folgt wieder $|S^{-1}(y)| \le (1+\varepsilon)|y|$ für alle y mit $|y| < r$. Also

$$|y| \le \frac{r}{1+\varepsilon} \;\Rightarrow\; |S^{-1}(y)| \le (1+\varepsilon)\frac{r}{1+\varepsilon} = r \;\Rightarrow\; S^{-1}(y) \in W_r \;\Rightarrow\; y \in S(W_r)\,.$$

Das heißt $W_{\frac{r}{1+\varepsilon}} \subset S(W_r)$ und daher $(1+\varepsilon)^{-n}\mathrm{vol}(W_r) = \mathrm{vol}(W_{\frac{r}{1+\varepsilon}}) \le \mathrm{vol}(S(W_r))$, was zu zeigen war.

2. Schritt im Beweis von Hilfssatz 1: Es ist zu zeigen, dass dasselbe δ für alle Würfel innerhalb K funktioniert. Dies ist ein typisches **Kompaktheitsargument**.
Im ersten Schritt war $S = A^{-1} \circ T = (DT_{|p})^{-1} \circ T$, und es kam an auf: $\max_{\substack{q\in \text{ Würfel}\\ \text{um } p}} \|DS_{|q}\| \le 1 + \varepsilon$. Man beachte nun (Kettenregel):

$$DS_{|q} = (DT_{|p})^{-1} \circ DT_{|q}$$

Damit folgt der Hilfssatz aus:

Behauptung: Sei $a(p,q) = \|(DT_{|p})^{-1}\, DT|_q\|$, $p, q \in K$. Dann gilt: $\forall \varepsilon > 0\ \exists \delta > 0$, so dass $\forall p, q \in K$

$$(**) \qquad |p - q| < \delta \quad\Longrightarrow\quad a(p,q) < 1 + \varepsilon$$

Beweis (der Behauptung): Die hierfür wesentlichen Eigenschaften von $a : K \times K \longrightarrow \mathbb{R}$ sind:

- ▷ a ist stetig (denn T ist *stetig* differenzierbar).
- ▷ $a(p,p) = 1$ für alle $p \in K$. Mit anderen Worten, $a_{|\Delta_K} \equiv 1$, wobei $\Delta_K \coloneqq \{(p,p) : p \in K\} \subset K \times K$ (»Diagonale«).

Die Behauptung $(**)$ lässt sich äquivalent umformulieren als $a(p,q) \geq 1+\varepsilon \Rightarrow |p-q| \geq \delta$, und dieses als

$$(p,q) \in a^{-1}([1+\varepsilon,\infty)) \overset{!}{\Rightarrow} |p-q| \geq \delta\,.$$

Im metrischen Raum $K \times K$ (mit der »Produktmetrik« $d((p,q),(p',q')) = |p-p'| + |q-q'|$) gilt nun:

- $A := a^{-1}([1+\varepsilon,\infty))$ ist abgeschlossen (denn a ist stetig und $[1+\varepsilon,\infty)$ ist abgeschlossen).
- $B := \Delta_K$ ist kompakt (denn Δ_K ist das Bild von K unter der stetigen Abbildung $p \mapsto (p,p)$),
- $A \cap B = \emptyset$ (denn $a_{|\Delta_K} \equiv 1$).

Erinnerung (Analysis II): Sei (X,d) metrischer Raum, $A, B \subset X$. Dann gilt:
A kompakt, B abgeschlossen, $A \cap B = \emptyset \Rightarrow \exists \delta > 0$, so dass für alle $x \in A$, $y \in B$ gilt: $d(x,y) \geq \delta$.
Wir verwenden dies mit $X = K \times K$ und A, B wie oben. Ist also $(p,q) \in a^{-1}([1+\varepsilon,\infty))$, so folgt mit diesem δ (mit $x = (p,q), y = (p,p)$)

$$\delta \leq d((p,q),(p,p)) = |p-p| + |q-p| = |p-q|\,, \qquad \square$$

was zu zeigen war. $\square$

Hilfssatz 2:
Sei $Q \subset U$ ein kompakter Quader. Dann ist

$$\mathrm{vol}(T(Q)) = \int_Q |\det T|$$

(Hier wird im Wesentlichen gezeigt, warum bei dem mittleren $\approx$ in der »Hauptidee« nichts schief geht.)

Beweis:

1. Fall: Die Seitenlängen von Q sind rationale Zahlen.

Sei N ein gemeinsamer Nenner, dann kann man Q als Vereinigung von Würfeln W_i schreiben mit Seitenlängen $\frac{1}{N}$, die sich nur in Randflächen überschneiden.

Sei $\varepsilon > 0$. Wähle $\delta > 0$ wie in Hilfssatz 1 für $K = Q$. Wähle Würfelchen W_i wie oben, mit $\frac{1}{N} < \delta$. Sei p_i der Mittelpunkt von W_i. Aus Hilfssatz 1 folgt

$$\begin{aligned}
\mathrm{vol}(T(Q)) &= \sum_i \mathrm{vol}(T(W_i)) \\
&\leq \sum (1+\varepsilon)^n \cdot |\det DT_{|p_i}| \cdot \mathrm{vol}(W_i) \\
&= (1+\varepsilon)^n \cdot \sum_i |\det DT_{|p_i}| \,\mathrm{vol}(W_i)
\end{aligned}$$

Die letzte Summe ist das Integral über Q der Treppenfunktion $\sum_i |\det DT_{|p_i}| \chi_{W_i}$, und diese konvergiert für $N \to \infty$ gleichmäßig gegen die stetige Funktion $|\det DT|$. Für $N \to \infty$ (d. h. Seitenlänge $\to 0$) gilt also $(\sum_i \ldots) \longrightarrow \int_Q |\det DT|$, also ergibt sich

$$\mathrm{vol}(T(Q)) \leq (1+\varepsilon)^n \cdot \int |\det DT|\,.$$

Mit der linken Ungleichung von Hilfssatz 1 folgt analog

$$(1+\varepsilon)^{-n} \int |\det DT| \leq \mathrm{vol}(T(Q))\,.$$

Da diese Ungleichungen für alle $\varepsilon > 0$ gelten, folgt $\mathrm{vol}(T(Q)) = \int_Q |\det DT|$.

2. Fall: Die Seitenlängen von Q sind beliebig: Da sich Q durch Quader mit rationalen Seitenlängen approximieren lässt, folgt die Behauptung aus dem 1. Fall und der Stetigkeit der beiden Seiten der Behauptung bzgl. Q. $\square$

Bemerkung: Wir haben verwendet: N Nullmenge $\Rightarrow$ $T(N)$ Nullmenge. (Beweis: Übung)

Beweis (der Transformationsformel): Wir schreiben die Behauptung $\int_V f = \int_U (f \circ T)|\det DT|$ zunächst um: Mit $g = f \circ T$ ist dies äquivalent zu

$$\int_V g \circ T^{-1} \stackrel{!}{=} \int_U g\,|\det DT| \tag{TF}$$

Wir wissen bereits, dass (TF) für $g = \chi_Q$, $Q \subset U$ kompakter Quader, gilt. Denn wegen $\chi_Q \circ T^{-1} = \chi_{T(Q)}$ ist das die Aussage von Hilfssatz 2.
Daraus folgt: (TF) gilt für Treppenfunktionen (da (TF) linear bezüglich g ist).

Zur Vollendung des Beweises brauchen wir also nur noch zu zeigen, dass bei der Approximation allgemeiner Funktionen durch Treppen alles glatt geht.

Ein kleines technisches Problem ist: $|\det DT_{|x}|$ kann beliebig groß werden, wenn $x \to \partial U$. Daher beschränken wir uns zunächst auf ein Kompaktum in U:

1. Schritt:
Sei $K \subset U$ kompakt. $g \in L^1(U)$ mit $g = 0$ außerhalb K. (Man sagt, g habe **kompakten Träger**.)
Schreibe $h(x) = |\det DT_{|x}|$. h stetig $\Rightarrow$ es existiert $C > 0$ mit $h(x) \le C$ für alle $x \in K$.
Seien φ_k Treppenfunktionen, $\varphi_k \to g$ bzgl. $\|\ \|_1$. Dann folgt $\varphi_k \cdot h \to g \cdot h$, denn $\|\varphi_k \cdot h - g \cdot h\|_1 = \|(\varphi_k - g) \cdot h\|_1 \le C\|\varphi_k - g\|_1$. Da $\int$ stetig bezüglich der Halbnorm $\|\cdot\|_1$ ist, folgt

$$\int \varphi_k \cdot |\det DT| \xrightarrow[k\to\infty]{} \int g \cdot |\det DT|\,.$$

Wir brauchen noch: $\int \varphi_k \circ T^{-1} \stackrel{!}{\to} \int g \circ T^{-1}$.
Kleines Problem: $\varphi_k \circ T^{-1}$ sind keine Treppen. Wir verwenden direkt die Definition der $\mathcal{L}^1$-Halbnorm:

$$\|\varphi_k - g\|_1 < \varepsilon$$

bedeutet: $\exists$ Hüllfunktion

$$H = \sum_{i=1}^{\infty} c_i \cdot \chi_{Q_i}\,,\quad c_i \ge 0\,,\ \text{ mit }\ I(H) < 2\varepsilon\ \text{ und}$$
$$|\varphi_k - g| \le H\,.$$

Daraus folgt: $|\varphi_k \circ T^{-1} - g \circ T^{-1}| \le H \circ T^{-1}$. Also

$$\begin{aligned}
\|\varphi_k \circ T^{-1} - g \circ T^{-1}\|_1 &\le \|H \circ T^{-1}\|_1 \\
&\le \sum_i c_i \underbrace{\|\chi_{Q_i} \circ T^{-1}\|_1}_{=\|\chi_{T(Q_i)}\|_1 = \mathrm{vol}(T(Q_i))} \ \text{(Lemma 1.2.4(e))} \\
&\le C \sum_i c_i \,\mathrm{vol}(Q_i)\ \ \text{(Hilfssatz 2)} \\
&= C \cdot I(H) \le 2C\varepsilon
\end{aligned}$$

Aus $\varphi_k \to g$ folgt also $\varphi_k \circ T^{-1} \to g \circ T^{-1}$, jeweils bzgl. $\|\ \|_1$, und daher $\int \varphi_k \circ T^{-1} \to \int g \circ T^{-1}$. Damit ist die Transformationsformel für g mit kompaktem Träger bewiesen.

2. Schritt: g allgemein.
Wähle eine kompakte Ausschöpfung von U, d. h. kompakte

$$K_1 \subset K_2 \subset \ldots \subset U \quad \text{mit} \quad \bigcup_i K_i = U\,.$$

Dann bilden die $T(K_i)$ eine kompakte Ausschöpfung von $V = T(U)$. Sei $g \circ T^{-1} \in \mathcal{L}^1(U)$. Nach dem 1. Schritt ist

$$\int_{K_i} g\,|\det DT| = \int_{T(K_i)} g\circ T^{-1}$$

für jedes i, und für $i\to\infty$ konvergiert die rechte Seite gegen $\int_V g\circ T^{-1}$. Damit ist die Folge der Integrale links beschränkt (mit demselben Argument auch für $|g|$ statt g), also ist nach dem Satz über die Integration durch Ausschöpfung die Funktion $g\,|\det DT|$ über U integrierbar, und es gilt die Transformationsformel.

Die Umkehrung (falls $g\,|\det DT|$ integrierbar ist, so auch $g\circ T^{-1}$) begründet man genauso. □

2.4 Weitere Beispiele zur Integralberechnung

Beispiele:

(1) Sei $D := \{(x,y) : x^2+y^2 \le 4\}$. Wir berechnen $\displaystyle\int_D \operatorname{sgn}(x^2-y^2+2)\,dx\,dy$.
Der Integrand und die Menge D sind symmetrisch bezüglich der Koord.-achsen, somit lässt sich das Integral in 4 gleiche Teile zerlegen. Mit $D' := \{(x,y)\in D : x\ge 0\,,\ y\ge 0\}$ ist

$$\int_D \operatorname{sgn}(x^2-y^2+2)\,dx\,dy = 4\int_{D'} \operatorname{sgn}(x^2-y^2+2)\,dx\,dy\,.$$

Weiter zerlegen wir D' in $D'_1 = \{(x,y)\in D' : x^2-y^2+2<0\}$, $D'_2 = \{(x,y)\in D' : x^2-y^2+2>0\}$ und die Nullmenge $\{(x,y)\in D' : (x^2-y^2+2)=0\}$. Für $(x,y)\in D'_1$ ist $x^2+2<y^2\le 4-x^2$, also $x\in[0,1)$, $y\in(\sqrt{x^2+2},\sqrt{4-x^2}]$, und für $(x,y)\in D'_2$ ist $y^2<2+x^2$, $y^2\le 4-x^2$, also $x\in[0,1)$, $y\in[0,\sqrt{2+x^2})$ oder $x\in[1,2]$, $y\in[0,\sqrt{4-x^2}]$, also folgt

$$\begin{aligned}\int_{D'} \operatorname{sgn}(x^2-y^2+2)\,dx\,dy &= \int_{D'_1}(-1)\,dx\,dy + \int_{D'_2} 1\,dx\,dy\\ &= -\int_0^1\int_{\sqrt{x^2+2}}^{\sqrt{4-x^2}} dy\,dx + \int_0^1\int_0^{\sqrt{x^2+2}} dy\,dx + \int_1^2\int_0^{\sqrt{4-x^2}} dy\,dx\\ &= \int_0^1 (2\sqrt{x^2+2}-\sqrt{4-x^2})\,dx + \int_1^2 \sqrt{4-x^2}\,dx\end{aligned}$$

Das könnte man zwar ausrechnen, wir lassen das aber mal...

(2) Dieses Beispiel zeigt, dass die Voraussetzung im Satz von Fubini, dass f über U integrierbar sein soll, wichtig ist: Sonst kann es vorkommen, dass beide zweifachen Integrale existieren, ohne gleich zu sein.
Sei $D := \{(x,y)\in\mathbb{R}^2 : x\ge 1\,,\ y\ge 1\}$.
Dann existieren die zweifachen Integrale:

$$I_1 = \int_1^\infty \Big(\int_1^\infty \frac{x^2-y^2}{(x^2+y^2)^2}\,dx\Big)\,dy\,,\quad I_2 = \int_1^\infty \Big(\int_1^\infty \frac{x^2-y^2}{(x^2+y^2)^2}\,dy\Big)\,dx$$

denn

$$I_1 = \int_1^\infty \Big(\int_1^\infty \frac{x^2-y^2}{(x^2+y^2)^2}\,dx\Big)\,dy = \int_1^\infty -\frac{x}{x^2+y^2}\Big|_{x=1}^\infty dy = \int_1^\infty \frac{1}{1+y^2}\,dy = \arctan y\Big|_1^\infty = \frac{\pi}{2}-\frac{\pi}{4} = \frac{\pi}{4}$$

und

$$I_2 = \int_1^\infty \Big(\int_1^\infty \frac{x^2-y^2}{(x^2+y^2)^2}\,dy\Big)\,dx = \int_1^\infty \frac{y}{x^2+y^2}\Big|_{y=1}^\infty dx = \int_1^\infty -\frac{1}{1+x^2}\,dx = -\arctan x\Big|_1^\infty = -\frac{\pi}{2}+\frac{\pi}{4} = -\frac{\pi}{4}$$

Also ist $I_1\ne I_2$. Daher kann das Integral $\displaystyle\int_D \frac{x^2-y^2}{(x^2+y^2)^2}\,dx\,dy$ nicht existieren (sonst wäre ja $I_1=I_2$ nach Fubini). Dies kann man auch direkt sehen:

Sei $D_n = [1,n] \times [1,n]$, damit bildet $\{D_n\}_{n\in\mathbb{N}}$ eine Ausschöpfung von D, und

$$\int_{D_n} \frac{|x^2-y^2|}{(x^2+y^2)^2}\,dx\,dy = \int_1^n \Big(\int_1^n \frac{|x^2-y^2|}{(x^2+y^2)^2}\,dx\Big)\,dy = \int_1^n \Big(\underbrace{\int_1^y \frac{x^2-y^2}{(x^2+y^2)^2}\,dx}_{x^2-y^2>0 \text{ auf } (1,y)} + \underbrace{\int_y^n -\frac{x^2-y^2}{(x^2+y^2)^2}\,dx}_{x^2-y^2<0 \text{ auf } (y,n)}\Big)\,dy$$

$$= -\int_1^n \left(\frac{x}{x^2+y^2}\Big|_1^y - \frac{x}{x^2+y^2}\Big|_y^n\right)\,dy = -\int_1^n \left(\frac{1}{2y} - \frac{1}{1+y^2} - \frac{n}{n^2+y^2} + \frac{1}{2y}\right)\,dy$$

$$= -\int_1^n \left(\frac{1}{y} - \frac{1}{1+y^2} - \frac{n}{n^2+y^2}\right)\,dy = -(\ln y - \arctan y - \arctan\frac{y}{n})\Big|_1^n = \underbrace{-\ln n + \arctan n - \arctan\frac{1}{n}}_{=:I_n}$$

und $\lim\limits_{n\to\infty} I_n = -\infty$

Somit folgt mit der Negation des Satzes 1.5.2 (Ausschöpfung), dass das Integral $\int_D \frac{|x^2-y^2|}{(x^2+y^2)^2}\,dx\,dy$ nicht existiert und somit auch $\int_D \frac{x^2-y^2}{(x^2+y^2)^2}\,dx\,dy$ nicht existiert.

Es gibt auch Beispiele, wo beide iterierten Integrale existieren und gleich sind, und trotzdem f nicht über D integrierbar ist. Basteln Sie sich eins!

(3) Sei $D = [0,1]$, $I_n = \left(\frac{1}{n+1}, \frac{1}{n}\right]$ und $f(x) = \begin{cases} -n & \text{wenn} \quad \frac{2n+1}{2n(n+1)} < x \le \frac{1}{n} \\ n & \text{wenn} \quad \frac{1}{1+n} < x \le \frac{2n+1}{2n(n+1)} \end{cases}$

Da $\frac{2n+1}{2n(n+1)} = \frac{1}{2}\left(\frac{1}{n} - \frac{1}{n+1}\right)$ der Mittelpunkt des Intervalls $\left[\frac{1}{n+1}, \frac{1}{n}\right]$ ist, ist $\int_{I_n} f(x)\,dx = 0$ für jedes $n \in \mathbb{N}$.

Allerdings existiert $\int_D f(x)\,dx$ nicht: Es ist $D = \cup_{n\in\mathbb{N}} I_n$.

Angenommen $\int_D f(x)\,dx$ existiert, so existiert auch $\int_D |f(x)|\,dx$. Aber

$$\begin{aligned}\int_D |f(x)|\,dx &= \int_{\cup_{n\in\mathbb{N}} I_n} |f(x)|\,dx \\ &= \sum_{n\in\mathbb{N}} \int_{I_n} |f(x)|\,dx = \sum_{n\in\mathbb{N}} \int_{\frac{1}{n+1}}^{\frac{1}{n}} n\;dx \\ &= \sum_{n\in\mathbb{N}} n\left(\frac{1}{n} - \frac{1}{n+1}\right) = \sum_{n\in\mathbb{N}} \left(\frac{1}{n+1}\right) = \infty\end{aligned}$$

(4) Sei $m \in \mathbb{R}$, $f(x) = \frac{1}{(x^2+y^2)^m}$ Es gilt:

(a) $\int_{x^2+y^2\le 1} f(x,y)\,dx\,dy$ ist integrierbar $\Leftrightarrow m < 1$.

(b) $\int_{x^2+y^2\ge 1} f(x,y)\,dx\,dy$ ist integrierbar $\Leftrightarrow m > 1$.

(c) $\int_{\mathbb{R}^2} f(x,y)\,dx\,dy$ existiert für kein m.

Beweis:

(a) $\int_{x^2+y^2\le 1} f(x,y)\,dx\,dy \overset{\text{Polarkoord.}}{=} \int_0^{2\pi}\int_0^1 \frac{1}{r^{2m}} r\;dr\,d\varphi = 2\pi\int_0^1 \frac{1}{r^{2m-1}}\,dr$

Dies ist genau für $2m-1 < 1 \Leftrightarrow m < 1$ integrierbar.

(b) $\int_{x^2+y^2\geq 1} f(x)\,dx\,dy \overset{\text{Polarkoord.}}{=} \int_0^{2\pi}\int_1^{\infty} \frac{1}{r^{2m}} r\;dr\,d\varphi = 2\pi\int_1^{\infty}\frac{1}{r^{2m-1}}\,dr$

Dies ist genau für $2m-1>1 \Leftrightarrow m>1$ integrierbar.

(c) Existierte es, so müssten auch $\int_{x^2+y^2\leq 1} f$ und $\int_{x^2+y^2\geq 1} f$ existieren, also müsste gleichzeitig $m>1$ und $m<1$ sein, was nicht geht.

Oder z. B. mit folgender Rechnung: Sei $A_n := \left\{(x,y)\in\mathbb{R}^2 : \frac{1}{n^2}\leq x^2+y^2\leq n^2\right\}$, damit ist A_n eine Ausschöpfung des $\mathbb{R}^2$

$$\begin{aligned}
\int_{A_n} f(x,y)\,dx\,dy &\overset{\text{Polarkoord.}}{=} \int_0^{2\pi}\int_{\frac{1}{n}}^{n}\frac{1}{r^{2m}} r\;dr\,d\varphi \\
&= 2\pi\int_{\frac{1}{n}}^{n}\frac{1}{r^{2m-1}}\,dr \\
&= \begin{cases} 2\pi\cdot\ln r\Big|_{\frac{1}{n}}^{n} & \text{falls } m=1 \\ \frac{2\pi}{-2m+2}\cdot\frac{1}{r^{2m-2}}\Big|_{\frac{1}{n}}^{n} & \text{sonst} \end{cases} \\
&= \begin{cases} 4\pi\cdot\ln n & \text{falls } m=1 \\ \frac{\pi}{1-m}\left(\frac{1}{n^{2(m-1)}}-n^{2(m-1)}\right) & \text{sonst} \end{cases}
\end{aligned}$$

Man sieht $\lim\limits_{n\to\infty}\int_{A_n} f(x,y)\,dx\,dy = \infty$, und somit existiert dies Integral nicht.

3 Untermannigfaltigkeiten des $\mathbb{R}^N$ und deren Volumen

3.1 Untermannigfaltigkeiten des $\mathbb{R}^N$

Untermannigfaltigkeiten des $\mathbb{R}^N$, im folgenden oft kurz Mannigfaltigkeiten genannt, sind die Verallgemeinerungen, auf beliebige Dimensionen, von Kurven (eindimensional) und Flächen (zweidimensional) im Raum. Diese sollen keine »Ecken« besitzen, sondern »glatt« sein.

Vorüberlegung: Wie kann man das mathematisch präzise formulieren?
Wir betrachten hierzu zunächst den Fall von Kurven, den wir schon aus Analysis II kennen.

Erinnerung: Kurven.
Sei $I \subset \mathbb{R}$ ein Intervall. Eine stetige Abbildung $\gamma : I \to \mathbb{R}^N$ heißt (parametrisierte) Kurve.
Das Bild $M = \gamma(I)$ heißt unparametrisierte Kurve.
Frage: Welche Bedingung müssen wir an γ stellen, damit M keine ›Ecken ‹ hat?
Offenbar sollten wir zumindest fordern, dass γ differenzierbar ist.
Reicht das?

Beispiel: Sei $\gamma(t) = (t^3, t^2)$ für $t \in \mathbb{R}$. Dies ist differenzierbar. Das Bild von γ ist die Neilsche Parabel $\gamma(\mathbb{R}) = \{(x, y) : x^2 = y^3\}$, diese ist bei $(0, 0) = \gamma(0)$ nicht glatt (hat dort eine Spitze).
(Der Begriff glatt wird später exakt definiert, ist hier intuitiv zu verstehen.)

Das Beispiel zeigt: Wenn wir uns nur für die Kurve *als Teilmenge des* $\mathbb{R}^2$, also die unparametrisierte Kurve interessieren, können wir anhand der Differenzierbarkeit einer Parametrisierung nicht ohne weiteres ablesen, ob M glatt ist oder nicht.

Wir sahen aber schon in Analysis II: Falls $\gamma'(t) \neq 0$ für alle t gilt, so ist das Bild von γ in dem Sinne glatt, dass es sich lokal als Graph einer differenzierbaren Funktion darstellen lässt.

Zusammenfassend:

- ▷ Wir beschreiben Kurven mit Hilfe von Parametrisierungen.
- ▷ Die Glattheit der Parametrisierung impliziert *nicht* die Glattheit der Kurve.
- ▷ Die zusätzliche Bedingung $\gamma'(t) \neq 0 \, \forall t$ impliziert Glattheit lokal.

Für Untermannigfaltigkeiten beliebiger Dimension n gehen wir analog vor: Wir beschreiben sie durch Parametrisierungen (sogenannten lokale Karten). Wir klären zunächst, was die korrekte Verallgemeinerung der Bedingung $\gamma'(t) \neq 0 \, \forall t$ ist.

3.1.1 Definition

Seien $n, N \in \mathbb{N}_0$. Sei $U \subset \mathbb{R}^n$ offen und $\varphi : U \to \mathbb{R}^N$ eine C^1-Abbildung.
φ heißt **Immersion** $:\Leftrightarrow$ $D\varphi_{|x}$ hat den Rang n für jedes $x \in U$.

Im Fall einer Kurve ($n = 1$) gilt:
φ ist Immersion $\Leftrightarrow \frac{\partial \varphi}{\partial x}(x) \neq 0$ für alle x, also $\varphi'(x) \neq 0$ in üblicher Notation.

Im Folgenden identifizieren wir meist lineare Abbildungen mit den sie repräsentierenden Matrizen. Damit ist

$$D\varphi_{|x} = (\frac{\partial\varphi}{\partial x_1}(x), \ldots, \frac{\partial\varphi}{\partial x_n}(x))$$

d.h. die Vektoren $\frac{\partial\varphi}{\partial x_i}(x) \in \mathbb{R}^N$ sind die Spalten von $D\varphi_{|x}$. Damit ergibt sich aus der linearen Algebra, dass die folgenden Bedingungen äquivalent sind:

(1) $D\varphi_{|x}$ hat Rang n.

(2) $\frac{\partial\varphi}{\partial x_1}(x), \ldots, \frac{\partial\varphi}{\partial x_n}(x)$ sind linear unabhängig.

(3) $D\varphi_{|x} : \mathbb{R}^n \to \mathbb{R}^N$ ist injektiv.

In diesem Fall muss $N \geq n$ sein.

Bemerkung: Die Vektoren $\frac{\partial\varphi}{\partial x_i}(x)$ haben folgende geometrische Bedeutung: Zeichne durch x eine Gerade, die parallel zur x_i-Achse verläuft. Deren Bild unter φ ist eine Kurve in $\mathbb{R}^N$, deren Tangentialvektor bei $\varphi(x)$ genau in Richtung $\frac{\partial\varphi}{\partial x_i}(x)$ zeigt. Die Länge von $\frac{\partial\varphi}{\partial x_i}(x)$ ist dadurch bestimmt, wie schnell sich φ ändert, wenn sich x mit Geschwindigkeit Eins entlang der Geraden bewegt.

Beispiele:

(1) $\varphi(t) = (\cos t, \sin t)$ ist eine Immersion $\mathbb{R} \to \mathbb{R}^2$, deren Bild ein Kreis ist.

(2) Durchläuft man eine Acht, ohne anzuhalten, definiert dies eine Immersion. (Es gibt hierfür explizite Formeln. Sehen Sie mal in einem Buch oder im Internet unter dem Stichwort Lemniskate nach.)

Beobachtung: Diese Immersionen sind nicht injektiv, aber sie sind lokal injektiv (d. h. jedes t hat eine Umgebung U, so dass $\varphi_{|U}$ injektiv ist). Dies gilt allgemein, wie wir sehen werden.

Folgender Begriff wird uns helfen, einen wichtigen qualitativen Unterschied zwischen der Acht und dem Kreis zu beschreiben.

3.1.2 Definition

Seien $(X, d_x), (Y, d_y)$ metrische Räume.

$f : X \to Y$ heißt **Homöomorphismus** $:\Leftrightarrow$ f stetig, bijektiv und f^{-1} stetig.

Beispiele:

(1) $\varphi : (0,1) \to (1,\infty)$, $x \mapsto \frac{1}{x}$ ist ein Homöomorphismus.

(2) $\varphi : [0, 2\pi) \to \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$, $t \mapsto (\cos t, \sin t)$ ist kein Homöomorphismus; es ist zwar bijektiv und stetig, aber φ^{-1} ist unstetig bei $(1,0)$.

Erinnerung:
Jedes $M \subset \mathbb{R}^N$ ist metrischer Raum, d. h. die Begriffe »offene Teilmenge von M« und »abgeschlossene Teilmenge von M« sind definiert.

Es gilt: $U \subset M$ offen in $M \Leftrightarrow$ Es existiert $U' \subset \mathbb{R}^N$ offen mit $U = U' \cap M$.
Zum Beispiel ist $\{(x,0) : 0 < x < 1\}$ zwar keine offene Menge von $\mathbb{R}^2$, aber als Teilmenge der x-Achse ist es offen.

3.1.3 Definition

Seien $n, N \in \mathbb{N}_0$. Eine Teilmenge $M \subset \mathbb{R}^N$ heißt **(Unter-) Mannigfaltigkeit** (kurz: (U)Mgfk) des $\mathbb{R}^N$ der Dimension n, falls gilt:
Für jedes $p \in M$ existieren eine offene Umgebung V von p in M, eine offene Menge $U \subset \mathbb{R}^n$ und eine Abbildung $\varphi : U \to \mathbb{R}^N$ mit:

(1) φ ist Immersion.

(2) $\varphi(U) = V$ und $\varphi : U \to V$ ist Homöomorphismus.

Solche φ heißen **lokale Karten** (oder Parametrisierungen) von M bei p.

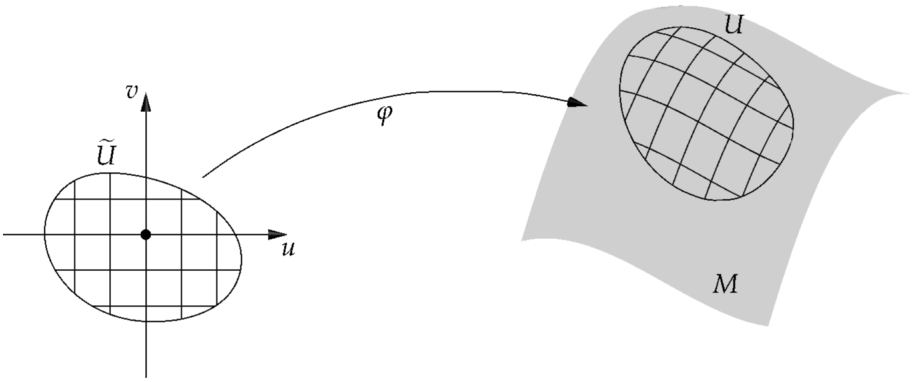


Abbildung 3.1. Lokale Karte einer 2-dimensionalen Untermannigfaltigkeit des $\mathbb{R}^3$

Man sagt auch **differenzierbare Mannigfaltigkeit**. Gibt es zu jedem Punkt eine lokale Karte, die eine C^k-Abbildung ist (k-mal stetig differenzierbar), so spricht man von einer $\boldsymbol{C^k}$**-Mannigfaltigkeit** (hierbei ist $k \in \mathbb{N} \cup \{\infty\}$. Den Fall $k = 0$ – stetige Karten, ohne die Immersionsbedingung, die dann keinen Sinn hat – betrachten wir hier nicht (sogenannte topologische Mannigfaltigkeiten)).

Bemerkung: Neben den Untermannigfaltigkeiten eines $\mathbb{R}^N$ gibt es noch den allgemeineren Begriff der (abstrakten) Mannigfaltigkeit, die nicht in irgendeinem $\mathbb{R}^N$ zu liegen braucht (man sagt auch: eingebettet ist). Dieser ist zum Beispiel von Bedeutung in folgenden Kontexten:

- ▷ Kosmologie: Nach gängiger relativistischer Theorie ist die Welt als 4-dimensionale Mannigfaltigkeit (die »Raum-Zeit«) zu beschreiben. Es macht wenig Sinn, sich diese als in einem höherdimensionalen Raum liegend vorzustellen. Es ist übrigens ein offenes Problem, wie die Raum-Zeit aussieht, d. h. welche Mannigfaltigkeit das sein könnte.
- ▷ Quotienten: Ähnlich wie bei Quotienten einer Gruppe nach einem Normalteiler ist es bei Betrachtung von Problemen mit Symmetrien oft nützlich, den Quotienten einer Mannigfaltigkeit bzgl. einer Gruppenoperation zu betrachten. Selbst wenn die ursprüngliche Mannigfaltigkeit eingebettet ist, liegt der Quotient nicht natürlich in einem $\mathbb{R}^N$. Beispiel $\mathbb{R}^2/\mathbb{Z}^2$ (als Gruppenquotient) ist eine Mannigfaltigkeit (topologisch dasselbe wie ein Torus).
- ▷ Eine andere wichtige Operation auf Mannigfaltigkeiten ist das Verkleben. Auch das führt einen oft aus dem eingebetteten Kontext heraus.

Da wir aber in dieser Vorlesung nur Untermannigfaltigkeiten des $\mathbb{R}^N$ betrachten (um die Abstraktion in Grenzen zu halten), verwenden wir dafür etwas unexakt auch den Begriff Mannigfaltigkeit.

Beispiele:

(1) Der Kreis $K = \{(x,y);\, x^2+y^2=1\} \subset \mathbb{R}^2$ ist Mgfk mit dim $= 1$. Als lokale Karten kann man die Immersion $\varphi(t) = (\cos t, \sin t)$ verwenden, eingeschränkt auf beliebige offene Intervalle der Länge $\leq 2\pi$ (sogenannte Winkel-Parametrisierung). Z.B. $(0, 2\pi) \to V := K \setminus \{(1,0)\}$. V ist offen in K, da $\{(1,0)\}$ abgeschlossen in K ist (einpunktige Mengen sind in jedem metrischen Raum abgeschlossen).

Analog $(-\pi, \pi) \to K \setminus \{(-1,0)\}$. Diese beiden Karten reichen aus, um ganz K abzudecken, d.h. die Anforderungen von Definition 3.1.3 für jedes p zu erfüllen.

Dass φ eingeschränkt auf diese Intervalle ein Homöomorphismus ist, sieht man so: Es ist offenbar stetig (nach Formel) und bijektiv (Eindeutigkeit des Winkels, vgl. Betrachtung in Analysis I). Dass die inverse Abbildung stetig ist, sieht man an der expliziten Formel (mit Fallunterscheidungen je nach Quadrant, z.B. $t = \arctan \frac{y}{x}$ für $x \neq 0$).

(2) Das Bild einer Acht in $\subset \mathbb{R}^2$ ist *keine* Mgfk, denn für den Kreuzungspunkt p gibt es keine lokale Karte.
Das ist anschaulich klar, da jede Umgebung von p in M zwei sich kreuzende Linien enthält, also nicht homöomorph zu einer Teilmenge von $\mathbb{R}$ sein kann.

(Dies kann man am besten mit dem später einzuführenden Begriff der Zusammenhangskomponenten präzisieren: Jede Umgebung V von p zerfällt bei wegnehmen von p in mindestens vier Teile (Zusammenhangskomponenten). Falls $\varphi : U \to \varphi(U) = V$ eine lokale Karte wäre, so wäre es ein Homöomorphismus, daher müsste auch gelten: U zerfällt bei Wegnehmen des Punktes $x = \varphi^{-1}(p)$ in ≥ 4 Teile; da man o. B. d. A. U als zusammenhängend annehmen kann (indem man ggf. U verkleinert, genauer durch die Zusammenhangskomponente von U ersetzt, die x enthält), also als offenes Intervall, zerfällt es aber nur in 2 Teile. Somit kann φ nicht existieren.)

(3) Die Sphäre $S^2 = \{(x,y,z) \in \mathbb{R}^3 :\ x^2+y^2+z^2 = 1\}$ ist eine 2-dimensionale UMgfk. des $\mathbb{R}^3$. Als Karten kann man die beiden stereographischen Projektionen

$$\mathbb{R}^2 \to S^2 \setminus \{N\}, \quad \mathbb{R}^2 \to S^2 \setminus \{S\}$$

(N, S = Nord- bzw. Südpol) verwenden (Übung). Es gibt aber auch andere Karten, z. B. $\varphi(x,y) = (x, y, \sqrt{1-x^2-y^2})$, definiert auf der offenen Kreisscheibe $U = \{(x,y) :\ x^2+y^2 < 1\}$. Dies ist eine Karte für alle Punkte in der offenen oberen Halbsphäre. Ähnliche Karten findet man für die untere, linke, rechte, vordere und hintere Halbsphäre, und das deckt alle Punkte ab.

Ein paar naheliegende Begriffe:

3.1.4 Definition

Ist $\varphi : U \to V \subset M$ eine lokale Karte für M, dann heißt V **Kartengebiet**. Falls $\{V_i\}_{i\in I}$ eine Menge von Kartengebieten ist, die M überdecken, d. h. $M = \bigcup_{i\in I} V_i$ mit Karten $\varphi_i : U_i \to V_i$, dann heißt die Menge $\{(\varphi_i, U_i, V_i)\}_{i\in I}$ ein **Atlas** für M.

Offenbar reicht es, einen Atlas für M anzugeben, um zu zeigen, dass M eine Mgfk. ist. Andererseits gibt es dann unendlich viele verschiedene Atlanten (Beispiel Sphäre oben).

Eine fundamentale Tatsache ist, dass man Mannigfaltigkeiten auch anders beschreiben kann. Wir führen hierfür folgende Sprechweisen ein:

3.1.5 Definition

Eine Teilmenge $V \subset \mathbb{R}^N$ heißt **Graph der Dimension n**, falls V Graph einer C^1-Abbildung von n Variablen mit $N-n$ Komponenten ist.

Genauer muss gelten: Es existiert eine offene Menge $U \subset \mathbb{R}^n$, eine Permutation π von $\{1,\dots,N\}$ und eine C^1-Abbildung $g : U \to \mathbb{R}^{N-n}$, so dass gilt:

$$V = \{(x_{\pi(1)},\dots,x_{\pi(N)}) : (x_1,\dots,x_n) \in U,\ (x_{n+1},\dots,x_N) = g(x_1,\dots,x_n)\}$$

Beispiel: Die obere Halbsphäre ist $V_o = \{(x,y,\sqrt{1-x^2-y^2}) : (x,y) \in U_o\}$ mit $U_o = \{(x,y) : x^2+y^2 < 1\}$, also der Graph der Funktion $\sqrt{1-x^2-y^2}$ auf U_o (hier ist keine Permutation nötig).
Die rechte Halbsphäre ist $V_r = \{(\sqrt{1-y^2-z^2},y,z) : (y,z) \in U_r\}$ mit $U_r = U_o$, mit derselben Funktion, aber der Permutation $1 \mapsto 2 \mapsto 3 \mapsto 1$ (oder umgekehrt? Egal, jedenfalls eine Permutation).

3.1.6 Definition

Eine Teilmenge $V \subset \mathbb{R}^N$ heißt **Lösungsmenge eines Systems von k unabhängigen Gleichungen**, falls gilt:
Es existiert eine offene Umgebung W von V in $\mathbb{R}^N$ und eine C^1-Abbildung $f : W \to \mathbb{R}^k$, $f = (f_1,\dots,f_k)$, so dass gilt

(1) $V = f^{-1}(0)$.

(2) $\nabla f_1(p),\dots,\nabla f_k(p)$ sind linear unabhängig für jedes $p \in V$.

Umformulierung von (1): (1) kann man auch als

$$V = \{q \in W : f_1(q) = 0,\dots,f_k(q) = 0\}$$

schreiben, daher der Name. (Statt Null können rechts beliebige andere Zahlen stehen, denn man kann diese ja auf die linke Seite der Gleichungen bringen. Man ersetzt also f durch $f-c$ für ein konstantes $c \in \mathbb{R}^k$. Gilt Bedingung (2) für f, so auch für $f-c$.)
Umformulierung von (2): Indem man

$$Df_{|p} = \begin{pmatrix} \nabla f_1(p) \\ \vdots \\ \nabla f_k(p) \end{pmatrix}$$

schreibt (wobei die Gradienten als Zeilenvektoren geschrieben sind), sieht man, wieder mit etwas linearer Algebra, dass folgende Bedingungen äquivalent sind:

(a) $Df_{|p}$ hat Rang k.

(b) $\nabla f_1(p),\dots,\nabla f_k(p)$ sind linear unabhängig.

(c) $Df_{|p} : \mathbb{R}^N \to \mathbb{R}^k$ ist surjektiv.

In diesem Fall muss $N \geq k$ sein. Eine Abbildung, die eine (und damit jede) dieser Bedingungen erfüllt, heißt **Submersion**. Vergleiche mit der Bemerkung nach Definition 3.1.1.

Beispiel: Nochmal die Sphäre S^2. Diese ist Lösungsmenge einer Gleichung:

$$S^2 = f^{-1}(0) \quad \text{mit } f(x,y,z) = x^2+y^2+z^2-1$$

Es gilt $\nabla f(x,y,z) = (2x,2y,2z)$, und dies ist ungleich null, also unabhängig, für $(x,y,z) \in S^2$.

Es gilt nun:

3.1.7 Satz

Sei $M \subset \mathbb{R}^N$ und $n \in \mathbb{N}_0$. Die folgenden Bedingungen sind äquivalent.

(1) M ist Untermannigfaltigkeit des $\mathbb{R}^N$ der Dimension n, d. h. M kann durch lokale Karten überdeckt werden.

(2) M ist überall lokal Graph der Dimension n.

(3) M ist überall lokal Lösungsmenge eines Systems von $N-n$ unabhängigen Gleichungen.

Hierbei bedeutet »überall lokal« in (2): Für jedes $p \in M$ existiert eine offene Umgebung V von p in M, die Graph der Dimension n ist. Analog für (3).

Beweis:

1. Die Äquivalenz (2) $\Leftrightarrow$ (3) hatten wir in Analysis II als unmittelbare Konsequenz des Satzes über implizite Funktionen bewiesen.

2. (2) $\Rightarrow$ (1): Wir nehmen o. B. d. A. an, dass die Permutation in der Definition von »Graph« die Identität ist. Schreibe $x = (x_1, \ldots, x_n)$. Ist die offene Umgebung V von p in M als Graph von $g : U \to \mathbb{R}^{N-n}$ gegeben, so ist $V = \{(x, g(x)) : x \in U\}$, und dies ist das Bild von $\varphi(x) = (x, g(x))$.

   Behauptung: φ ist eine Immersion.

   Beweis: Als Spaltenvektor ist $\varphi(x) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \\ g(x) \end{pmatrix}$, also $D\varphi = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \\ \frac{\partial g_1}{\partial x_1} & & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & & \cdots & \vdots \\ \frac{\partial g_{N-n}}{\partial x_1} & & \cdots & \frac{\partial g_{N-n}}{\partial x_n} \end{pmatrix}$.

   Da die ersten n Zeilen linear unabhängig sind, ist der Rang gleich n, also ist φ Immersion.

   Behauptung: $\varphi : U \to V$ ist Homöomorphismus.

   Beweis: φ ist stetig, da g stetig ist. Offenbar ist φ bijektiv. Das Inverse von φ ist durch $(y_1, \ldots, y_n, y_{n+1}, \ldots, y_N) \mapsto (y_1, \ldots, y_n)$ gegeben, also stetig.

3. (1) $\Rightarrow$ (2): Wir geben nur den Beweis für $n = 2$, $N = 3$, der allgemeine Fall ist ganz analog. Wir bezeichnen Punkte im Definitionsbereich von φ mit x, Punkte im Bildbereich mit y.

   (i) Wir betrachten zuerst den linearen Fall, d. h. $\varphi = A : \mathbb{R}^2 \to \mathbb{R}^3$ ist eine lineare Abbildung. Da A injektiv ist, ist das Bild V von A ein zweidimensionaler Unterraum. Was ist die Bedingung dafür, dass V ein Graph über der $y_1 - y_2$-Ebene ist?

   Sei $\pi(y_1, y_2, y_3) = (y_1, y_2)$ die vertikale Projektion, dann lautet die Bedingung: $p \circ A : \mathbb{R}^2 \to \mathbb{R}^2$ muss bijektiv sein (denn zu jedem $x \in \mathbb{R}^2$ muss es ja genau ein $p \in V$ geben mit $x = \pi(p)$).

   Falls $A = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$ ist (mit Zeilenvektoren $a, b, c \in \mathbb{R}^n$), so hat $p \circ A$ die Matrix $\begin{pmatrix} a \\ b \end{pmatrix}$.

   Resultat: Bild(A) ist Graph über y_1, y_2 genau dann, wenn die beiden ersten Zeilen von A linear

unabhängig sind.
Da A Rang 2 hat, sind zwei der Vektoren a, b, c linear unabhängig, also ist Bild(A) Graph über einer der Koordinatenebenen.

(ii) Der allgemeine Fall. Sei $\varphi : U \to \mathbb{R}^3$ mit $U \subset \mathbb{R}^2$ Immersion. Sei $x \in U$ derart, dass $\varphi(x) = p$. $D\varphi_{|x} : \mathbb{R}^2 \to \mathbb{R}^3$ ist linear mit Rang 2. Seien o. B. d. A. die ersten zwei Zeilen von $D\varphi_{|x}$ linear unabhängig. Schreibe $\varphi = \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \end{pmatrix} = \begin{pmatrix} \widetilde{\varphi} \\ \varphi_3 \end{pmatrix}$ mit $\widetilde{\varphi} = \begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix} : U \to \mathbb{R}^2$. Wegen der Annahme ist $D\widetilde{\varphi}_{|x}$ invertierbar.
Aus dem Satz über die Umkehrfunktion folgt: Es gibt Umgebungen $U_1 \subset U$ von x, U_2 von $\widetilde{\varphi}(x)$, so dass $\widetilde{\varphi} : U_1 \to U_2$ ein Diffeomorphismus ist. Setze $g = \varphi_3 \circ \widetilde{\varphi}^{-1} : U_2 \to \mathbb{R}$. Dann ist

$$\begin{aligned} V_1 := \varphi(U_1) &= \{\varphi(x) : \ x \in U_1\} \\ &= \left\{ \begin{pmatrix} \widetilde{\varphi}(x) \\ \varphi_3(x) \end{pmatrix} | x \in U_1 \right\} \\ &= \left\{ \begin{pmatrix} y \\ \varphi_3(\widetilde{\varphi}^{-1}(y)) \end{pmatrix} | y \in U_2 \right\} \\ &= \left\{ \begin{pmatrix} y \\ g(y) \end{pmatrix} | y \in U_2 \right\} = \text{ Graph von } g \end{aligned}$$

□

Was Sie sich von diesem Beweis mindestens merken sollten: (etwa im Fall $n = 2$, $N = 3$)

- ▷ Ist M als Lösungsmenge der Gleichung $f(x, y, z) = 0$, mit $\nabla f \neq 0$, gegeben, so erhält man eine Darstellung von M als Graph wie folgt (in der Nähe eines Punktes p):
  Wähle eine Variable, nach der die partielle Ableitung bei p nicht verschwindet; sagen wir z, also $\partial f / \partial z(p) \neq 0$.
  Dann kann man die Gleichung $f = 0$ nahe p nach z auflösen (nach dem Satz über implizite Funktionen).
  Man erhält $z = g(x, y)$, d. h., M ist nahe p der Graph von g.

- ▷ Ist M als Graph von $g = g(x, y)$ gegeben, so kann es alternativ mittels der lokalen Karte $\varphi(x, y) = (x, y, g(x, y))$ geschrieben werden.

Dies ist wichtig, denn: Mannigfaltigkeiten sind oft als Lösungsmengen von Gleichungen gegeben. Will man deren Fläche berechnen, braucht man aber eine Darstellung mittels lokaler Karten (wie wir sehen werden). Daher braucht man diese beiden Schritte.

Beispiele: Hier ein paar weitere Beispiele von Untermannigfaltigkeiten des $\mathbb{R}^N$:

(1) $n = 0$: M = Punkt oder diskrete Menge von Punkten. (Beachte: $\mathbb{R}^0 = \{0\}$.)

(2) $n = 1$: M = offene Wege, geschlossene Kurven und Vereinigungen von solchen; *nicht*: Intervall mit Endpunkten, sich selbst schneidende Kurve

(3) $n = 2$: M = Sphäre, jeder Graph einer C^1-Funktion mit zwei Variablen, Möbius-Band (ohne Randlinien), Torus, Kleinsche Flasche (die gibt es nicht im $\mathbb{R}^3$, wohl aber im $\mathbb{R}^4$).

(4) $n = N$: Die n-dimensionalen Untermannigfaltigkeiten des $\mathbb{R}^n$ sind genau die offenen Teilmengen von $\mathbb{R}^n$.

Beispiele:

(1) $M = \{(x,y,z) \in \mathbb{R}^3 : x^2+y^2 = z^2\}$ (Doppelkegel) nach der dritten Bedingung der obigen Charakterisierung: Mit $f(x,y,z) = x^2+y^2-z^2$ ist $M = f^{-1}(0)$. Es gilt $\nabla f(x,y,z) = (2x, 2y, -2z)$, dies ist $= 0 \Leftrightarrow (x,y,z) = 0$. Also ist $M \setminus \{(0,0,0)\}$ eine Mannigfaltigkeit, aber M potentiell nicht. Man kann sich überlegen, dass es bei $p = (0,0,0)$ wirklich keine lokale Karte für M gibt, also ist M keine Mannigfaltigkeit.

(2) $M = \{(x,y,z) \in \mathbb{R}^3 : x^2+y^2 = z^2,\ z \geq 0\}$ (Kegel) ist zwar Graph der stetigen Funktion $g(x,y) = \sqrt{x^2+y^2}$, aber g ist in $(0,0)$ nicht differenzierbar. Wiederum kann man sich überlegen, dass M keine differenzierbare Mannigfaltigkeit ist. Man muss aber genau argumentieren, die Nicht-Differenzierbarkeit von g reicht nicht aus, siehe Beispiel (3) unten! (Der Kegel ist aber eine topologische Mannigfaltigkeit.)

$M = \{(x,y,z) \in \mathbb{R}^3 : x^2+y^2 = z^2+a\}$, $a \neq 0$ (Hyperboloid). Nach der dritten Bedingung: mit $f(x,y,z) = x^2+y^2-z^2$ ist $M_2 = f^{-1}(a)$. Es gilt $\nabla f(x,y,z) = (2x, 2y, -2z)$, dies verschwindet nur bei $(x,y,z) = 0$. Aber der Punkt $(x,y,z) = (0,0,0)$ liegt nicht in M, da $0^2+0^2 \neq 0^2+a$ für $a \neq 0$. Also ist M eine Mannigfaltigkeit.

Hier noch ein paar **Beispiele zur Vorsicht:**

(3) $M = \{(x,y) : y^3 = x\}$. Man kann M als Graph von $y = x^{1/3}, x \in \mathbb{R}$, schreiben (wobei natürlich $x^{1/3} = -(-x)^{1/3}$ für $x < 0$ gemeint ist). Diese Funktion ist bei 0 nicht differenzierbar. Trotzdem ist M eine differenzierbare Mannigfaltigkeit, da M auch als Graph von $x = y^3$ geschrieben werden kann!

(4) $M = \{(x,y) : x^2 = 0\}$. Das ist offenbar die y-Achse, also eine 1-dimensionale Mannigfaltigkeit. Aber hier ist $f(x,y) = x^2$, und $\nabla f(x,y) = (2x, 0) = (0,0)$ für alle $(x,y) \in M$, d. h. die Unabhängigkeitsbedingung ist nicht erfüllt.

(5) $M = \{(x,y) : x^2+y^2 = 0\}$. Offenbar ist $M = \{(0,0)\}$, also eine 0-dimensionale Mannigfaltigkeit. Hier ist $f(x,y) = x^2+y^2$ und $\nabla f(x,y) = (2x, 2y) = (0,0)$ bei $(x,y) = (0,0)$.

In all diesen Beispielen war M eine Mannigfaltigkeit, obwohl die gegebene Darstellung nicht von der Art war, wie in den Definitionen gefordert.

> *Mannigfaltigkeit zu sein, ist eine Eigenschaft einer Menge, nicht einer speziellen Darstellung dieser Menge!*

Die in den Definitionen gegebenen Bedingungen an Parametrisierungen, Gleichungen etc. sind hinreichend, nicht notwendig für das Vorliegen einer Mgfk. Allerdings ist notwendig und hinreichend, dass es eine (möglicherweise andere) Parametrisierung, Gleichung etc. *gibt,* die diesen Bedingungen genügt.

Im Beispiel (4) kann man M alternativ als Nullstellenmenge der Funktion $g(x,y) = x$ beschreiben, und im Beispiel (5) ist M die Lösungsmenge des Gleichungssystems $x = 0, y = 0$ (also $g(x,y) = (x,y)$), und diese Darstellungen erfüllen die Unabhängigkeitsbedingung.

3.2 Der Tangentialraum einer Mannigfaltigkeit

Bevor wir zur Flächenberechnung kommen, müssen wir noch eine Frage klären:
Wie charakterisiert und berechnet man die Tangentialvektoren an eine Fläche? Es sind genau die Tangenten an Kurven, die in der Fläche verlaufen! Also:

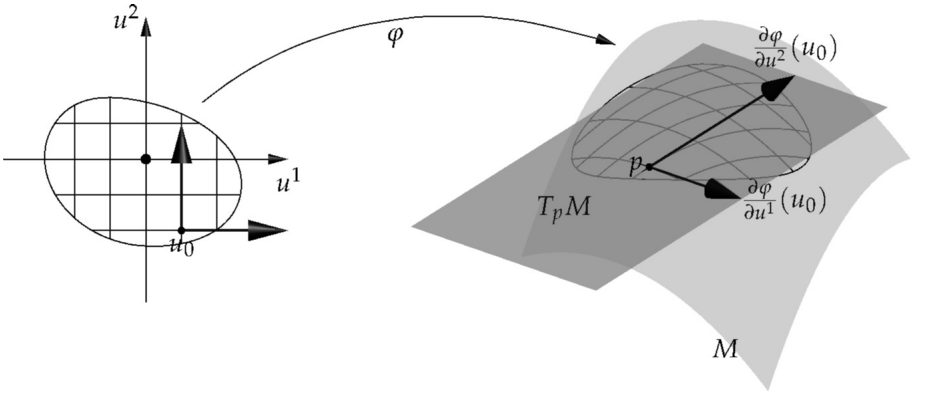


Abbildung 3.2. Tangentialraum im Punkt p einer 2-dimensionalen Untermannigfaltigkeit des $\mathbb{R}^3$: $\frac{\partial \varphi}{\partial u^1}(u_0)$ und $\frac{\partial \varphi}{\partial u^2}(u_0)$ bilden eine Basis des Tangentialraumes

3.2.1 Definition

Sei M Untermannigfaltigkeit des $\mathbb{R}^N$, $p \in M$. Der **Tangentialraum an M in p** ist definiert durch

$$T_pM = \{v \in \mathbb{R}^N : \text{Es gibt ein offenes Intervall } I \subset \mathbb{R}, 0 \in I, \text{ und eine } C^1\text{-Kurve } \gamma : I \to \mathbb{R}^N \text{ mit } \gamma(I) \subset M, \gamma(0) = p, \gamma'(0) = v\}$$

Elemente von T_pM heißen **Tangentialvektoren** an M im Punkt p.

Man stellt sich einen Tangentialvektor als an p festgeklebt vor und damit den Tangentialraum als Unterraum eines Exemplars von $\mathbb{R}^N$, dessen Ursprung bei p sitzt.

Wie berechnet man den Tangentialraum in den verschiedenen Darstellungsmöglichkeiten einer Mannigfaltigkeit?

3.2.2 Satz

Sei M n-dimensionale Untermannigfaltigkeit des $\mathbb{R}^N$, $p \in M$.

(a) T_pM ist ein n-dimensionaler Untervektorraum des $\mathbb{R}^N$.

(b) (Darstellung mit lokaler Karte) Falls $\varphi : U \xrightarrow{\subset \mathbb{R}^n} V \subset M$, $p \in V$, lokale Karte ist, mit $\varphi(x_0) = p$, so gilt

$$T_pM = \text{Bild}(D\varphi_{|x_0}).$$

(c) (Darstellung als Lösungsmenge von Gleichungen) Falls $f : W \xrightarrow{\subset \mathbb{R}^N} \mathbb{R}^{N-n}$, $M \cap W = f^{-1}(0)$, $p \in M \cap W$, $Df_{|p}$ surjektiv, dann ist

$$T_pM = \text{Kern}(Df_{|p}).$$

(b) kann auch so formuliert werden: T_pM ist der von $\frac{\partial \varphi}{\partial x_1}(x_0), \ldots, \frac{\partial \varphi}{\partial x_n}(x_0)$ aufgespannte Unterraum. Diese Vektoren bilden eine Basis von T_pM.

Beweis:

(b): Kurven $\gamma : I \longrightarrow M$, $\gamma(0) = p$
entsprechen Kurven $\tilde{\gamma} : I \longrightarrow U$, $\tilde{\gamma}(0) = x_0$

mittels $\tilde{\gamma} = \varphi^{-1} \circ \gamma$ (evtl. muss man I verkleinern, um im Kartengebiet von φ zu bleiben). Nach der Kettenregel ist

$$\gamma = \varphi \circ \tilde{\gamma} \;\Rightarrow\; \gamma'(0) = D\varphi_{|_{\tilde{\gamma}(0)}}(\tilde{\gamma}'(0))$$

Die Menge der möglichen $\tilde{\gamma}'(0)$ (über alle $\tilde{\gamma}$) ist $\mathbb{R}^n$, denn zu $w \in \mathbb{R}^n$ kann man $\tilde{\gamma}(t) = x_0 + tw$ nehmen. Daraus folgt (b).

(a): Da $D\varphi_{|x_0}$ eine lineare Abbildung ist, ist nach (b) T_pM ein Vektorraum. Außerdem φ Immersion $\Rightarrow$ $D\varphi_{|x_0} : \mathbb{R}^n \longrightarrow \mathbb{R}^N$ injektiv $\Rightarrow$ dim Bild $(D\varphi_{|x_0}) = n$.

(c): Ist γ wie in der Definition von T_pM, so gilt $f(\gamma(t)) = 0$ für alle t, da $f(q) = c$ für alle $q \in M$. Leitet man nach t ab, verwendet die Kettenregel und setzt $t = 0$, so folgt $Df_{|_{\gamma(0)}}(\gamma'(0)) = 0$, also $Df_{|p}(v) = 0$. Da dies für alle γ, also alle $v \in T_pM$ gilt, folgt $T_pM \subset \operatorname{Kern} Df_{|p}$. Da beides lineare Unterräume der Dimension n sind (nach (a) und weil $\dim \operatorname{Kern} Df_{|p} = N - (N-n) = n$ wegen der Surjektivität), folgt (c). □

Bemerkung: Teil (b) lässt sich leicht mit der Taylor-Formel verstehen: Es ist

$$\varphi(x_0 + w) = \underbrace{\varphi(x_0) + D\varphi_{|x_0}(w)}_{\substack{\text{lineare Approximation} \\ \text{zu } \varphi \text{ nahe } x_0}} + o(\|w\|) \qquad (w \to 0)$$

Das heißt, der Vektor zwischen den Punkten $\varphi(x_0)$ und $\varphi(x_0 + w)$ ist für kleine w approximativ der Tangentialvektor $D\varphi_{|x_0}(w)$.

Spezialfall: Tangentialraum an einen Graphen

Der Fall $n = 1$, $N = 2$: Sei $I \subset \mathbb{R}$ Intervall, $g \in C^1(I)$.
$M = \{(x,y) : y = g(x), x \in I\} = \{(x, g(x)) : x \in I\}$. M ist Bild der Immersion $\varphi(x) = (x, g(x))$, $x \in I$.
Es gilt $D\varphi_{|x} = \varphi'(x) = \begin{pmatrix} 1 \\ g'(x) \end{pmatrix}$. Also folgt

$$T_pM = \operatorname{span}\left\{ \begin{pmatrix} 1 \\ g'(x) \end{pmatrix} \right\}, \quad \text{falls } p = (x, g(x)).$$

Der Fall $n = 2$, $N = 3$ (Fläche im $\mathbb{R}^3$): Sei $U \subset \mathbb{R}^2$ offen, $g \in C^1(U)$.
$M = \{(x,y,z) : z = g(x,y), (x,y) \in U\}$. Dies wird parametrisiert durch die Immersion $\varphi(x,y) = (x,y,g(x,y))$, $(x,y) \in U$.
Es gilt $D\varphi = \begin{pmatrix} \frac{\partial \varphi_1}{\partial x} & \frac{\partial \varphi_1}{\partial y} \\ \vdots & \vdots \\ \frac{\partial \varphi_3}{\partial x} & \frac{\partial \varphi_3}{\partial y} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ g_x & g_y \end{pmatrix}$, wobei wir kurz $g_x = \dfrac{\partial g}{\partial x}$ und $g_y = \dfrac{\partial g}{\partial y}$ schreiben, also

$$T_pM = \operatorname{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ g_x(x,y) \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ g_y(x,y) \end{pmatrix} \right\}, \quad \text{falls } p = (x,y,g(x,y)).$$

3.3 Das Volumen von Parallelotopen

Unser Ziel ist es, die »Oberfläche« einer n-dimensionalen Untermannigfaltigkeit des $\mathbb{R}^N$ zu berechnen. Genauer wollen wir ein Maß auf einer Untermannigfaltigkeit einführen. Dies erlaubt es dann, von Oberfläche zu sprechen und Funktionen zu integrieren.

Als Vorbereitung betrachten wir Parallelotope: Zu $a_1, \dots, a_n \in \mathbb{R}^N$ sei

$$P(a_1, \dots, a_n) = \left\{ \sum_{i=1}^{n} a_i t_i, t_i \in [0,1] \text{ für alle } i \right\}$$

der von $a_1, \dots, a_n$ aufgespannte **Spat** (oder **Parallelotop**).

Wir sind hier in der prekären Situation, dass wir etwas berechnen wollen, was noch gar nicht definiert ist. Unser Ziel ist es also, gleichzeitig eine Definition und eine Berechnungsmethode zu finden. Die Definition kann natürlich nicht beliebig sein. Sie sollte gewissen »offensichtlichen« Anforderungen genügen. Wir lassen uns hier von der Vorstellung leiten (etwa im Fall $n = 2$, $N = 3$), dass die Fläche des Parallelogramms $P(a_1, a_2)$ gleich dem Volumen eines Körpers der Dicke 1 mit Querschnitt $P(a_1, a_2)$ sein sollte. Das heißt, man wählt einen Vektor $v \perp a_1, a_2$ mit $\|v\| = 1$ und berechnet $\mathrm{vol}(P(v, a_1, a_2))$.

In der folgenden Definition wird dies analog in höheren Dimensionen durchgeführt, und der Satz gibt eine Berechnungsmethode an.

Bemerkung: Alternativ kann man sich von der Idee leiten lassen, dass die Fläche invariant unter Rotationen sein soll. Man rotiert dann $P(a_1, a_2)$ so, dass es in der (x_1, x_2)-Ebene landet; dort haben wir bereits einen Flächenbegriff. Dieses Verfahren führt auf dasselbe Resultat.

3.3.1 Satz (und Definition)

Seien $a_1, \dots, a_n \in \mathbb{R}^N$.

(1) Definiere

$$\mathbf{vol_n}(P(a_1, \dots, a_n)) := \mathrm{vol}_N(P(a_1, \dots, a_n, v_1, \dots, v_{N-n}))$$

für eine Orthonormalbasis $v_1, \dots, v_{N-n}$ von $(\mathrm{span}\{a_1, \dots, a_n\})^\perp$. Dies ist wohldefiniert, also unabhängig von der Wahl der $v_1, \dots, v_{N-n}$.

(2) Es gilt:

$$\mathrm{vol}_n(P(a_1, \dots, a_n)) = \sqrt{\det A^t A}$$

mit der $N \times n$-Matrix $A = (a_1 \cdots a_n)$.

Konkret ist

$$A^t A = \begin{pmatrix} < a_1, a_1 > & \cdots & < a_1, a_n > \\ \vdots & \dots & \vdots \\ < a_n, a_1 > & \cdots & < a_n, a_n > \end{pmatrix},$$

wobei $<,>$ das Skalarprodukt ist. Dies ist eine $n \times n$-Matrix. Sie heißt **Gram-Matrix** von A oder von $a_1, \dots, a_n$. Beachte, dass A nicht quadratisch ist (falls $n < N$), also selbst keine Determinante hat.

Beweis: Sei $B = (a_1, \dots, a_n, v_1, \dots, v_{N-n})$. Aus Satz 2.3.1 folgt $\mathrm{vol}_N(P(a_1, \dots, a_n, v_1, \dots, v_{N-n})) = |\det B|$. Wie sieht man, dass dies unabhängig von der Wahl der v_i ist, und wie berechnet man es? Mit folgendem

hübschen Trick: Berechne B^tB! Es gilt

$$
\begin{aligned}
B^tB &= \begin{pmatrix} a_1^t \\ \vdots \\ a_n^t \\ v_1^t \\ \vdots \\ v_{N-n}^t \end{pmatrix} \cdot \begin{pmatrix} a_1 & \dots & a_n & v_1 & \dots & v_{N-n} \end{pmatrix} \\
&= \begin{pmatrix} a_1^t \cdot a_1 & \cdots & a_1^t \cdot a_n & a_1^t \cdot v_1 & \cdots & a_1^t \cdot v_{N-n} \\ \vdots & & & & & \vdots \\ a_n^t \cdot a_1 & \cdots & a_n^t \cdot a_n & a_n^t \cdot v_1 & \cdots & a_n^t \cdot v_{N-n} \\ v_1^t \cdot a_1 & \cdots & v_1^t \cdot a_n & v_1^t \cdot v_1 & \cdots & v_1^t \cdot v_{N-n} \\ \vdots & & & & & \vdots \\ v_{N-n}^t \cdot a_1 & \cdots & v_{N-n}^t \cdot a_n & v_{N-n}^t \cdot v_1 & \cdots & v_{N-n}^t \cdot v_{N-n} \end{pmatrix} \\
&= \begin{pmatrix} < a_1, a_1 > & \cdots & < a_1, a_n > & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ < a_n, a_1 > & \cdots & < a_n, a_n > & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 1 & \cdots & 0 \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 1 \end{pmatrix} \\
&= \begin{pmatrix} A^tA & 0 \\ 0 & I \end{pmatrix}
\end{aligned}
$$

Daraus folgt $\det B^tB = \det A^tA \cdot \det I = \det A^tA$. Da B eine quadratische Matrix ist, gilt $\det B^tB = \det B^t \cdot \det B = (\det B)^2$, also folgt $|\det B| = \sqrt{\det A^tA}$, und da dies offensichtlich unabhängig von $v_1, \dots, v_{N-n}$ ist, ist der Satz bewiesen. □

Bemerkung: Im Fall $n = N$ ist A eine quadratische Matrix, also ist $\det A^tA = \det A^t \det A = (\det A)^2$, also $\mathrm{vol}_n(P(a_1, \dots, a_n)) = |\det A|$. Das heißt, Satz 3.3.1 ist in diesem Spezialfall mit Satz 2.3.1 konsistent.

3.3.2 Korollar

Sei Q ein Spat in $\mathbb{R}^n$ und $A : \mathbb{R}^n \longrightarrow \mathbb{R}^N$ linear. Dann gilt

$$\mathrm{vol}_n(A(Q)) = \sqrt{\det A^t A} \cdot \mathrm{vol}_n(Q)$$

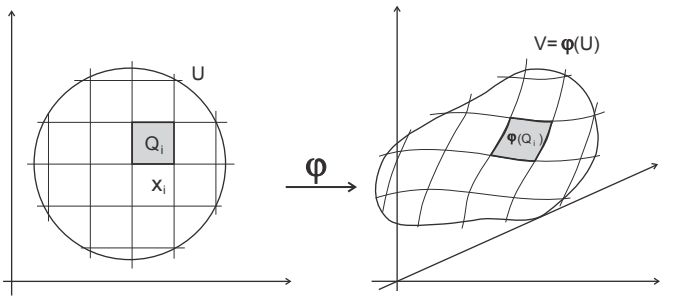
Beweis: Sei Q von den Vektoren $b_1, \dots, b_n \in \mathbb{R}^n$ aufgespannt und $B = (b_1, \dots, b_n)$. Dann ist $A(Q)$ von den Vektoren $Ab_1, \dots, Ab_n$ aufgespannt, diese bilden die Spalten der Matrix AB und daher gilt (B ist quadratisch)

$$
\begin{aligned}
\mathrm{vol}_n(A(Q))^2 &= \det((AB)^tAB) = \det(B^tA^tAB) = \det B^t \det(A^tA) \det B \\
&= \det(A^tA)\, \mathrm{vol}_n(Q)^2
\end{aligned}
$$

□

3.4 Volumen von und Integration über Untermannigfaltigkeiten des $\mathbb{R}^N$

Ähnlich wie schon bei den Parallelotopen sind wir in der prekären Lage, gleichzeitig eine Definition und eine Berechnungsmethode für die »Oberfläche« einer Untermannigfaltigkeit geben zu müssen. Wir lassen uns von einem ähnlichen Bild wie im Beweis der Transformationsformel leiten:



Integration über ein Kartengebiet

Heuristische Vorüberlegung: Ist $\varphi : U \to V \subset M$ eine lokale Karte, so zerlegen wir U in kleine Würfel Q_i. Das Bild $\varphi(Q_i)$ ist dann eine Teilmenge von M, die von dem Parallelotop $\tilde{Q}_i = D\varphi_{|x_i}(Q_i)$ (genauer einer Verschiebung davon) approximiert wird. Dessen Fläche (n-dimensionales Volumen) kennen wir bereits, das motiviert die Definition:

3.4.1 Definition

Sei $\varphi : U \subset \mathbb{R}^n \to \mathbb{R}^N$ eine C^1-Abbildung.

$$g_{\varphi}(x) := \det\left[(D\varphi_{|x})^t\, D\varphi_{|x}\right]$$

heißt **Gramsche Determinante** und die $n \times n$-Matrix $D\varphi_{|x})^t\, D\varphi_{|x}$ heißt **Gram-Matrix** von φ in x.

Fortsetzung der Vorüberlegung: Nach Korollar 3.3.2 ist $\mathrm{vol}_n(\tilde{Q}_i) = \sqrt{g_{\varphi}(x_i)}\ \mathrm{vol}_n(Q_i)$.
Ist nun $f : V \longrightarrow \mathbb{R}$, so liegt es nahe, das Integral von f über V als Grenzwert von $\sum_i f(\varphi(x_i)) \cdot \mathrm{vol}_n(\tilde{Q}_i)$, für ein immer feiner werdendes Würfelnetz, zu definieren. Nun ist

$$\sum_i f(\varphi(x_i)) \cdot \mathrm{vol}_n(\tilde{Q}_i) = \sum_i f(\varphi(x_i))\sqrt{g_{\varphi}(x_i)}\ \mathrm{vol}_n(Q_i) \to \int_U f(\varphi(x))\sqrt{g_{\varphi}(x)}\, dx\,,$$

(wenn das Würfelnetz immer feiner wird) und das motiviert die folgende Definition.

3.4.2 Definition (Integration einer Funktion über ein Kartengebiet einer Mannigfaltigkeit)

Sei M eine n-dimensionale Untermannigfaltigkeit des $\mathbb{R}^N$. Sei $V \subset M$ ein Kartengebiet, $f : V \longrightarrow \mathbb{R}$. Sei $\varphi : U \longrightarrow V$ eine lokale Karte, dann definiere

$$\int_V f\, dS := \int_U f(\varphi(x))\sqrt{g_{\varphi}(x)}\, dx$$

falls die rechte Seite existiert, d. h. falls

$$(f \circ \varphi) \cdot \sqrt{g_{\varphi}} \in \mathcal{L}^1(U)$$

Weiterhin heißt

$$\mathbf{vol}_n(V) := \int_V 1 \cdot dS$$

das ***n*-dimensionale Volumen** von V.

Die Bezeichnung dS stammt von »surface area«.

Eine wesentliche Rechtfertigung für die Definition liegt in folgendem Lemma. Eine weitere wird in Satz 3.4.9 folgen.

3.4.3 Lemma

$\int_V f\,dS$ ist wohldefiniert, d. h.:

Falls $\Phi : U' \longrightarrow V$ eine weitere lokale Karte ist, so gilt $\int_U (f\circ\varphi)\cdot\sqrt{g_\varphi} = \int_{U'} (f\circ\psi)\cdot\sqrt{g_\psi}$

Wir brauchen hier und später immer wieder:

3.4.4 Lemma

Seien $\varphi : U \longrightarrow V$ und $\psi : U' \longrightarrow V$ lokale Karten, dann ist der **Kartenwechsel**

$$\psi^{-1}\circ\varphi : U \longrightarrow U'$$

ein Diffeomorphismus.

Beweis:

- ▷ Bijektiv, weil φ, ψ bijektiv.
- ▷ Stetig differenzierbar: nicht schwierig mittels der Darstellung als lokaler Graph (Übung).
- ▷ Da das Inverse gleich $\varphi^{-1}\circ\psi$ ist und somit dieselbe Form hat, ist es ebenfalls stetig differenzierbar.□

Beweis (von Lemma 3.4.3): Sei $\kappa = \psi^{-1}\circ\varphi : U \longrightarrow U'$. Wende die Transformationsformel für die Funktion $(f\circ\psi)\cdot\sqrt{g_\psi}$ an:

$$\int_U \underbrace{((f\circ\psi)\circ\kappa)}_{\substack{f\circ\psi\circ\psi^{-1}\circ\varphi\\ =f\circ\varphi}} \sqrt{g_\psi\circ\kappa}\,|\det D\kappa|\,dx = \int_{\kappa(U)} (f\circ\psi)\sqrt{g_\psi}\,dy$$

Es bleibt zu zeigen, dass $\sqrt{g_\psi\circ\kappa}\,|\det D\kappa| = \sqrt{g_\varphi}$ gilt. Wende die Kettenregel für $\psi\circ\kappa = \varphi$ an: $D\psi|_{\kappa(x)}\circ D\kappa|_x = D\varphi|_x$. Dieselbe Rechnung wie im Beweis des Korollars 3.3.2 ergibt, da $D\kappa$ eine $n\times n$-Matrix ist,

$$\begin{aligned} g_\varphi &= \det((D\varphi)^t D\varphi) = \det((D\psi\,D\kappa)^t D\psi\,D\kappa)\\ &= \det((D\kappa)^t (D\psi)^t D\psi\,D\kappa)\\ &= \det(D\kappa)^t\cdot\det((D\psi)^t D\psi)\cdot\det(D\kappa)\end{aligned}$$

Der Übersicht halber wurden hier die Argumente weggelassen. Wenn g_φ bei $x\in U$ ausgewertet wird, so wird $D\kappa$ ebenfalls bei x ausgewertet, aber $D\psi$ bei $\kappa(x)$. Mit $\det((D\psi)^t D\psi) = g_\psi$ folgt also $g_\varphi = (g_\psi\circ\kappa)\cdot(\det D\kappa)^2$, und nach Ziehen der Wurzel folgt die Behauptung. □

Wichtige Spezialfälle:

Kurvenlänge: Sei $\gamma : I \longrightarrow \mathbb{R}^N$ eine C^1-Kurve. Wir nehmen an, dass $\gamma'(t)\neq 0$ für alle t ist, d. h. dass γ eine Immersion ist.

Es ist $D\gamma = \gamma'$ und $(\gamma')^t\gamma' = \langle\gamma',\gamma'\rangle = \|\gamma'\|^2$, also $\sqrt{g_\gamma} = \|\gamma'\|$.

$$\boxed{L(\gamma) = \int_I \|\gamma'(t)\|dt \quad \text{(Kurvenlänge)}}$$

Beispiel: $\gamma(t) = (\cos(t),\sin(t))$, $T\in(0,2\pi)$

$$\|\gamma'(t)\|^2 = \|(-\sin(t),\cos(t)\|^2 = \sin^2(t)+\cos^2(t) = 1$$

$$\Rightarrow\ L(\gamma) = \text{ Umfang des Kreises } = \int_0^{2\pi} 1\,dt = 2\pi$$

Beachte: Betrachtet man etwa das Intervall $(0,4\pi)$ statt $(0,2\pi)$, erhält man die doppelte Länge, obwohl das Bild von γ immer noch derselbe Kreis ist. Denn der Kreis wird nun zweimal durchlaufen.

Länge eines Graphen: Ist die Kurve als Graph einer Funktion $h : I \to \mathbb{R}$ (oder auch $h : I \to \mathbb{R}^{n-1}$) gegeben, so ist $\gamma(t) = (t, h(t))$, also $\gamma'(t) = (1, h'(t))$ und daher

$$\boxed{\text{Kurvenlänge des Graphen von } h : I \to \mathbb{R}^{n-1} : \quad \int_I \sqrt{1 + \|h'(t)\|^2}\, dt}$$

Das führt selbst bei recht einfachen Kurven zu nicht explizit berechenbaren Integralen.

Beispiel: $h(t) = t^3,\ t \in [0,1] \quad \rightsquigarrow \quad \int_0^1 \sqrt{1 + 9t^4}\, dt$

Fläche eines Graphen: Wir berechnen die Fläche des Graphen von $h : U \longrightarrow \mathbb{R}$, $U \subset \mathbb{R}^2$:

$$\varphi(x,y) = (x,y,h(x,y))$$

$$D\varphi = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ h_x & h_y \end{pmatrix} = (\varphi_x, \varphi_y)$$

$$g_\varphi = \det((D\varphi)^t D\varphi) = \det \begin{pmatrix} \langle \varphi_x, \varphi_x \rangle & \langle \varphi_x, \varphi_y \rangle \\ \langle \varphi_y, \varphi_x \rangle & \langle \varphi_y, \varphi_y \rangle \end{pmatrix} = \det \begin{pmatrix} 1 + h_y^2 & h_x h_y \\ h_x h_y & 1 + h_y^2 \end{pmatrix}$$

$$= 1 + h_x^2 + h_y^2 = 1 + \|\nabla h\|^2$$

Also

$$\boxed{\text{Fläche des Graphen von } h : U \longrightarrow \mathbb{R} : \quad \int_U \sqrt{1 + \|\nabla h\|^2}}$$

(Übung: Dieselbe Formel gilt auch für vol_n des Graphen einer Funktion $h : U \to \mathbb{R}$ mit $U \subset \mathbb{R}^n$.)

Der Fall $n = N$: Was sind die n-dimensionalen Untermannigfaltigkeiten des $\mathbb{R}^n$? Die offenen Mengen! Eine Karte reicht: $\varphi = \mathrm{id}_U : U \to U$ (also $U = V = M$).

Es ergibt sich $D\varphi_{|x} = \mathrm{id}$ für alle x, also $g_\varphi \equiv 1$, also

$$\int_M f\, dS = \int_M f(x)\, dx.$$

Also stimmt der neue Integralbegriff in diesem Fall mit dem alten überein. Verwendet man eine andere Karte, etwa $\psi : U' \to M$, so gilt $\sqrt{g_\psi(x)} = |\det D\psi_{|x}|$, da $D\psi_{|x}$ eine quadratische Matrix ist. Die Aussage von Lemma 3.4.3, dass bei der Berechnung von $\int_M f\, dS$ mittles ψ dasselbe herauskommt, ist also genau die Transformationsformel!

Integration über eine Untermannigfaltigkeit des $\mathbb{R}^N$

Beispiel: Fläche der Sphäre: $S^2 = \{(x,y,z) : x^2 + y^2 + z^2 = 1\}$.

1. Schritt: Finde lokale Karten.
Am einfachsten: Stelle S^2 oder Teile davon als Graph dar. Wie geht das? Auflösen der Gleichung $x^2 + y^2 + z^2 = 1$ nach z:

$$z^2 = 1 - x^2 - y^2$$

$$z = \pm\sqrt{1 - x^2 - y^2}$$

Das $\pm$ bedeutet, dass man zwei Funktionen braucht, für die obere und die untere Halbsphäre. Zerlege $S^2 = \underset{\uparrow \{z>0\}}{S^2_+} \cup \underset{\uparrow \{z<0\}}{S^2_-} \cup \underset{\uparrow \{z=0\}}{\ddot{A}}$ ($\ddot{A}$ = Äquator).

$S^2_+ =$ Graph von $h(x,y) = \sqrt{1-x^2-y^2}$ auf $\{(x,y) : x^2+y^2 < 1\}$. Es ist

$$h_x = \frac{-x}{\sqrt{1-x^2-y^2}} \qquad h_y = \frac{-y}{\sqrt{1-x^2-y^2}}$$

$$\|\nabla h\|^2 = \frac{x^2+y^2}{1-x^2-y^2}$$

$$1 + \|\nabla h\|^2 = \frac{1-x^2-y^2+x^2+y^2}{1-x^2-y^2} = \frac{1}{1-x^2-y^2}$$

Also: Fläche von $S^2_+ = \int\limits_{x^2+y^2<1} \frac{1}{\sqrt{1-x^2-y^2}}\, dx\, dy = 2\pi$ (Übung). Wegen Symmetrie ist die Fläche von S^2_- ebenfalls 2π. Da $\mathrm{vol}_2(\ddot{A}) = 0$ (siehe unten), folgt

$$\mathrm{vol}_2(S^2) = 4\pi\,.$$

Das Beispiel zeigt den Weg, wie man eine Funktion $f : M \longrightarrow \mathbb{R}$ integriert: Indem man M in Kartengebiete zerlegt.

> **Bemerkung:** In den meisten Büchern wird hier die sogenannte »Zerlegung der Eins« verwendet. Das ist jedoch an dieser Stelle unnötig kompliziert (overkill). Später werden wir die Zerlegung der Eins in anderem Kontext kennenlernen.

Zunächst die nötigen Begriffe:

3.4.5 Definition

Sei $M \subset \mathbb{R}^N$ n-dimensionale Untermannigfaltigkeit und $(\varphi_i, U_i, V_i)_{i\in I}$ ein Atlas für M.

(a) $X \subset M$ heißt ***n*-messbar** $:\Longleftrightarrow$ $\varphi_i^{-1}(X \cap V_i)$ ist messbare Teilmenge des $\mathbb{R}^n$ für alle i.

(b) $X \subset M$ heißt ***n*-Nullmenge** $:\Longleftrightarrow$ $\varphi_i^{-1}(X \cap V_i)$ ist Nullmenge in $\mathbb{R}^n$ für alle i.

Man sieht leicht, dass diese Definition unabhängig von der Wahl des Atlasses ist. (Zu zeigen: Falls (a) bzw. (b) gilt und $\varphi : U \longrightarrow V \subset M$ eine beliebige Karte ist, so ist auch $\varphi^{-1}(X \cap V)$ messbar bzw. Nullmenge. Grund: $\kappa = \varphi^{-1} \circ \varphi_i$ ist Diffeomorphismus, also κ(messbar) $=$ messbar, κ(Nullmenge) $=$ Nullmenge).

> **Bemerkung:** Man kann die Begriffe der n-Messbarkeit und n-Nullmenge auch für beliebige Teilmengen von $\mathbb{R}^N$ definieren (auch für solche, die nicht in eine n-dimensionale Untermannigfaltigkeit passen). Letzterer ist relativ einfach (eine einfache Modifikation der Bedingung (1) in Satz 1.4.3), siehe z. B. Königsberger, Analysis II, 11.6. Das zugehörige Maß heißt **Hausdorff-Maß** der Dimension n. Übrigens gibt es analog auch das Hausdorff-Maß der Dimension d für beliebige positive *reelle* (!) Zahlen d. Für $d \notin \mathbb{N}_0$ führt dies auf den Begriff des Fraktals.

Im Beispiel oben braucht man:

3.4.6 Lemma

Sei $M \subset \mathbb{R}^N$ eine n-dimensionale Untermannigfaltigkeit und $N \subset M$ eine Untermannigfaltigkeit einer Dimension $< n$. Dann ist N eine n-Nullmenge.

Beweis: Übung. □

Um das Integral über Mannigfaltigkeiten definieren zu können, brauchen wir noch ein technisches Lemma:

3.4.7 Lemma

Sei $M \subset \mathbb{R}^N$ Untermannigfaltigkeit.

(a) Jeder Atlas von M hat einen abzählbaren Teilatlas.

(b) Es existiert eine Zerlegung von M in abzählbar viele Teile X_i, $i \in \mathbb{N}$, mit:

- X_i messbar für alle i.
- Jedes X_i ist in einem Kartengebiet V_i enthalten.

Zerlegung heißt: $\cup_i X_i = M$, $X_i \cap X_j = \varnothing$ für $i \neq j$.

Beweis:
(a): Es ist zu zeigen, dass jede offene Überdeckung $(V_i)_{i\in I}$ von M eine abzählbare Teilüberdeckung besitzt. Dies geht mit einem typischen Argument der »mengentheoretischen Topologie«, z. B. so:

Sei $\mathcal{K}$ die Menge der offenen Kugeln in $\mathbb{R}^N$ mir rationalem Radius, deren Mittelpunkte nur rationale Koordinaten haben. $\mathcal{K}$ ist abzählbar. Sei $\mathcal{K}' = \{K \cap M : K \in \mathcal{K}\}$ und $\mathcal{K}''$ die Menge derjenigen $K \in \mathcal{K}'$, die in mindestens einem V_i enthalten sind. Da die Kugeln in $\mathcal{K}$ beliebig klein sein können und die V_i offen sind, ist jedes V_i die Vereinigung derjenigen $K \in \mathcal{K}'$, die in ihm enthalten sind. Daher ist $\mathcal{K}''$ eine abzählbare Überdeckung von M. Wählt man nun zu jedem $K \in \mathcal{K}''$ ein $i_K \in I$ mit $K \subset V_{i_K}$, so folgt, dass die V_{i_K}, $K \in \mathcal{K}''$ eine abzählbare Teilüberdeckung von $(V_i)_{i\in I}$ bilden.

(b): Nach (a) existiert ein abzählbarer Atlas. Seien $V_1, V_2 \ldots$ seine Kartengebiete. Setze

$$\begin{aligned} X_1 &= V_1 \\ X_2 &= V_2 \backslash V_1 \\ X_3 &= V_3 \backslash (V_1 \cup V_2) \\ &\vdots \end{aligned}$$

Da offene Mengen in $\mathbb{R}^n$ und damit Kartengebiete in M sowie Differenzen messbarer Mengen messbar sind, sind alle X_i messbar, und es gilt offenbar $X_i \subset V_i$ für alle i. □

Für eine messbare Teilmenge X eines Kartengebietes V schreiben wir $\int_X f\,dS = \int_V f \cdot \chi_X\,dS$ und nennen f integrierbar über X, wenn $f \cdot \chi_X$ integrierbar über V ist.

3.4.8 Definition

Sei $M \subset \mathbb{R}^N$ n-dimensionale Untermannigfaltigkeit, $f : M \longrightarrow \mathbb{R}$. Wähle eine Zerlegung $X_1, X_2, \ldots$ wie in Lemma 3.4.7. Dann definiere

$$\int_M f\,dS := \sum_{i=1}^{\infty} \int_{X_i} f\,dS$$

Genauer:

$$\begin{aligned} & f \in \mathcal{L}^1(M) \quad \text{(integrierbar über } M\text{)} \\ & :\Longleftrightarrow\ f \text{ integrierbar über } X_i \text{ für alle } i \text{ und } \sum_{i=1}^{\infty} \int_{X_i} |f|\,dS < \infty \end{aligned}$$

Für messbare Teilmengen $X \subset M$ definieren wir natürlich wieder $\mathrm{vol}_n(X) := \int_M \chi_X\,dS$.

Bemerkung: Es ist leicht zu zeigen, dass die Definition unabhängig von der Wahl der X_i, V_i ist.

Man kann nun die Konzepte und Aussagen der Integrationstheorie leicht auf die Integration über Mannigfaltigkeiten übertragen, z. B.:

- ▷ Die Menge der n-messbaren Teilmengen einer n-dimensionalen Untermannigfaltigkeit des $\mathbb{R}^N$ ist eine σ-Algebra und vol_n ist ein Maß darauf. n-Nullmengen sind genau die Nullmengen bzgl. dieses Maßes.
- ▷ Ist $X \subset M$ n-messbar und $f \in \mathcal{L}^1(M)$, so ist $\int_X f\,dS$ definiert.
  Ist $X \subset M$ zusätzlich n-Nullmenge, so ist $\int_X f\,dS = 0$.
- ▷ Sind $f, g \in \mathcal{L}^1(M)$ und $f = g$ außerhalb einer n-Nullmenge, so gilt $\int_M f\,dS = \int_M g\,dS$.
- ▷ Ist $f : M \to \mathbb{R}$ stetig außerhalb einer n-Nullmenge und außerhalb einer kompakten Teilmenge von M gleich null, so ist $f \in \mathcal{L}^1(M)$.
- ▷ Die Konvergenzsätze gelten analog.
- ▷ etc.

Das Volumen von Tuben

Zum Abschluss dieses Kapitels noch ein kleines Juwel.

Wem die heuristische Hinführung zur Definition des Integrals über (bzw. Volumens von) Untermannigfaltigkeiten nicht behagte, die mag Befriedigung darin finden, dass mit dieser Definition eine anschaulich klare Aussage wirklich stimmt:

Dickt man eine Fläche M im $\mathbb{R}^3$ auf die Dicke 2ε auf, so ist das Volumen des erhaltenen Körpers M_ε etwa $2\varepsilon\,\mathrm{vol}_2(M)$, genauer:

$$\lim_{\varepsilon\to 0} \frac{\mathrm{vol}_3(M_\varepsilon)}{2\varepsilon} = \mathrm{vol}_2(M)\,.$$

Das sollte nicht überraschen, ist es doch sehr ähnlich zu der Idee bei unserem Ausgangspunkt zur Flächenberechnung, der Definition von vol_n für Parallelotope, Definition 3.3.1. Der Faktor 2 kommt daher, dass wir an beiden Seiten von M eine Schicht der Dicke ε haben.

Dies hätte man natürlich auch als Definition von vol_2 verwenden können, es eignet sich aber weniger für Berechnungen.

Etwas Analoges gilt für beliebige Dimensionen, und sogar noch etwas viel Präziseres. Die im Satz definierte Menge M_ε nennt man manchmal (vom Fall $n = 1$, $N = 3$ her) **Tube** um M.

3.4.9 Satz (Weylsche Tubenformel)

Sei $M \subset \mathbb{R}^N$ kompakte n-dimensionale Untermannigfaltigkeit. Sei $M_\varepsilon := \{x \in \mathbb{R}^N : \mathrm{dist}(x, M) < \varepsilon\}$ und $v(\varepsilon) = \mathrm{vol}_N(M_\varepsilon)$. Dann gibt es ein $\varepsilon_0 > 0$, so dass $v(\varepsilon)$ für $0 < \varepsilon < \varepsilon_0$ durch ein Polynom in ε gegeben ist, der Form

$$v(\varepsilon) = \varepsilon^{N-n}(a_0 + a_1\varepsilon + \ldots + a_n\varepsilon^n)$$

Hierbei ist $a_0 = C_{N-n}\,\mathrm{vol}_n(M)$, wobei $C_k =$ Volumen der Einheitskugel in $\mathbb{R}^k$.
Insbesondere ist

$$\lim_{\varepsilon\to 0} \frac{v(\varepsilon)}{C_{N-n}\varepsilon^{N-n}} = \mathrm{vol}_n(M)\,.$$

Wegen $C_1 = 2$ erhält man für $n = 2$, $N = 3$ die vorhergehende Behauptung.

Bemerkung: Die Formel für a_0 ist leicht zu verstehen (stellen Sie sich zunächst $n = 1$, $N = 3$ vor): In erster Näherung kann man vernachlässigen, dass M »gekrümmt« ist. Ist M nicht gekrümmt, also etwa $M = U \times \{0\}^{N-n}$ mit einer offenen Teilmenge $U \subset \mathbb{R}^n$, so ist $M_\varepsilon = (U \times K_\varepsilon(0)) \cup R$, wobei $K_\varepsilon(0)$ die Kugel im $\mathbb{R}^{N-n}$ ist, also Volumen $C_{N-n}\varepsilon^{N-n}$ hat, und R ein Rest ist, der vom Rand von U stammt. Das Volumen von R ist von der Ordnung ε^{N-n+1} (das ist zumindest für $n = 1, N = 3$ klar, da R für ein Intervall U aus zwei Halbkugeln vom Radius ε besteht).

(Genau genommen kann M gar nicht flach sein, da es als kompakt angenommen wurde. Dies ist ja auch nur eine Heuristik.)

Höchst überraschend ist jedoch, dass nicht nur die Grenzwertbeziehung am Ende des Satzes gilt, sondern dass die zunächst kompliziert anmutende Funktion v die einfache Gestalt eines Polynoms hat.

Die Bedeutung der anderen Koeffizienten ist dabei weniger einfach zu verstehen. Im Fall $n = 2$, $N = 3$ ist $v(\varepsilon) = a_0\varepsilon + a_1\varepsilon^2 + a_2\varepsilon^3$, und a_2 hat eine überraschend einfache Bedeutung:

$$a_2 = \text{ Eulercharakteristik von } M$$

Die Eulercharakteristik kann man dabei wie folgt erklären: Eine kompakte Fläche M im $\mathbb{R}^3$ sieht entweder aus wie eine verformte Sphäre, ein verformter Torus, ein verformter Brezel mit 2 oder 3 oder ... Löchern (anders gesagt: man klebe k »Henkel« an die Sphäre, dann erhält man für $k = 0, 1, 2, \ldots$ alle Möglichkeiten). Die Euler-Charakteristik der Fläche ist dann definiert als $2 - 2k$. Dies ist eine *topologische Invariante*, d.h. sie hängt nicht von der »Verformung« ab, d.h. davon, wie genau M im Raum liegt.

Beweis (Skizze im Fall $n = 2$, $N = 3$): Wir verwenden folgende Tatsachen:

(1) Es gibt ein Einheitsnormalenvektorfeld auf M, d. h. eine stetige Abbildung $\nu : M \to \mathbb{R}^3$, so dass für alle $p \in M$ gilt: $\|\nu(p)\| = 1, \nu(p) \perp T_pM$.

(2) Es gibt ein $\varepsilon_0 > 0$, so dass

$$M \times (-\varepsilon, \varepsilon) \to M_\varepsilon$$
$$(p, t) \mapsto p + t\nu(p)$$

für $\varepsilon < \varepsilon_0$ ein Diffeomorphismus ist. Dies bedeutet: Ist $\varphi : U \to V \subset M$ eine lokale Karte für M, dann definiert

$$\Phi(x, t) = \varphi(x) + t\nu(x), \quad x \in U,\ t \in (-\varepsilon, \varepsilon)$$

eine lokale Karte für M_ε, und eine Zerlegung von M in Kartengebiete (bis auf eine 2-Nullmenge) liefert eine entsprechende Zerlegung von M_ε in Kartengebiete (bis auf eine Nullmenge).

Dies sollte intuitiv recht klar sein. (1) ist leicht zu zeigen (werden wir später tun), (2) aber etwas schwieriger.

Sei nun $\varphi : U \to V \subset M$ eine lokale Karte für M und Φ wie oben definiert. Sei $V(\epsilon) = \Phi(U \times (-\varepsilon, \varepsilon))$.

Es genügt, zu zeigen, dass $\mathrm{vol}(V(\epsilon))$ ein Polynom in ε der angegebenen Form ist, dessen ε-Koeffizient gleich $2\,\mathrm{vol}_2(V)$ ist.

Es ist

$$\begin{aligned} D\Phi &= (\partial_{x_1}\Phi, \partial_{x_2}\Phi, \partial_t\Phi) \\ &= (\partial_{x_1}\varphi + t\,\partial_{x_1}\nu, \partial_{x_2}\varphi + t\partial_{x_2}\nu, \nu) \end{aligned}$$

Nach der Leibniz-Formel für die Determinante ist $f(x, t) = \det D\Phi_{|(x,t)}$ ein Polynom in t, dessen Koeffizienten Funktionen von x sind:

$$\det D\Phi_{|(x,t)} = b_0(x) + tb_1(x) + t^2b_2(x)$$

Hierbei ist $b_0(x) = f(x, 0) = \det(\partial_{x_1}\varphi(x), \partial_{x_2}\varphi(x), \nu(x))$, und wegen der Eigenschaften von ν folgt $|b_0(x)| = \sqrt{g_\varphi(x)}$ (vgl. Satz 3.3.1 und Definition 3.4.1), insbesondere $b_0(x) \neq 0$ für alle x. O. B. d. A. sei

$b_0(x) > 0$ (hierzu muss man evtl. alle $\nu(x)$ umdrehen, d.h. durch $-\nu(x)$ ersetzen). Aus Stetigkeitsgründen ist dann auch $f(x,t) > 0$ für genügend kleine $|t|$.

Zusammenfassend haben wir bisher gezeigt: $\det D\Phi_{|(x,t)}$ ist für kleine $|t|$ positiv und ein quadratisches Polynom in t mit konstantem Koeffizienten $\sqrt{g_\varphi(x)}$.
Also folgt:

$$\begin{aligned}
\mathrm{vol}_3(V(\varepsilon)) &= \int_U \int_{-\varepsilon}^{\varepsilon} |D\Phi_{|(x,t)}|\, dt\, dx \\
&= \int_U \int_{-\varepsilon}^{\varepsilon} (b_0(x) + t b_1(x) + t^2 b_2(x))\, dt\, dx \\
&= \int_U \Big(2\varepsilon b_0(x) + \frac{2}{3}\varepsilon^3 b_2(x)\Big)\, dx \\
&= \varepsilon \cdot 2\Big(\int_U b_0(x)\, dx\Big) + \varepsilon^3 \cdot \frac{2}{3}\Big(\int_U b_2(x)\, dx\Big)
\end{aligned}$$

Wegen $b_0(x) = \sqrt{g_\varphi(x)}$ ist der erste Koeffizient gleich $2\,\mathrm{vol}_2(V)$, und damit sind wir fertig. (Wir sehen sogar, dass $a_1 = 0$ ist.) □

Literatur zur Weylschen Tubenformel:
H. Weyl, American Journal of Mathematics 61, 461–472 (1939)
Alfred Gray, Tubes. Birkhäuser 2003.

4 Hauptsätze der Differential- und Integralrechnung in mehreren Dimensionen

Wir haben nun die Integration als Mittel zur Berechnung von Volumina kennengelernt. In einer Dimension kennen wir eine weitere Bedeutung der Integration: als Umkehrung der Differentiation. Dies ist die Aussage des Hauptsatzes der Differential- und Integralrechnung. Gibt es davon eine mehrdimensinale Verallgemeinerung?

Der Hauptsatz hat zwei eng verwandte Teile:

(1) $\frac{d}{dx}\int_a^x f(t)\,dt \;=\; f(x)$

(2) $\int_a^b F'(t)\,dt \;=\; F(b) - F(a)$

Aus (1) folgt, dass jede stetige Funktion eine Stammfunktion besitzt. (2) hat praktische Bedeutung: Kennt man eine Stammfunktion F einer Funktion f, so kann man mittels F bestimmte Integrale über $f = F'$ ausrechnen. Dies ist genau das, was Sie beim praktischen Integrieren verwenden.
Wir werden zu beiden Teilen Verallgemeinerungen kennenlernen. Dies sind die sogenannten Integralsätze. Der zweite Teil ist eng mit dem Verfahren der partiellen Integration verwandt, denn dieses erhält man durch Anwenden von (2) auf ein Produkt $F = u \cdot v$ zweier Funktíonen. Daher könnte dieses Kapitel auch **»Partielle Integration in mehreren Dimensionen«** überschrieben sein.

4.1 Kurvenintegrale und Gradientenfelder

Die Ableitung einer Funktion $u : U \to \mathbb{R}$, $U \subset \mathbb{R}^n$ offen, lässt sich auf verschiedene Weisen verstehen. Wir kennen aus Analysis II diese Ableitungsbegriffe:

- ▷ Gradient: $\nabla u : U \to \mathbb{R}^n;\ x \mapsto \left(\frac{\partial u}{\partial x_1}(x), \ldots, \frac{\partial u}{\partial x_n}(x)\right)^T$.
- ▷ Differential: $Du : U \to L(\mathbb{R}^n, \mathbb{R});\ x \mapsto Du_{|x}$. Hierbei ist $Du_{|x} : \mathbb{R}^n \to \mathbb{R}$ eine lineare Abbildung.

∇u und Du enthalten dieselbe Information und sind verknüpft durch die Relation $Du_{|x}(v) = \langle \nabla u(x),\, v\rangle\ \forall v \in \mathbb{R}^n$.

∇u ist ein Vektorfeld, Du eine sogenannte 1-Form. Wir werden uns im Folgenden mit der Sichtweise der Vektorfelder begnügen. Diese ist intuitiver.

> **Bemerkung:** Für ein höheres Verständnis der Zusammenhänge wird später aber die Sichtweise von 1-Formen wichtig sein.

Frage: Gegeben eine offene Menge $U \subset \mathbb{R}^n$ und ein stetiges Vektorfeld V auf U, gibt es eine »Stammfunktion« zu V, also eine Funktion $u : U \to \mathbb{R}$ mit $V = \nabla u$?

Und wenn ja, wie findet man eine solche Funktion u?

Schon bei der ersten Frage unterscheidet sich der Fall $n > 1$ vom Fall $n = 1$. Die Antwort ist: Im Allgemeinen NEIN (falls $n > 1$). Denn:

Eine Integrabilitätsbedingung

4.1.1 Satz

Sei $U \subset \mathbb{R}^n$ offen und $V : U \to \mathbb{R}^n$ ein $\mathcal{C}^1$-Vektorfeld auf U.
Falls $V = \nabla u$ für eine $\mathcal{C}^2$-Funktion $u : U \to \mathbb{R}$ ist, so folgt mit $V = (V_1, \ldots, V_n)$:

$$\frac{\partial V_i}{\partial x_j} = \frac{\partial V_j}{\partial x_i} \qquad \text{für alle } i, j \in \{1, \ldots, n\}$$

Ein Vektorfeld, das eine Stammfunktion besitzt, nennt man auch **integrabel** und daher diese Gleichungen **Integrabilitätsbedingungen**.

Beweis: Dies ist nur eine Umformulierung der Tatsache, dass es bei $\mathcal{C}^2$-Funktionen egal ist, in welcher Reihenfolge man sie nach verschiedenen Variablen ableitet:

$$V = \nabla u \implies \begin{cases} V_i = \dfrac{\partial u}{\partial x_i} \Rightarrow \dfrac{\partial V_i}{\partial x_j} = \dfrac{\partial^2 u}{\partial x_j \partial x_i} \\ V_j = \dfrac{\partial u}{\partial x_j} \Rightarrow \dfrac{\partial V_j}{\partial x_i} = \dfrac{\partial^2 u}{\partial x_i \partial x_j} \end{cases}$$

□

Gilt auch die Umkehrung? D.h., ist die angegebene Bedingung an V nicht nur notwendig, sondern sogar hinreichend für die Existenz einer Stammfunktion? Die Antwort lautet: Je nachdem, wie U aussieht. Darin steckt der Schlüssel zu ganz viel höherer Mathematik (Algebraische Topologie)! Dies werden wir weiter unten erläutern.

Beispiele:

(1) Sei $v \in \mathbb{R}^n$ und $V(x) = v$ für alle $x \in \mathbb{R}^n$ (ein konstantes Vektorfeld)
Durch Raten findet man die Stammfunktion $u(x) = \langle v,\ x \rangle = \sum_{i=1}^{n} v_i x_i$, denn $\frac{\partial u}{\partial x_i}(x) = v_i$, also $\nabla u(x) = v$.

(2) $V(x, y) = (y, x)$ (für $x, y \in \mathbb{R}$) $\rightsquigarrow u(x, y) = x \cdot y$. In der Tat ist $\frac{\partial V_1}{\partial y} = 1 = \frac{\partial V_2}{\partial x}$.

(3) $V(x, y) = (y, 1)$ $\quad \frac{\partial V_2}{\partial x} = 0 \neq 1 = \frac{\partial V_1}{\partial y}$. Es gibt keine Stammfunktion!

Wie im eindimesionalen Fall gilt: Existiert eine Stammfunktion, gibt es unendliche viele davon (denn $\nabla u = \nabla(u + r)$ für beliebige $r \in \mathbb{R}$).
In den obigen Beispielen haben wir die Lösung nur geraten. Lässt sich eine Stammfunktion auch systematisch bestimmen?

Kurvenintegrale

Angenommen, für eine Funktion $u : U \to \mathbb{R}$ sind das Gradientenfeld $V = \nabla u$ sowie ein Funktionswert $u(p)$, $p \in I$, gegeben.

Frage: Wie berechnet man $u(q)$ für *beliebige* $q \in U$?

Die Idee ist, das Problem auf den eindimensionales Fall zurückzuführen, wo wir wissen, wie das geht: Mittels des Hauptsatzes.

Sei hierzu $\gamma : [a,b] \to U$ eine $\mathcal{C}^1$-Kurve von p nach q. Wir betrachten die Funktionswerte auf der Kurve, also die Funktion $t \mapsto u(\gamma(t)) =: F(t)$. Wegen $\gamma(a) = p$, $\gamma(b) = q$ folgt $u(q) - u(p) = u(\gamma(b)) - u(\gamma(a)) = F(b) - F(a)$. Nun ist F eine $\mathcal{C}^1$-Funktion einer Variablen, und mit dem Hauptsatz folgt $F(b) - F(a) = \int_a^b F'(t)\,dt$. Mit $F = u \circ \gamma$ folgt aus der Kettenregel $F'(t) = \langle \nabla u(\gamma(t)), \gamma'(t) \rangle\,dt$. Insgesamt erhalten wir

$$u(q) - u(p) = \int_a^b \langle \nabla u(\gamma(t)), \gamma'(t) \rangle\,dt$$

Dies motiviert die folgende Definition.

4.1.2 Definition

Sei $I \subset \mathbb{R}$ ein Intervall und $\gamma : I \to \mathbb{R}^n$ eine $\mathcal{C}^1$-Kurve. Sei $V : \text{Bild}(\gamma) \to \mathbb{R}^n$ ein stetiges Vektorfeld. Dann sei

$$\int_\gamma \boldsymbol{V} \coloneqq \int_I \langle V(\gamma(t)),\ \gamma'(t) \rangle\,dt$$

das **Kurven-/Wegintegral** über V entlang γ.

Ist $\gamma : [a,b] \to U$ nur **stückweise** stetig differenzierbar ($\gamma \in \mathcal{C}^1_{st}$), d.h. gibt es Punkte t_i mit $a = t_0 < t_1 < \ldots < t_{k-1} < t_k = b$ ($k \in \mathbb{N}$) und $\gamma|_{[t_{i-1},\, t_i]} \in \mathcal{C}^1$ für alle $i \in \{1, \ldots, k\}$, so definieren wir $\int_\gamma V$ durch dieselbe Formel. Dass der Integrand auf der Nullmenge $\{t_1, \ldots, t_{k-1}\}$ undefiniert ist, ist irrelevant.

Erinnerung: Kurven sind per Definition immer stetig, d.h. im $\mathcal{C}^1_{st}$- Fall muss γ bei den t_i stetig sein, nur der Tangentialvektor darf springen.

Bemerkung: Aus der Definition ist sofort klar, dass

$$\int_\gamma V = \sum_{i=1}^{k} \int_{\gamma|_{[t_{i-1},\, t_i]}} V$$

für einen beliebigen $\mathcal{C}^1_{st}$-Weg γ und beliebige $a = t_0 < t_1 < \ldots < t_{k-1} < t_k = b$ (nicht notwendig die Sprungstellen von γ') gilt.

Eine wichtige allgemeine Eigenschaft von Kurvenintegralen ist folgende:

> $\int_\gamma V$ hängt nur davon ab, in welcher *Richtung* $\text{Bild}(\gamma)$ durchlaufen wird, aber nicht davon, wie *schnell* dies passiert.

Genauer:

4.1.3 Lemma

Sei $U \subset \mathbb{R}^n$ offen, $\gamma : I \to U$ eine $\mathcal{C}^1_{st}$-Kurve und $\tilde{\gamma} : J \to U$ eine Umparametrisierung von γ, das heißt es gebe einen Diffeomorphismus $\kappa : J \to I$ mit $\tilde{\gamma} = \gamma \circ \kappa$. Dann gilt:

$$\int_\gamma V = \text{sgn}(\kappa') \cdot \int_{\tilde{\gamma}} V$$

Dabei sei $\text{sgn}(\kappa') = -1$, falls $\forall s \in J:\ \kappa'(s) < 0$ und $\text{sgn}(\kappa') = 1$, falls $\forall s \in J:\ \kappa'(s) > 0$ (einer der beiden Fälle tritt auf jeden Fall ein, da für einen Diffeomorphismus $\kappa'(s) \neq 0\ \forall s$ gilt).

Beweis: Es ist

$$\int_\gamma V = \int_I \langle V(\gamma(t)),\ \gamma'(t) \rangle\,dt$$

und mit der Transformationsformel folgt:

$$\int_I \langle V(\gamma(t)),\ \gamma'(t)\rangle\, dt = \int_J \langle V(\gamma(\kappa(s))),\ \gamma'(\kappa(s))\rangle \cdot |\kappa'(s)|\, ds$$

Da $|\kappa'(s)| = \operatorname{sgn}(\kappa') \cdot \kappa'(s)$ gilt und man in Skalarprodukte einen Faktor »reinziehen« kann, folgt weiter:

$$\int_J \langle V(\gamma(\kappa(s))),\ \gamma'(\kappa(s))\rangle \cdot |\kappa'(s)|\, ds = \operatorname{sgn}(\kappa') \cdot \int_J \langle V(\gamma(\kappa(s))),\ \gamma'(\kappa(s)) \cdot \kappa'(s)\rangle\, ds$$

Nach der Kettenregel gilt $\frac{d}{ds}\tilde{\gamma}(s) = \frac{d}{ds}\gamma(\kappa(s)) = \gamma'(\kappa(s)) \cdot \kappa'(s)$, d. h.:

$$\operatorname{sgn}(\kappa') \cdot \int_J \langle V(\gamma(\kappa(s))),\ \gamma'(\kappa(s)) \cdot \kappa'(s)\rangle\, ds \ = \ \operatorname{sgn}(\kappa') \cdot \int_J \langle V(\tilde{\gamma}(s)),\ \tilde{\gamma}'(s)\rangle\, ds \ = \ \operatorname{sgn}(\kappa') \cdot \int_{\tilde{\gamma}} V$$

Hierbei sind immer die endlich vielen Punkte t bzw. s auszulassen, bei denen γ' bzw. $\tilde{\gamma}'$ nicht definiert ist. □

Wir fassen die Überlegungen vor Definition 4.1.2 zusammen.

4.1.4 Satz

Sei $u : U \to \mathbb{R}$ eine $\mathcal{C}^1$-Funktion. Für $p, q \in U$ und jede $\mathcal{C}^1_{st}$-Kurve $\gamma : p \rightsquigarrow q$ gilt:

$$u(q) - u(p) = \int_\gamma \nabla u$$

Dies ist eine Version des zweiten Teil des Hauptsatzes!

Beweis: Dies wurde vor Definition 4.1.2 bewiesen, falls γ eine $\mathcal{C}^1$-Kurve ist. Ist γ bloß stückweise $\mathcal{C}^1$, also $\mathcal{C}^1$ auf Teilintervallen $[t_{i-1}, t_i]$, $i = 1, \ldots, k$, so folgt mit dem bereits Bewiesenen, dass $\int_{\gamma_{|[t_{i-1},t_i]}} \nabla u = u(\gamma(t_i)) - u(\gamma(t_{i-1}))$, und Summation über $i = 1, \ldots, k$ liefert die Behauptung auch in diesem Fall (Teleskopsumme auf der rechten Seite!). □

Man sieht: Wenn $V = \nabla u$ gilt, hängt das Ergebnis des Integrals gar nicht vom Verlauf der Kurve γ ab, sondern nur von deren Endpunkten! Das heißt, $\int_\gamma V$ ist gleich für *alle* $\gamma : p \rightsquigarrow q$.

Schreibweise: Sei $\gamma : [a, b] \to U \subset \mathbb{R}^n$ eine Kurve. Schreibe $\boldsymbol{\gamma : p \rightsquigarrow q}$, falls $\gamma(a) = p$ und $\gamma(b) = q$. Außerdem sei $\mathbf{Anfang}(\boldsymbol{\gamma}) = \gamma(a)$ und $\mathbf{Ende}(\boldsymbol{\gamma}) = \gamma(b)$.

> **Bemerkung:** In diesem Fall schreibt man manchmal einfach $\int_p^q V$ statt $\int_\gamma V$. Aber Vorsicht: Diese Schreibweise ist nur für Gradientenfelder V sinnvoll!

Wir können damit auch die eingangs gestellte Frage beantworten: Die gesuchte Stammfunktion u zu ∇u mit bekanntem $u(p)$ ist gegeben durch

$$u(x) = u(p) + \int_{\gamma : p \rightsquigarrow x} \nabla u$$

Der Rest des Kapitels widmet sich folgender

Frage: Wie sieht man einem Vektorfeld V an, ob es ein Gradientenfeld ist?

Wir werden zwei Teilantworten finden und diese dann in ihrer Kombination verfeinern.

Gradientenfelder und Integrale über geschlossene Kurven

4.1.5 Satz

Sei $U \subset \mathbb{R}^n$ offen und $V : U \to \mathbb{R}^n$ ein $\mathcal{C}^1$-Vektorfeld. Folgende Aussagen sind dann äquivalent:

(a) Es gibt ein $u \in \mathcal{C}^2(U)$ mit $V = \nabla u$ (»V ist ein Gradientenfeld«)

(b) Für alle $\mathcal{C}^1_{st}$-Wege γ_1, γ_2 mit $\text{Anfang}(\gamma_1) = \text{Anfang}(\gamma_2)$ und $\text{Ende}(\gamma_1) = \text{Ende}(\gamma_2)$ gilt:

$$\int_{\gamma_1} V \;=\; \int_{\gamma_2} V$$

(c) Für alle **geschlossenen $\mathcal{C}^1_{st}$-Wege** γ, d. h. $\text{Anfang}(\gamma) = \text{Ende}(\gamma)$, gilt:

$$\int_{\gamma} V \;=\; 0$$

Beweis:

▷ (a) $\Rightarrow$ (b): Dies folgt direkt aus Satz 4.1.4.

▷ (b) $\Rightarrow$ (c): Sei γ geschlossen, also $\text{Anfang}(\gamma) = \text{Ende}(\gamma) = p$. Zudem sei $\gamma_1 : [0,1] \to U$ der konstante Weg, $\gamma_1(t) = p$ für alle t. Also ist $\text{Anfang}(\gamma) = \text{Anfang}(\gamma_1)$, $\text{Ende}(\gamma) = \text{Ende}(\gamma_1)$ und außerdem $\gamma_1'(t) = 0$ für alle t. Damit folgt nach (b):

$$\int_{\gamma} V \;=\; \int_{\gamma_1} V \;=\; \int_0^1 \langle V(\gamma_1(t)),\, 0\rangle \, dt \;=\; \int_0^1 0 \, dt \;=\; 0$$

▷ (c) $\Rightarrow$ (b): Seien $\gamma_1 : [a,b] \to U$, $\gamma_2 : [c,d] \to U$ Wege, deren Anfangspunkte sowie deren Endpunkte übereinstimmen. Sei γ so, dass zunächst γ_1 vorwärts und anschließend γ_2 rückwärts durchlaufen wird. Dann ist γ ein geschlossener Weg und es folgt:

$$0 \;=\; \int_{\gamma} V \;=\; \int_{\gamma_1} V - \int_{\gamma_2} V \quad\Longrightarrow\quad \int_{\gamma_1} V \;=\; \int_{\gamma_2} V$$

Es bleibt zu zeigen, dass (b) $\Rightarrow$ (a) gilt. Dazu führen wir zunächst einen nützlichen Begriff ein. □

4.1.6 Definition

Ein metrischer Raum (X, d) heißt **wegzusammenhängend**, falls für alle $p, q \in X$ eine stetige Kurve $\gamma : I \to X;\ p \rightsquigarrow q$ existiert.

Man kann jeden metrischen Raum in maximale wegzusammenhängende Teile zerlegen: Definiert man $p \sim q :\Leftrightarrow$ Es gibt ein stetiges $\gamma : I \to X;\ p \rightsquigarrow q$, so erhält man eine Äquivalenzrelation. Die Äquivalenzklassen davon nennt man **Wegzusammenhangskomponenten** von X. Man könnte also auch sagen, dass X wegzusammenhängend ist, falls X nur genau eine Wegzusammenhangskomponente hat (nämlich sich selbst).

Man sieht leicht: Ist $U \subset \mathbb{R}^n$ offen und wegzusammenhängend, so kann man je zwei Punkte in U sogar durch eine $\mathcal{C}^1$-Kurve verbinden.

Beweis (Fortführung von 4.1.5): (b) $\Rightarrow$ (a): O. B. d. A. sei U wegzusammenhängend (sonst führe man die Argumentation auf jeder Wegzusammenhangskomponente einzeln durch). Wir wollen zeigen, dass es ein u mit $\nabla u = V$ gibt. Setzen wir also

$$u(x) = u(p) + \int_{\gamma} V \quad \text{für eine beliebige } \mathcal{C}^1\text{-Kurve } \gamma : p \rightsquigarrow x$$

Aufgrund der Voraussetzung (b) ist dies wohldefiniert. Sei $x \in U$. Wir müssen zeigen, dass $\nabla u(x) = V(x)$ gilt, d.h.

$$\lim_{\|h\|\to 0} \frac{u(x+h) - u(x) - \langle V(x),\, h\rangle}{\|h\|} \;=\; 0 \qquad \text{bzw.} \qquad u(x+h) - u(x) - \langle V(x),\, h\rangle \;=\; \mathrm{o}(h)$$

Sei γ ein Weg von p nach x und $\gamma_1 : [0,1] \to U$ der direkte Verbindungsweg von x zu $x+h$: $\gamma_1(t) = x + th$. Für $\|h\|$ genügend klein verläuft γ_1 ganz in U. Da man von p nach $x+h$ gelangt, indem man erst γ und dann γ_1 durchläuft, gilt

$$u(x+h) = u(p) + \int_\gamma V + \int_{\gamma_1} V = u(x) + \int_{\gamma_1} V \quad \Longrightarrow \quad u(x+h) - u(x) = \int_{\gamma_1} V$$

Es gilt $\gamma_1'(t) = h \;\forall t$, also $\int_{\gamma_1} V = \int_0^1 \langle V(x+th),\, h\rangle$. Aufgrund der Linearität des Skalarprodukts und $\langle V(x),\, h\rangle = \int_0^1 \langle V(x),\, h\rangle\, dt$ folgt also

$$u(x+h) - u(x) - \langle V(x),\, h\rangle = \int_0^1 \langle V(x+th),\, h\rangle\, dt - \langle V(x),\, h\rangle = \int_0^1 \langle V(x+th) - V(x),\, h\rangle\, dt\,.$$

Um zu zeigen, dass dies nach Teilen durch $\|h\|$ gegen 0 geht, verwenden wir die Stetigkeit von V:

$$\forall_{\varepsilon>0}\; \exists_{\delta>0}\; \forall_{z:\|z\|<\delta} \; \|V(x+z) - V(x)\| < \varepsilon$$

Für $\|h\| < \delta$ und $t \in [0,1]$ gilt $\|th\| < \delta$ und daher mit der Cauchy-Schwarzschen Ungleichung:

$$|\langle V(x+th) - V(x),\, h\rangle| \;\le\; \|V(x+th) - V(x)\| \cdot \|h\| < \varepsilon \cdot \|h\|$$

Zusammengefasst heißt das:

$$\|h\| < \delta \quad \Longrightarrow \quad |u(x+h) - u(x) - \langle V(x),\, h\rangle| < \varepsilon \cdot \|h\|$$

Da dies für alle $\varepsilon > 0$ (insbesondere für beliebig kleine) gilt, folgt die Behauptung. □

Beispiel (Physikalische Bedeutung von Kurvenintegralen): Nehmen wir an, V beschreibt ein Kraftfeld und γ einen Weg im Raum, der in einem gewissen Zeitraum von einem Teilchen durchlaufen wird (in der Physik schreibt man meist F statt V). Das Wegintegral von V entlang γ beschreibt dann die Arbeit, die an dem Teilchen auf seinem Weg vom Kraftfeld verrichtet wird. Diese verrichtete Arbeit äußert sich in der Änderung der Energie des Teilchens. Ist sie positiv, hat das Teilchen am Ende mehr Energie, sonst weniger. Es ist eine physikalische Tatsache, dass das uns umgebende Gravitationsfeld ein Gradientenfeld ist, damit gelten alle gerade gezeigten Folgerungen. Betrachtet man etwa (unter Vernachlässigung von anfallender Reibung) einen Fahrradfahrer im Gravitationsfeld G so sieht man:

- ▷ Fährt er auf einem horizontalen Weg γ_1, so ist $\gamma_1'(t) \perp G(\gamma_1(t))$ für alle t, da G senkrecht nach unten zeigt, also $\displaystyle\int_{\gamma_1} G = 0$: Die Energie ändert sich nicht, z.B. behält der Radfahrer seine Geschwindigkeit bei, wenn wir Reibung vernachlässigen.

- ▷ Fährt er einem Weg γ_2 nach oben, wobei der Weg einen Winkel $\varphi > \pi/2$ mit G bildet, so ist $\langle G(t),\, \gamma_2'(t)\rangle = \|G(t)\| \cdot \|\gamma_2'(t)\| \cdot \cos\varphi < 0$ für alle t. Das Integral ist damit auch negativ, d.h. der Radfahrer verliert Energie: er wird langsamer. Will er seine Geschwindigkeit beibehalten, muss er Energie aufwenden.

- ▷ Sei γ der Weg, der durch das Hintereinander-Durchlaufen von γ_1 und γ_2 entsteht: erst geradeaus, dann aufwärts. Sei $\tilde\gamma$ der Weg mit selbem Anfangs- und Endpunkt wie γ, der aber mit konstanter Steigung ansteigt. Da G ein Gradientenfeld ist, verliert der Radfahrer bei beiden Wegen γ und $\tilde\gamma$ dieselbe Energie, d.h. er muss genauso viel Energie aufwenden, um seine Geschwindigkeit beizubehalten.

- ▷ Fährt er den gleichen Weg wieder zurück, erhält er die beim Hochfahren verbrauchte Energie komplett wieder zurück.

Die Integrabilitätsbedingung ist lokal hinreichend

Wir wenden uns nun der Frage zu, ob es eine Umkehrung zu Satz 4.1.1 gibt.

4.1.7 Satz

Sei $U \subset \mathbb{R}^n$ offen und $V : U \to \mathbb{R}^n$ ein $\mathcal{C}^1$-Vektorfeld, für das gilt:

$$\frac{\partial V_i}{\partial x_j} = \frac{\partial V_j}{\partial x_i} \quad \text{für alle } i,j \quad (*)$$

Dann hat jeder Punkt $p \in U$ eine Umgebung U', auf der eine $\mathcal{C}^2$-Funktion $u : U' \to \mathbb{R}$ mit $V = \nabla u$ existiert (»V ist *lokal* ein Gradientenfeld«).

Genauer gilt dies für jede sternförmige Menge $U' \subset U$. Eine Menge $U' \subset \mathbb{R}^n$ heißt dabei **sternförmig**, wenn es ein $p \in U'$ gibt, so dass für alle $x \in U'$ die ganze Strecke $\overline{px}$ in U' enthalten ist.

Beweis: Wir führen einen konstruktiven Beweis: Ohne Beschränkung der Allgemeinheit sei $p = 0$. Sei U' eine sternförmige, in U enthaltene Menge mit $0 \in U'$. Zu $x \in U'$ setze

$$\gamma : [0,1] \to U';\ \gamma(t) = tx \quad \text{und} \quad u(x) = \int_\gamma V = \int_0^1 \langle V(tx),\ x\rangle\, dt\,.$$

Dies ist definiert, da U' sternförmig ist und damit der Weg γ ganz in U' enthalten ist. Wir müssen zeigen, dass $\nabla u(x) = V(x)$ gilt. Leitet man u nach x_i ab, ergibt sich:

$$\frac{\partial u}{\partial x_i}(x) = \frac{\partial}{\partial x_i}\int_0^1 \langle V(tx),\ x\rangle\, dt = \int_0^1 \frac{\partial}{\partial x_i}\langle V(tx),\ x\rangle\, dt$$

(Das Vertauschen von Ableitung und Integral ist erlaubt, da der Integrand C^1 ist und über eine kompakte Menge integriert wird, siehe Korollar 1.5.6.)
Wir formen nun zunächst den Ausdruck unter dem Integral um. Dabei sehen wir, dass er sich erstaunlicherweise als eine Ableitung nach t darstellen lässt:

$$\frac{\partial}{\partial x_i}\langle V(tx),\ x\rangle = \sum_{j=1}^n \frac{\partial}{\partial x_i}(V_j(tx)x_j) = \sum_{j=1}^n \left[t\frac{\partial V_j}{\partial x_i}(tx)x_j + V_j(tx)\frac{\partial x_j}{\partial x_i}\right]$$

$$\overset{*}{=} t\cdot\sum_{j=1}^n \frac{\partial V_i}{\partial x_j}(tx)x_j + V_i(tx) = t\cdot\langle \nabla V_i(tx),\ x\rangle + V_i(tx) = \frac{\partial}{\partial t}(t\cdot V_i(tx))$$

Setzt man dies nun ein, erhält man mit dem Hauptsatz der Differential- und Integralrechnung folgendes:

$$\int_0^1 \frac{\partial}{\partial x_i}\langle V(tx),\ x\rangle\, dt = \int_0^1 \frac{\partial}{\partial t}(t\cdot V_i(tx))\, dt = 1\cdot V_i(1\cdot x) - 0\cdot V_i(0\cdot x) = V_i(x)$$

Insgesamt liefert das genau die Behauptung:

$$\frac{\partial u}{\partial x_i} = V_i(x) \iff \nabla u = V$$

□

Dass trotz der Integrabilitätsbedingung eine Stammfunktion nicht auf ganz U existieren muss, zeigt folgendes Beispiel.

Beispiel: $V(x,y) = \left(\frac{-y}{x^2+y^2},\ \frac{x}{x^2+y^2}\right) = \frac{1}{x^2+y^2}\cdot(-y,\ x)$ erfüllt die Integrabilitätsbedingung (nachrechnen!). Es gibt aber *global* kein $u : \mathbb{R}^2 \setminus \{(0,0)\} \to \mathbb{R}$ mit $V = \nabla u$.

Beweis: Angenommen, es gäbe doch so ein u. Sei

$$\gamma_0 : [0,2\pi] \to \mathbb{R}^2;\ \gamma_0(t) = (\cos t, \sin t)\,.$$

Offenbar ist γ_0 ein geschlossener Weg, also müsste $\int\limits_{\gamma_0} V = 0$ sein! Nachrechnen ergibt aber:

$$\int\limits_{\gamma_0} V = \int_0^{2\pi} \langle(-\sin t, \cos t),\ (-\sin t, \cos t)\rangle\, dt = \int_0^{2\pi} (\sin^2 t + \cos^2 t)\, dt = \int_0^{2\pi} 1\, dt = 2\pi \neq 0$$

Nach Satz 4.1.7 existiert u aber zumindest lokal, in diesem Fall zum Beispiel auf $\mathbb{R}^2$ ohne eine *Halbgerade*, z.B. auf $U' = \mathbb{R}^2 \setminus \{(x,0) : x \geq 0\}$, denn U' ist sternförmig bezüglich des Punktes $p = (-1,0)$. Wie findet man eine Stammfunktion?
Wie im Beweis von Satz 4.1.7 können wir $u(q) = \int\limits_{\gamma: p \rightsquigarrow q} V$ wählen. Da wir nach Satz 4.1.5 wissen, dass das Wegintegral über V nur von Anfangs- und Endpunkt des Weges abhängt, steht es uns frei, γ so zu wählen, dass sich besonders leicht integrieren lässt. Sei $q' = \frac{q}{\|q\|}$ der Einheitsvektor in derselben Richtung wie q. Wir wählen γ so, dass wir erst entlang einem Weg γ_1 auf dem Einheitskreis von p nach q' gehen und dann auf einem Weg γ_2 radial von q' nach q. Da γ_2' überall orthogonal zu V steht, ist $\int\limits_{\gamma_2} V = 0$. Den Weg γ_1 erhalten wir als Einschränkung des oben verwendeten γ_0: Sei φ der Polarkoordinatenwinkel von q', also $\gamma_0(\varphi) = q'$. Wegen $p = \gamma_0(\pi)$ folgt wie in der Rechnung oben

$$u(q) = \int_{\gamma_1} V = \int_\pi^\varphi 1\, dt \ = \ \varphi - \pi$$

Da man zu einer Stammfunktion beliebige Konstanten addieren kann, erhalten wir:

Die Funktion $q \mapsto \varphi(q)$ (= Winkel in der Polarkoordinatendarstellung von q) ist eine Stammfunktion von V.

Dies erklärt auch, warum V keine Stammfunktion auf ganz $\mathbb{R}^2 \setminus 0$ hat: Diese müsste die Form $\varphi + \text{const}$ haben, doch dies hat einen Sprung, wenn man den Nullpunkt einmal umkreist.

Globale Bedingungen

Frage: Welche Bedingungen muss man zusätzlich zur Integrabilitätsbedingung stellen, damit ein Vektorfeld *global*, d.h. auf seinem ganzen Definitionsbereich, ein Gradientenfeld ist?

Exemplarisch betrachten wir $\mathbb{R}^2 \setminus 0$.

4.1.8 Satz

Sei $U = \mathbb{R}^2 \setminus 0$ und V ein C^1-Vektorfeld auf U. Dann sind äquivalent:

(1) V ist ein Gradientenfeld, d.h. es gibt ein $u \in C^2(U)$ mit $V = \nabla u$;

(2) $\frac{\partial V_1}{\partial y} = \frac{\partial V_2}{\partial x}$ auf U und

$$\int_{\gamma_0} V = 0$$

für den Weg $\gamma_0 : [0.2\pi] \to U,\ t \mapsto (\cos t, \sin t)$.

Diese Bedingungen sind effektiv nachprüfbar, etwa im Unterschied zu der Bedingung (c) in Satz 4.1.5, wo man die Integrale über alle Wege betrachten muss.

Beweis: Seien $X_\pm = \{(x,0) : \pm x \geq 0\}$ die positive und negative x-Achse und $U_\pm = \mathbb{R}^2 \setminus X_\pm$. Die Teilmengen U_+ und U_- von U sind sternförmig, also existieren C^2-Funktionen $u_\pm$ auf $U_\pm$ mit $\nabla u_\pm = V$ auf $U_\pm$. Wir wollen zeigen, dass man u_+ und u_- zu einer Funktion auf ganz u ‚zusammenkleben' kann.

Der Schnitt $U_+ \cap U_- = \mathbb{R}^2 \setminus \{x - \text{Achse}\}$ hat zwei Zusammenhangskomponenten: $H_> = \{y > 0\}$ und $H_< = \{y < 0\}$. Die Funktion $v = u_+ - u_-$ ist auf diesem Schnitt definiert, und dort gilt $\nabla v = \nabla u_+ - \nabla u_- = V - V = 0$, also ist $u_+ - u_-$ auf $H_>$ und auf $H_<$ konstant:

$$u_+ - u_- = c_> \text{ auf } H_>, \quad u_+ - u_- = c_< \text{ auf } H_<$$

Mit den bisher verwendeten Informationen wissen wir nichts darüber, ob die Konstanten $c_>$ und $c_<$ gleich sind oder nicht.

Behauptung: Aus $\int_{\gamma_0} V = 0$ folgt, dass $c_> = c_<$ gilt.

Beweis: Sei $p_> = (0,1)$ und $p_< = (0,-1)$. Wir können von $p_<$ nach $p_>$ entweder auf dem linken Halbkreis gehen (nenne diesen Weg γ_+) oder auf dem rechten Halbkreis (nenne diesen Weg γ_-). Der Weg γ_+ liegt in U_+, also gilt $\int_{\gamma_+} V = u_+(p_>) - u_+(p_<)$. Analog liegt γ_- in U_-, also folgt $\int_{\gamma_-} V = u_-(p_>) - u_-(p_<)$. Da γ_0 aus γ_- und γ_+ (rückwärts durchlaufen) zusammengesetzt ist (bis auf eine Reparametrisierung) und da $\int_{\gamma_0} V = 0$ gilt, folgt $\int_{\gamma_+} V = \int_{\gamma_-} V$, also

$$u_+(p_>) - u_+(p_<) = u_-(p_>) - u_-(p_<)$$

Umstellen ergibt $u_+(p_>) - u_-(p_>) = u_+(p_<) - u_-(p_<)$. Wegen $p_> \in H_>$, $p_< \in H_<$ folgt $c_> = c_<$, was zu zeigen war.

Setze $c = c_> = c_-$ und

$$u = \begin{cases} u_+ & \text{auf } U_+ \\ u_- + c & \text{auf } U_- \end{cases}$$

Wegen $u_+ = u_- + c$ auf $U_+ \cap U_-$ ist dies auf ganz $U = U_- \cup U_+$ wohldefiniert. Weiterhin ist $\nabla u = V$ auf U, denn auf U_+ folgt dies aus $\nabla u_+ = V$ und auf U_- aus $\nabla(u_- + c) = \nabla u_- = V$. Daher ist V ein Gradientenfeld. □

Statt γ_0 kann man auch andere Wege nehmen, die den Zweck im Beweis erfüllen oder homotop zu γ_0 sind, siehe unten.

> **Bemerkung:** Satz 4.1.8 hat weitreichende Verallgemeinerungen. Sind etwa $p_1, \dots, p_N \in \mathbb{R}^2$ verschiedene Punkte der Ebene, so ist ein C^1-Vektorfeld V auf $\mathbb{R}^2 \setminus \{p_1, \dots, p_N\}$ genau dann ein Gradientenfeld, wenn es die Integrabilitätsbedingung erfüllt und wenn $\int_{\gamma_i} V = 0$ für $i = 1, \dots, N$ gilt, wobei γ_i einen Kreis mit Mittelpunkt p_i durchläuft, der keinen der anderen p_j in seinem Innern oder auf der Kreislinie enthält.
>
> Dies lässt sich ähnlich beweisen. In angemessener Allgemeinheit (auch in höheren Dimensionen) lässt sich der Satz aber am besten mit der Sprache der algebraischen Topologie und der Differentialformen formulieren und beweisen (Satz von de Rham).

Die zunächst etwas unmotiviert erscheinende Rechnung im Beweis von Satz 4.1.7 wird in etwas größerer Allgemeinheit transparenter und hat dann weitere interessante Konequenzen. Wir betrachten eine **parametrisierte Familie von Kurven** in $U \subset \mathbb{R}^n$, d.h. eine Abbildung $\gamma : I \times [a,b] \to U$ für ein Intervall I. Die Variablen bezeichnen wir mit $s \in I$, $t \in [a,b]$. Für jedes $s \in I$ ist $\gamma(s,\cdot)$ eine Kurve in U. Wenn s variiert, variiert die Kurve. Wir berechnen, wie sich das Kurvenintegral mit s ändert:

4.1.9 Satz

Sei $\gamma : I \times [a,b] \to U$ eine C^2-Abbildung. Sei $V : U \to \mathbb{R}^n$ ein C^1-Vektorfeld, das die Integrabilitätsbedingung $\frac{\partial V_i}{\partial x_j} = \frac{\partial V_j}{\partial x_i}$ $\forall i,j$ erfüllt. Dann gilt für alle s

$$\frac{d}{ds} \int\limits_{\gamma(s,\cdot)} V = \langle V(\gamma(s,\cdot)), \partial_s \gamma(s,\cdot) \rangle |_{t=a}^{t=b}$$

Beweis: Der Übersichtlichkeit halber lassen wir die Argumente von γ in der Rechnung weg. Die Behauptung folgt sofort aus der Identität

$$\partial_s \sum_i V_i(\gamma)\partial_t\gamma_i = \partial_t \sum_j V_j(\gamma)\partial_s\gamma_j\,,$$

wenn man dies über $t \in [a,b]$ integriert und rechts den Hauptsatz anwendet (alle Summen gehen von 1 bis n). Diese Identität ist leicht nachzuprüfen: Mit der Kettenregel erhält man

$$\text{linke Seite } = \sum_i \sum_j \frac{\partial V_i}{\partial x_j}(\gamma)\partial_s\gamma_j\partial_t\gamma_i + \sum_i V_i(\gamma)\partial_s\partial_t\gamma_i$$

$$\text{rechte Seite } = \sum_j \sum_i \frac{\partial V_j}{\partial x_i}(\gamma)\partial_t\gamma_i\partial_s\gamma_j + \sum_j V_j(\gamma)\partial_t\partial_s\gamma_j$$

Wegen $\frac{\partial V_i}{\partial x_j} = \frac{\partial V_j}{\partial x_i}$ und $\partial_s\partial_t\gamma = \partial_t\partial_s\gamma$ sind beide Ausdrücke gleich. □

Die Rechnung im Beweis von Satz 4.1.7 ist ein Spezialfall davon: Für festes i und feste $x_1,\dots,\hat{x_i},\dots,x_n$ war dort $\gamma(x_i,t) = tx$.

Eine viel weitreichendere Anwendung betrifft die Invarianz des Kurvenintegrals unter Veränderung des Weges. Zunächst ein Begriff:

4.1.10 Definition

Sei $U \subset \mathbb{R}^n$ und $\gamma_0,\gamma_1 : [a,b] \to U$ C^2-Kurven mit $\gamma_0(a) = \gamma_1(a)$ und $\gamma_0(b) = \gamma_1(b)$. Eine C^2-Abbildung $\gamma : [0,1]\times[a,b] \to U$ heißt **Homotopie** von γ_0 nach γ_1, falls gilt:

(1) $\gamma_0 = \gamma(0,\cdot)$ und $\gamma_1 = \gamma(1,\cdot)$

(2) $\gamma(s,a) = \gamma_0(a)$ und $\gamma(s,b) = \gamma_0(b)$ für alle $s \in [0,1]$.

γ_0,γ_1 heißen **homotop**, wenn es eine Homotopie von γ_0 nach γ_1 gibt. Man schreibt dann $\boldsymbol{\gamma_0 \simeq \gamma_1}$.

$\gamma_0 \simeq \gamma_1$ bedeutet also, dass man die Kurve γ_0 in glatter (C^2) Weise in die Kurve γ_1 deformieren kann, wobei Anfangs- und Endpunkt festgehalten werden. (Man sagt auch glatt homotop, denn in manchen Kontexten ist es günstiger, von einer Homotopie nur Stetigkeit zu verlangen.) Es ist leicht zu zeigen, dass dies eine Äquivalenzrelation auf der Menge der Kurven definiert.

Wesentlich ist, dass die Homotopie nur Werte in U annimmt. Zum Beispiel ist anschaulich klar, dass die Kurven γ_- und γ_+ im Beweis von Satz 4.1.8 in $\mathbb{R}^2$ homotop sind, in $U = \mathbb{R}^2 \setminus 0$ aber nicht. Wie zeigt man das exakt? Mit Hilfe von:

4.1.11 Korollar (Homotopieinvarianz des Kurvenintegrals unter Integrabilitätsbedingung)

Sei $U \subset \mathbb{R}^n$ offen und seien $\gamma_0,\gamma_1 : [a,b] \to U$ zwei C^2-Kurven mit $\gamma_0(a) = \gamma_1(a)$ und $\gamma_0(b) = \gamma_1(b)$. Sei V ein Vektorfeld auf U, das die Integrabilitätsbedingung erfüllt. Dann gilt

$$\int_{\gamma_0} V = \int_{\gamma_1} V$$

Beweis: Sei γ eine Homotopie von γ_0 nach γ_1. Sei $\gamma_s = \gamma(s,\cdot)$ und $f(s) = \int\limits_{\gamma_s} V$. Nach Satz 4.1.9 ist für jedes s

$$\frac{d}{ds}f(s) = \langle V(\gamma_s), \partial_s\gamma_s\rangle\big|_{t=a}^{t=b}$$

und dies ist gleich Null, da $\gamma_s(a)$ und $\gamma_s(b)$ als Funktionen von s konstant sind, also $\partial_s\gamma_s(a) = \partial_s\gamma_s(b) = 0$ für alle s gilt. Also ist $f' \equiv 0$, also f konstant, also $f(0) = f(1)$. □

Bemerkung: Für geschlossene Wege gilt das Korollar auch, wenn die Bedingung (b) in Definition 4.1.10 durch die Bedingung ersetzt wird, dass $\gamma(s,\cdot)$ für jedes s ein geschlossener Weg ist, d.h. $\gamma_s(a) = \gamma_s(b)$ für alle s. Denn dann ist $\langle V(\gamma_s), \partial_s\gamma_s\rangle|_{t=a}^{t=b} = \langle V(\gamma_s(b)), \partial_s\gamma_s(b)\rangle - \langle V(\gamma_s(a)), \partial_s\gamma_s(a)\rangle = 0$.

Beispiel: Die Kurven γ_- und γ_+ im Beweis von Satz 4.1.8 sind nicht homotop in $\mathbb{R}^2 \setminus 0$, da für das integrable Vektorfeld $V(x,y) = \left(\frac{-y}{x^2+y^2}, \frac{x}{x^2+y^2}\right)$ gilt, dass $\int_{\gamma_-} V \neq \int_{\gamma_+} V$ (denn die Differenz ist 2π nach der Rechnung oben).

Damit können wir folgende Frage beantworten: Für welche Gebiete $U \subset \mathbb{R}^n$ gilt die Aussage von Satz 4.1.7 global?

4.1.12 Definition

$U \subset \mathbb{R}^n$ heißt **einfach zusammenhängend**, falls jeder geschlossene Weg in U homotop zu einem konstanten Weg ist.

Zum Beispiel ist $\mathbb{R}^n$ einfach zusammenhängend: Ist $\gamma_1 : [a,b] \to \mathbb{R}^n$ geschlossen mit $\gamma_1(a) = \gamma_1(b) = 0$, so setze $\gamma(s,t) = s\gamma_1(t)$. Dies ist eine Homotopie nach $\gamma_0 \equiv 0$. Ist der Anfangs- und Endpunkt nicht der Nullpunkt, geht es ähnlich. Analog sieht man, dass sternförmige Mengen einfach zusammenhängend sind.

4.1.13 Korollar

Sei $U \subset \mathbb{R}^n$ offen und einfach zusammenhängend. Dann ist jedes C^1-Vektorfeld auf U, das die Integrabilitätsbedingung erfüllt, ein Gradientenfeld auf ganz U.

Beweis: Wir müssen zeigen, dass $\int_{\gamma_0} V = 0$ für jeden geschlossenen Weg in U gilt. Nun ist γ_0 zu einem konstanten Weg γ_1 homotop. Nach Korollar 4.1.11 ist $\int_{\gamma_0} V = \int_{\gamma_1} V$. Das Kurvenintegral über einen konstanten Weg ist immer gleich Null, damit folgt die Behauptung. □

Beispiele:

Sternförmige Gebiete sind einfach zusammenhängend, daher erhalten wir erneut Satz 4.1.7.

$\mathbb{R}^2 \setminus 0$ ist nicht einfach zusammenhängend. Das ist zwar anschaulich klar, aber direkt mit der Definition schwer zu zeigen. Es folgt nun daraus, dass das Vektorfeld $V(x,y) = \left(\frac{-y}{x^2+y^2}, \frac{x}{x^2+y^2}\right)$ die Integrabilitätsbedingung erfüllt, aber kein Gradientenfeld ist.

Zusammenfassung zur Frage, wie man einem Vektorfeld V ansieht, ob es ein Gradientenfeld ist:

Erste Antworten darauf geben die Sätze 4.1.5 und 4.1.7. Die Bedingung (b) oder (c) in Satz 4.1.5 ist zwar notwendig und hinreichend, aber sie ist praktisch schwer nachzuprüfen. Dagegen ist die Integrabilitätsbedingung in Satz 4.1.7 notwendig, aber nur lokal hinreichend – dafür aber durch einfache Rechnung nachzuprüfen.

Daher sollte man so vorgehen:

- ▷ Zunächst prüft man nach, ob V die Integrabilitätsbedingung erfüllt. Wenn ja, ist es zumindest lokal ein Gradientenfeld.
- ▷ Ist zusätzlich U einfach zusammenhängend (hat also keine ‚Löcher') – z.B. sternförmig –, so ist V nach Korollar 4.1.13 ein Gradientenfeld.
- ▷ Sonst wählt man geschlossene Wege um jedes ‚Loch' in U und prüft nach, ob das Integral von V über diese Wege gleich Null ist. Ist dies nicht der Fall, ist V kein Gradientenfeld.

▷ Ist aber jedes dieser Integrale gleich Null, so folgt, dass dann das Integral von V über *jeden* geschlossenen Weg verschwindet und daher V ein Gradientenfeld ist. Dies wurde examplarisch in Satz 4.1.8 bewiesen, im Allgemeinen muss man natürlich zunächst definieren, was ‚Loch' bedeutet.

Noch eine abschließende Bemerkung: Die genauen Regularitätsanforderungen (C^1,C^2 etc.) in diesem Kapitel sind ziemlich irrelevant. Sie wurden immer so gewählt, dass die gemachten Aussagen sinnvoll waren. Daher sollten Sie sie beim ersten Lernen ignorieren. Das wirklich Interessante passiert woanders!

4.2 Mannigfaltigkeiten mit Rand

Die offene Kreisscheibe $\{x \in \mathbb{R}^2 : \|x\| < 1\}$ ist eine Mannigfaltigkeit, die abgeschlossene Kreisscheibe $\{x \in \mathbb{R}^2 : \|x\| \leq 1\}$ aber nicht, da Randpunkte keine Umgebungen V besitzen, für die eine lokale Karte $\varphi : U \to V$ mit einer offenen Teilmenge $U \subset \mathbb{R}^2$ existiert. (Warum nicht? Übung.)

Wohl aber gibt es eine Karte, wenn wir für U offene Teilmengen der Halbebene zulassen. Daher diskutieren wir diesen Begriff zuerst.

Vorbemerkung zum Begriff des Halbraums
Sei $\mathbb{R}^n_+ = \{(x_1, \ldots, x_n) \in \mathbb{R}^n : x_n \geq 0\}$ der **Halbraum**. Sein Rand ist $\partial\mathbb{R}^n_+ = \{x \in \mathbb{R}^n : x_n = 0\}$, kann also mit $\mathbb{R}^{n-1}$ identifiziert werden.

Erinnerung: $U \subset \mathbb{R}^n_+$ ist offen in $\mathbb{R}^n_+ \Leftrightarrow \exists W \subset \mathbb{R}^n$ offen mit $U = W \cap \mathbb{R}^n_+$

Sei $U \subset \mathbb{R}^n_+$ offen in $\mathbb{R}^n_+$ und $\varphi : U \to \mathbb{R}^N$. Dann ist definiert, was Differenzierbarkeit und das Differential von φ bedeutet, auch bei Randpunkten $p \in U \cap \partial\mathbb{R}^n_+$:

$$\lim_{h \to 0, h \in U} \frac{\varphi(p+h) - \varphi(p) - D\varphi_{|p}(h)}{\|h\|} = 0$$

Damit ist auch der Begriff einer Immersion auf U definiert (Injektivität von $D\varphi_{|p}$ für jedes $p \in U$).

Bemerkung: Ursprünglich hatten wir $D\varphi_{|p}$ nur für Punkte p definiert, die im *Innern* des Definitionsbereiches von φ liegen. Dies war deshalb wichtig, damit $D\varphi_{|p}$, wenn es existiert, durch die angegebene Definition eindeutig bestimmt ist. Man sieht aber leicht, dass diese Eindeutigkeit auch im vorliegenden Fall stimmt, denn für jedes genügend kleine h ist entweder $h \in U$ oder $-h \in U$.

4.2.1 Definition

Eine Teilmenge $M \subset \mathbb{R}^N$ heißt **n-dimensionale Untermannigfaltigkeit mit Rand** $:\Leftrightarrow \forall p \in M$ existiert eine Umgebung V von p in M, eine in $\mathbb{R}^n_+$ offene Menge U und eine **lokale Karte** $\varphi : U \to V$ (d.h. φ ist Immersion und Homöomorphismus).

Beispiele: (1) $n = 1$: Jedes Intervall I ist eine eindimensionale Untermannigfaltigkeit mit Rand von $\mathbb{R}$. Außerdem ist $I \times \{a\}$ für beliebiges $a \in \mathbb{R}$ eine eindimensionale Untermgfk mit Rand von $\mathbb{R}^2$.

(2) $n = 2$: $\{x \in \mathbb{R}^2 : \|x\| \leq 1\}$ ist eine Umgfk mit Rand von $\mathbb{R}^2$. Eine lokale Karte nahe einem Randpunkt p erhält man z.B. mittels Polarkoordinaten: Ist $p = (\cos\theta_0, \sin\theta_0)$, so ist

$$\varphi(\theta, s) = ((1-s)\cos\theta, (1-s)\sin\theta),$$

definiert auf $U = (\theta_0 - \pi, \theta_0 + \pi) \times [0, \frac{1}{2})$, eine lokale Karte. Der Rand $s = 0$ von U entspricht genau dem Rand der Kreisscheibe.

(3) das Quadrat $Q = \{(x,y) \in \mathbb{R}^2 : 0 \le x,y \le 1\}$ ist *keine* Mgfk. mit Rand (es ist eine »Mannigfaltigkeit mit Ecken«). $Q \setminus \{(0,0),(0,1),(1,0),(1,1)\}$ ist aber eine Mgfk mit Rand.

4.2.2 Definition

Sei $M \subset \mathbb{R}^N$ eine Untermannigfaltigkeit mit Rand. Der **Rand ∂M** von M ist
$\partial M = \{p \in M :$ es existiert eine lokale Karte $\varphi : U \to V$ für p mit $\varphi^{-1}(p) \in U \cap \partial\mathbb{R}^n_+\}$

Bemerkung: Vorsicht: Falls $\dim M < N$, dann ist ∂M nicht der Rand im topologischen Sinn!

Beispiel: ▷ $M = [0,1] \times \{0\} \subset \mathbb{R}^2$. Dies ist eine 1-dimensionale Untermannigfaltigkeit mit Rand von $\mathbb{R}^2$, und $\partial M = \{(0,0),(1,0)\}$. Der topologische Rand von M ist aber $\overline{M} \setminus \overset{\circ}{M} = M$.

▷ Eine Mannigfaltigkeit ist eine Mannigfaltigkeit mit Rand, deren Rand leer ist!

4.2.3 Satz

Sei $M \subset \mathbb{R}^N$ n-dimensionale Untermannigfaltigkeit mit Rand. Dann ist ∂M eine $(n-1)$-dimensionale Untermannigfaltigkeit (ohne Rand).

Beweis (Beweisskizze): Lokale Karten für ∂M : $\tilde{\varphi} = \varphi_{|_{U \cap \partial\mathbb{R}^n_+}}$, für lokale Karten $\varphi : U \to V$ von M. Dass $\tilde{\varphi}$ eine Immersion ist, ist einfach: $D\tilde{\varphi}_{|x} = (D\varphi_{|x})_{|x_n=0}$, und mit $D\varphi_{|x}$ ist auch seine Einschränkung auf einen Unterraum injektiv. Um zu zeigen, dass $\tilde{\varphi}$ ein Homöomorphismus ist, muss man ein wenig arbeiten. Im Wesentlichen muss man zeigen, dass aus $p \in \partial M$ folgt, dass für *jede* lokale Karte $\psi : U' \to V'$ mit $p \in V'$ gilt, dass $\psi^{-1}(p) \in U' \cap \partial\mathbb{R}^n_+$. Siehe die Literatur. □

Wichtiger Spezialfall:
Die n-dimensionalen Untermannigfaltigkeiten des $\mathbb{R}^n$ sind die offenen Mengen, vgl. die Bemerkung am Ende von Abschnitt 3.4.
Die n-dimensionalen Untermannigfaltigkeiten *mit Rand* von $\mathbb{R}^n$ nennt man auch **Gebiete mit glatten Rand**. (Siehe die Beispiele oben.)

Tangentialraum, Normalvektor einer Mannigfaltigkeit mit Rand

4.2.4 Definition

Sei $M \subset \mathbb{R}^N$ eine Untermannigfaltigkeit mit Rand. Sei $p \in M$. Der **Tangentialraum an M im Punkt p** ist definiert als

$$T_pM = \{v \in \mathbb{R}^N : \exists\, \varepsilon > 0 \text{ und eine } C^1\text{-Kurve } \gamma : [0,\varepsilon) \to M \text{ mit } \gamma(0) = p, \gamma'(0) = v \text{ oder } -v\}$$

T_pM enthält also die Vektoren, die bei p an M tangential sind, und deren Negative.

Beachte: Im Unterschied zu Definition 3.2.1 braucht die Kurve hier von p nur in einer Richtung loszulaufen. Zum Ausgleich darf der Tangentialvektor entweder $\gamma'(0)$ oder $-\gamma'(0)$ sein. Das bewirkt folgendes:

▷ für einen inneren Punkt p kommt dasselbe heraus wie bei Definition 3.2.1;

▷ für einen Randpunkt p kommt ebenfalls ein n-dimensionaler Vektorraum (nicht bloß ein Halbraum!) heraus.

Für $p \in \partial M$ kann man auch $T_p\partial M$ betrachten. Das sind die Vektoren, die bei p an ∂M tangential sind. Dies ist ein $(n-1)$-dimensionaler Unterraum von T_pM.

Beispiel: ▷ $M = \{x : \|x\| \leq 1\}$ mit $n = 3, N = 3$:
$T_pM = \mathbb{R}^3$ für jedes $p \in M$ und $T_p\partial M = \{v \in \mathbb{R}^3 : v \perp p\}$ für $p \in \partial M$.

▷ $M = \{(x, y, 0) : x^2 + y^2 \leq 1\} \subset \mathbb{R}^3$ mit $n = 2, N = 3$:
$T_pM = \mathbb{R}^2 \times \{0\}$ für jedes $p \in M$ und $T_{(x,y,0)}\partial M = \{(v, w, 0) \in \mathbb{R}^3 : vx + wy = 0\}$ für $(x, y, 0) \in \partial M$, d.h. $x^2 + y^2 = 1$.

▷ (Halbsphäre) $M = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1,\ z \geq 0\}$ mit $n = 2, N = 3$. Übung: Das ist eine Untermannigfaltigkeit mit Rand.

$T_pM = \{v \in \mathbb{R}^3 : v \perp p\}$ für jedes $p \in M$ und $T_p\partial M = \{(v, w, 0) \in \mathbb{R}^3 : vx + wy = 0\}$ für $p = (x, y, 0) \in \partial M$.

4.2.5 Lemma (Äußeres Einheitsnormalenfeld)

Sei $M \subset \mathbb{R}^N$ eine n-dimensionale Untermannigfaltigkeit mit Rand. Dann gibt es eine eindeutige Abbildung $\nu : \partial M \to \mathbb{R}^N$, so dass für alle $p \in \partial M$ gilt:

(a) $\|\nu(p)\| = 1$

(b) $\nu(p) \perp T_p\partial M$ und $\nu(p) \in T_pM$

(c) $\nu(p)$ zeigt nach aussen, d.h. $\nu(p) = -\gamma'(0)$ für eine geeignete Kurve $\gamma : [0, \epsilon) \to M, \gamma(0) = p$

ν ist stetig. Es heißt **äußeres Einheitsnormalenfeld** von M. Falls $n = N$ und $\varphi : U \xrightarrow{\subset \mathbb{R}^n_+} V \subset M$ eine lokale Karte mit $\varphi(x_0) = p$ ist, so gilt

$$\nu(p) = \frac{N(p)}{\|N(p)\|} \text{ oder } \nu(p) = -\frac{N(p)}{\|N(p)\|}, \quad N(p) := \frac{\partial \varphi}{\partial x_1}(x_0) \times \ldots \times \frac{\partial \varphi}{\partial x_{n-1}}(x_0)$$

Beweis: Wegen $T_p\partial M \subset T_pM$ mit $\dim T_p\partial M = n - 1$ und $\dim T_pM = n$ gibt es genau zwei Vektoren, die (a) und (b) erfüllen. Einer davon erfüllt (c).

Wir beweisen nun die Formel im Fall $n = N$. Da $\varphi_{|U \cap \partial \mathbb{R}^n_+}$ eine lokale Karte für $T_p\partial M$ ist, gilt $T_p\partial M = \operatorname{span}\left\{\frac{\partial \varphi}{\partial x_1}(x_0), \ldots, \frac{\partial \varphi}{\partial x_{n-1}}(x_0)\right\}$. Diese Vektoren sind linear unabhängig, da φ eine Immersion ist, also ist ihr Kreuzprodukt $N(p) \neq 0$. Da $N(p)$ senkrecht auf allen $\frac{\partial \varphi}{\partial x_i}(x_0)$ und damit auf $T_p\partial M$ steht, ist es ein Normalenvektor, also ist $N(p)/\|N(p)\|$ ein Einheitsnormalenvektor, und die Formel folgt.

Aus der Formel folgt die Stetigkeit von ν im Fall $n = N$, da $\varphi \in C^1 \Rightarrow \frac{\partial \varphi}{\partial x_i}$ stetig $\Rightarrow N$ stetig $\Rightarrow \nu$ stetig. Im allgemeinen Fall gibt es eine ähnliche Formel (anspruchsvolle Übung), die ebenfalls die Stetigkeit beweist. □

Bemerkung: Es gibt $(n-1)$-dimensionale Untermannigfaltigkeiten des $\mathbb{R}^n$, die kein stetiges Feld von Einheitsnormalenvektoren haben, z.B. das Möbiusband. Eine $(n-1)$-dimensionale Untermannigfaltigkeit $H \subset \mathbb{R}^n$ nennt man auch **Hyperfläche**. Falls ein stetiges Einheitsnormalenfeld $\nu : H \to \mathbb{R}^n$ existiert, so nennt man H **orientierbar**, ν eine **Orientierung** von H und das Paar (H, ν) eine **orientierte Hyperfläche**.

Wir sahen, dass für jedes Gebiet $M \subset \mathbb{R}^n$ mit glattem Rand der Rand ∂M orientierbar ist und das äußere Einheitsnormalenfeld eine Orientierung bildet. Eine andere Orientierung ist durch das innere Einheitsnormalenfeld gegeben. Das Möbiusband ist nicht orientierbar.

4.3 Der Satz von Gauß

Der Satz von Gauß ist eine der wichtigsten Verallgemeinerungen des Hauptsatzes. Für seinen Beweis (und für viele weitere Zwecke) ist ein Hilfsmittel nützlich, die Partition der Eins.

Partitionen der Eins

4.3.1 Definition

Sei (X, d) metrischer Raum, $f : X \to \mathbb{R}$. Der **Träger** (support) von f ist definiert als

$$\operatorname{supp} f := \overline{\{x \in X : f(x) \neq 0\}}$$

Beachte: Man nimmt den Abschluss der Menge, wo f nicht verschwindet. Der Träger ist also per Definition immer abgeschlossen.

Beispiel: $\sin : \mathbb{R} \to \mathbb{R}$ hat Träger $\mathbb{R}$. Die Funktion $f(x) = \begin{cases} 0, & x \leq 0 \\ x, & x > 0 \end{cases}$ hat Träger $[0, \infty)$.

Frage: Gibt es eine Funktion $f : \mathbb{R} \to \mathbb{R}$, die C^∞ (∞ oft differenzierbar) ist und für die $\operatorname{supp} f$ beschränkt, also kompakt ist? (Natürlich soll f nicht konstant null sein.)

Mit »Ecken« ist das leicht zu konstruieren, z.B.

$$f(x) = \begin{cases} 1 - |x| & \text{für } |x| \leq 1 \\ 0 & \text{für } |x| > 1 \end{cases}$$

(Zeltdach). Wie geht das C^∞, d.h. wie kriegt man die Ecken weg?

Das Hauptproblem ist der glatte Übergang von $\neq 0$ zu 0. Eine solche Funktion kennen wir aus Analysis I. Erinnerung:

$$g(x) = \begin{cases} e^{-1/x} & \text{für } x > 0, \\ 0 & \text{für } x \leq 0 \end{cases}$$

ist C^∞. Offenbar ist $\operatorname{supp} g = [0, \infty)$. Für $a < b$ setze $h_{a,b}(x) = g(x - a) \cdot g(b - x)$, dann ist h offenbar C^∞ und es gilt: $\operatorname{supp} h_{a,b} = [a, b]$. Damit erhält man sofort

4.3.2 Lemma

Sei $U \subset \mathbb{R}^n$ offen, $x_0 \in U$. Dann existiert eine C^∞-Funktion ρ mit $\operatorname{supp} \rho \subset U$, $\rho(x_0) > 0$ und $\rho(x) \geq 0 \; \forall x$.

Beweis: Wähle $[a_1, b_1] \times \ldots \times [a_n, b_n] \subset U$ mit $a_i < b_i \; \forall i$ und setze $\rho(x) = h_{a_1,b_1}(x_1) \ldots h_{a_n,b_n}(x_n)$. □

Bemerkung: Da $\operatorname{supp} \rho$ abgeschlossen ist, bedeutet $\operatorname{supp} \rho \subset U$, dass $\rho = 0$ in einer Umgebung von ∂U gilt.

(Beweis: Man nehme als Umgebung das Komplement von $\operatorname{supp} \rho$. Dieses ist offen, da $\operatorname{supp} \rho$ abgeschlossen ist, und enthält ∂U, da ∂U wegen der Offenheit von U disjunkt zu U ist.)

4.3.3 Satz (Partition der Eins)

Sei $K \subset \mathbb{R}^N$ kompakt und $(U_i)_{i\in I}$ eine offene Überdeckung von K. Dann gibt es eine Umgebung U von K und C^∞-Funktionen $\rho_1, \dots, \rho_r : U \to \mathbb{R}$ mit

(1) $0 \le \rho_j(x) \le 1 \ \forall j \ \forall x$

(2) $\sum_{j=1}^r \rho_j(x) = 1 \ \forall x \in K$

(3) Für jedes j liegt $\operatorname{supp}\rho_j$ in einem der U_i.

Die Funktionen $\{\rho_1, \dots, \rho_r\}$ nennt man eine **Partition der Eins für K, die der Überdeckung U untergeordnet ist**.

Vorstellung: Zerlege K in Teile $K_1, K_2, \cdots, K_r$, so dass jedes Teil »klein« (d.h. in einem U_i enthalten) ist. Die charakteristischen Funktionen $\rho_i = \chi_{K_i}$ erfüllen dann (1)-(3), sind aber unstetig. Der Satz sagt, dass man ihre Sprünge so »glätten« kann, dass (1)-(3) erhalten bleiben.

(Sozusagen eine Zerlegung von K mit fließenden Übergängen.)

Beachte: Endlich viele ρ_j reichen, auch wenn unendlich viele U_i vorkommen (vgl. Satz von Heine-Borel).

Beweis: Zu jedem $x \in K$ wähle :

- i_x mit $x \in U_{i_x}$
- ρ_x nach Lemma 4.3.2 mit $\operatorname{supp}\rho_x \subset U_{i_x}$, $\rho_x \ge 0$, $\rho_x(x) > 0$

Dann bilden die $V_x = \{p : \rho_x(p) > 0\}$ mit $x \in K$ eine offene Überdeckung von K.
$\overset{Heine-Borel}{\Rightarrow}$ $\exists$ endliche Teilüberdeckung $V_{x_1}, \dots, V_{x_r}$. Setze $\tilde\rho_j = \rho_{x_j}$, $j = 1, \dots, r$.
Nach Konstruktion ist $s(x) := \sum_{j=1}^r \tilde\rho_j(x) > 0 \ \forall x \in K$.
$\Rightarrow s > 0$ auf einer Umgebung U von K (setze $U = \{x : s(x) > 0\}$, dies ist als Urbild der offenen Menge $(0,\infty) \subset \mathbb{R}$ unter der stetigen Funktion s offen).
Setze $\rho_j(x) = \frac{\tilde\rho_j(x)}{s(x)}$. (1) und (3) sind per Definition erfüllt, und auf U und damit auf K gilt $\sum \rho_j = \sum \frac{\tilde\rho_j}{s} = \frac{1}{s}\sum \tilde\rho_j = \frac{1}{s}\cdot s = 1$. □

Der mehrdimensionale Hauptsatz und der Satz von Gauß

4.3.4 Satz (Mehrdimensionaler Haupsatz der Differential- und Integralrechnung)

Sei $M \subset \mathbb{R}^n$ ein kompaktes Gebiet mit glattem Rand, $\nu : \partial M \to \mathbb{R}$ das äußere Einheitsnormalenfeld. Sei $f : M \to \mathbb{R}$ eine C^1 -Funktion. Dann gilt für $i = 1, \dots, n$:

$$\int_M \frac{\partial f}{\partial x_i}\, dx = \int_{\partial M} f \cdot \nu_i \, dS$$

Hierbei ist $\nu = (\nu_1, \dots, \nu_n)$.

Bemerkung:

- "f ist C^1-Funktion auf M" ist analog zum Fall offener Teilmengen $U \subset \mathbb{R}^n_+$ definiert: Für Differenzierbarkeit am Rand nimmt man nur die nach innen zeigenden Richtungen für h.

  Man kann zeigen, dass dies äquivalent ist zu: Für jede Karte $\varphi : U \to V$ ist $f \circ \varphi : U \to \mathbb{R}$ eine C^1-Funktion.

Ebenfalls äquivalent dazu ist: Es gibt eine offene Umgebung W von M und eine C^1 - Fortsetzung von f auf W, d.h. eine C^1-Funktion $\tilde{f} : W \to \mathbb{R}$ mit $\tilde{f}_{|M} = f$.

▷ Der Satz gilt auch unter schwächeren Annahmen.

- f ist C^1 in $\overset{\circ}{M}$ und stetig auf M
- M muss nicht glatten Rand haben. Genauer: M heißt C^1-Polyeder (nach Königsberger, Analysis II, Kapitel 12), falls der topologische Rand $\partial M = \{$ reguläre Randpunkte $\} \cup \{$ singuläre Randpunkte $\}$ und die Menge der singulären Randpunkte eine $(n-1)$-Nullmenge ist. Hierbei heißt p regulärer Randpunkt $:\Leftrightarrow$ es gibt eine lokale Karte $\varphi : U \to V, p \in V, U \subset \mathbb{R}^n_+$ offen und singulär := nicht regulär.
  Eine Untermgfk mit Rand hat also nur reguläre Randpunkte.

Der Satz gilt auch für C^1 -Polyeder, wobei rechts nur über den regulären Rand integriert wird.

Zum Beispiel $n = 2$: Ist ∂M eine stückweise C^1-Kurve, so ist M ein C^1-Polyeder, daher ist der Satz anwendbar.

Bemerkung: Wenn Sie Königsberger ganz genau lesen, stellen Sie fest, dass er einen allgemeineren Begriff von $(n-1)$-Nullmenge hat als wir; d. h., unsere $(n-1)$-Nullmengen sind auch solche in seinem Sinne, aber nicht umgekehrt. Damit gilt der Satz also noch etwas allgemeiner als oben formuliert. (Wir haben den Begriff k-Nullmenge nur für Teilmengen von k-dimensionalen Untermannigfaltigkeiten definiert. Den allgemeineren Begriff erhält man durch eine Modifizierung der in Satz 1.4.3 gegebenen äquivalenten Charakterisierung von Nullmengen im $\mathbb{R}^n$.)

Beispiel: $n = 1, \quad M = [a,b]$. Da ν nach außen zeigt, ist $\nu(a) = -1, \nu(b) = 1$. Also ist in Satz 4.3.4:
linke Seite $= \int_a^b f'(x)dx$,
rechte Seite $= f(b) - f(a)$,
und man erhält genau den (zweiten) Hauptsatz der Differential- und Integralrechnung.

Beweis (von Satz 4.3.4): Strategie: Wir beweisen den Satz zunächst für zwei Spezialfälle: Erstens für den Fall, wo der Träger von f im Innern $\overset{\circ}{M}$ enthalten ist, d.h. f in einer Umgebung des Randes verschwindet. Zweitens für den Fall, wo der Träger von f in einer kleinen Umgebung eines Randpunktes enthalten ist. Mittels einer Partition der Eins folgt aus diesen beiden Fällen der Satz für beliebige f.

1. Fall: $\operatorname{supp} f \subset \overset{\circ}{M}$. Da $f_{|\partial M} = 0$, ist die rechte Seite im Satz gleich null, also ist zu zeigen:

Behauptung:

$$\int_M \frac{\partial f}{\partial x_i} dx = 0$$

Beweis (der Behauptung): Sei $\tilde{f} = \begin{cases} f(x) & , x \in M \\ 0 & , x \in \mathbb{R}^n \setminus M \end{cases}$ Dann ist $\tilde{f}$ C^1-Funktion auf ganz $\mathbb{R}^n$, denn $\mathbb{R}^n$ ist Vereinigung der offenen Mengen $\overset{\circ}{M}$ und $\mathbb{R}^n \setminus \operatorname{supp} f$, und auf der ersten ist $\tilde{f} = f$, also C^1, und auf der zweiten ist $\tilde{f} = 0$, also auch C^1.
Wähle nun $R > 0$ so groß, dass M in dem Würfel $[-R,R]^n$ enthalten ist. Es gilt $\int_M \frac{\partial f}{\partial x_i} dx = \int_{[-R,R]^n} \frac{\partial \tilde{f}}{\partial x_i} dx$.
Um zu zeigen, dass dies gleich null ist, wenden wir den ein-dimensionalen Hauptsatz und Fubini an: Der Einfachheit halber sei $i = 1$. Dann gilt für beliebige $(x_2, \ldots, x_n)$, dass $\int_{-R}^{R} \frac{\partial \tilde{f}}{\partial x_1}(x_1, \ldots, x_n)\, dx_1 =$

$\tilde{f}(R, x_2, \ldots, x_n) - \tilde{f}(-R, x_2, \ldots, x_n) = 0$; integriert man dies über $(x_2, \ldots, x_n)$, folgt $\int\limits_{[-R,R]^n} \frac{\partial \tilde{f}}{\partial x_i}\, dx_1\, dx_2 \cdots dx_n =$ 0, was zu zeigen war. □

Vorbemerkung zum 2. Fall: Falls $p \in \partial M$, so gibt es eine Umgebung V_p von p in M der Form $V_p = \{(x', x_n) \in U \times (a, b) : a < x_n < g(x')\}$ für eine C^1 -Funktion $g : U \to (a, b)$. Hierbei schreiben wir $x' = (x_1, \ldots, x_{n-1})$, und U ist eine offene Menge im $\mathbb{R}^{n-1}$. Eventuell muss man die Variablen erst permutieren oder die Spiegelung $x_n \mapsto -x_n$ vornehmen, um diese Form zu erhalten.
(Im Wesentlichen ist das die Tatsache, dass sich die Mannigfaltigkeit ∂M in einer Umgebung von p als Graph einer Funktion g (wie oben angegeben) schreiben lässt.)
Wir nehmen an, dass für jedes $p \in \partial M$ eine solche Umgebung V_p fest gewählt sei.

2.Fall: $\operatorname{supp} f \subset V_p$ für ein $p \in \partial M$.
Erinnerung: Der Normalenvektor in $p = (x', g(x'))$ ist

$$\nu(p) = \frac{(-g_{x_1}(x'), \ldots, -g_{x_n}(x'), +1)}{\sqrt{1 + \|\nabla g(x')\|^2}}$$

(Warum? Entweder mittels der aus Analysis II bekannten Tatsache, dass der Gradient einer Funktion h von n Variablen auf ihren Niveaumengen senkrecht steht, falls er ungleich Null ist: Verwende dies mit $h(x', x_n) = x_n - g(x')$, dies hat $\nabla h = (-\nabla g, 1)$ und Nullniveaumenge $h^{-1}(0) = \partial M \cap V_p$. Normalisierung liefert den Ausdruck für ν.

Oder, komplizierter, mittels des Kreuzproduktes (Übung): $\partial M \cap V_p$ ist nahe p durch $\varphi(x') = (x', g(x'))$ parametrisiert. Dann ist ein Normalenvektor durch $\varphi_{x_1} \times \ldots \times \varphi_{x_{n-1}} = \pm(-\nabla g, 1)$ gegeben.)

Wegen $\operatorname{supp} f \subset V_p$ und der Form von V_p ist $\int\limits_M \frac{\partial f}{\partial x_i} dx = \int\limits_{V_p} \frac{\partial f}{\partial x_i} dx = \int\limits_U \int\limits_a^{g(x')} \frac{\partial f}{\partial x_i}(x', x_n)\, dx_n\, dx'$, und da $\partial M \cap V_p$ Graph der Funktion g mit Definitionsbereich U ist, gilt

$$\int\limits_{\partial M} f \cdot \nu_i\, dS = \int\limits_{\partial M \cap V_p} f \cdot \nu_i\, dS = \int\limits_U f(x', g(x')) \cdot \nu_i(x', g(x')) \underbrace{\sqrt{1 + \|\nabla g(x')\|^2}}_{=\sqrt{\text{Gram-Determinante}}}\, dx',$$

also ist zu zeigen:

Behauptung:

$$\int\limits_U \int\limits_a^{g(x')} \frac{\partial f}{\partial x_i}(x', x_n)\, dx_n\, dx' = \int\limits_U f(x', g(x')) \cdot \nu_i(x', g(x')) \sqrt{1 + \|\nabla g(x')\|^2}\, dx'$$

Beweis (der Behauptung): Der Beweis verläuft für $i = n$ und $i < n$ unterschiedlich.

Fall 2a): $i = n$. Anwendung des ein-dimensionalen Hauptsatzes auf das innere (dx_n) Integral links ergibt $\int\limits_a^{g(x')} \frac{\partial f}{\partial x_n}(x', x_n) dx_n = f(x', g(x')) - \underbrace{f(x', a)}_{=0}$, daher ist das linke Integral in der Behauptung gleich $\int\limits_U f(x', g(x'))dx'$, und wegen $\nu_n = 1/\sqrt{1 + \|\nabla g\|^2}$ ist das gerade die rechte Seite.

Fall 2b): $i < n$.
Zunächst eine Nebenrechnung:

$$\frac{\partial}{\partial x_i} \int_a^{g(x')} f(x', x_n) dx_n = \frac{\partial g}{\partial x_i}(x') f(x', g(x')) + \int_a^{g(x')} \frac{\partial f}{\partial x_i}(x', x_n) dx_n$$

Siehe Lemma 4.3.5 am Ende des Beweises. Beachte, dass x_i unter den x'-Variablen vorkommt. Setzt man nun $H(x') = \int_a^{g(x')} f(x', x_n) dx_n$, so folgt durch Umstellen

$$\int_a^{g(x')} \frac{\partial}{\partial x_i} f(x', x_n) dx_n = \frac{\partial}{\partial x_i} H(x') - \frac{\partial g}{\partial x_i}(x') \cdot f(x', g(x')),$$

und Integration über U liefert schließlich

$$\int_U \int_a^{g(x')} \frac{\partial f}{\partial x_i}(x', x_n)\, dx_n\, dx' = \underbrace{\int_U \frac{\partial H}{\partial x_i}(x')\, dx'}_{=0, \text{ da } \operatorname{supp} H \subset U \text{ (wie im 1.Fall)}} - \int_U \frac{\partial g}{\partial x_i}(x') \cdot f(x', g(x'))\, dx'$$

$$= \int_U f(x', g(x')) \cdot \underbrace{\frac{-\frac{\partial g}{\partial x_i}}{\sqrt{1 + \|\nabla g\|^2}}}_{=\nu_i} \cdot \sqrt{1 + \|\nabla g\|^2}\, dx',$$

was zu zeigen war. □

Abschluss des Beweises des Hauptsatzes: Sei $f : M \to \mathbb{R}$ eine C^1 - Funktion. Die offenen Mengen $\overset{\circ}{M}$, V_p $(p \in \partial M)$ bilden eine offene Überdeckung von M.

M kompakt $\Rightarrow \exists$ Partition der Eins, die dieser Überdeckung untergeordnet ist, d.h. C^∞ - Funktionen $\rho_1, \ldots, \rho_r$ auf einer Umgebung von M mit

- ▷ $\sum \rho_i = 1$ auf M
- ▷ $\forall_i$ entweder $\operatorname{supp} \rho_i \subset \overset{\circ}{M}$ oder $\exists_p \operatorname{supp} \rho_i \subset V_p$

Setze $f_i = \rho_i f$, dann ist $f = \sum_{i=1}^{r} f_i$. Da $f \in C^1$ und $\rho_i \in C^\infty$, ist $f_i \in C^1$, für jedes i. Für jedes f_i trifft Fall 1 oder Fall 2 zu, daher ist der Satz für jedes f_i bereits bewiesen. Da beide Seiten der Behauptung $\int_M \frac{\partial f}{\partial x_i} = \int f \nu\, dS$ linear bezüglich f sind, folgt der Satz für f. □

4.3.5 Lemma

Für C^1-Funktionen g, h gilt

$$\frac{d}{dz} \int_a^{g(z)} h(x, z)\, dx = g'(z) h(g(z), z) + \int_a^{g(z)} \frac{\partial h}{\partial z}(x, z)\, dx$$

Beweis: Der Trick besteht darin, die beiden Vorkommnisse der Variablen z zu entkoppeln, indem man zunächst eine neue Variable für die obere Integrationsgrenze einführt. Die Formel folgt dann aus der Kettenregel: Setze

$$F(y, z) = \int_a^y h(x, z) dx, \text{ dann ist } \frac{\partial F}{\partial y}(y, z) = h(y, z) \text{ (Hauptsatz!)}, \frac{\partial F}{\partial z}(y, z) = \int_a^y \frac{\partial h}{\partial z}(x, z)\, dx$$

(Vertauschen von Ableiten und Integral), also folgt

$$\frac{d}{dz}\int_a^{g(z)} h(x,z)dx = \frac{d}{dz}F(g(z),z) = g'(z)\frac{\partial F}{\partial y}(g(z),z) + \underbrace{\frac{\partial z}{\partial z}}_{=1}\frac{\partial F}{\partial z}(g(z),z)$$

und damit die Behauptung. □

Bemerkung: Vergleiche die Berechnung von $\frac{d}{dx}x^x$ auf ähnliche Weise!

4.3.6 Satz (Integralsatz von Gauß)

$M \subset \mathbb{R}^n$ sei ein kompaktes Gebiet mit glattem Rand und $V : M \to \mathbb{R}^n$ ein C^1 - Vektorfeld. Dann gilt:

$$\int_M \text{div } V\, dx = \int_{\partial M} < V, \nu > dS$$

wobei für $V = (V_1, \ldots, V_n)$ die **Divergenz** von V definiert ist durch

$$\text{div } V = \sum_{i=1}^n \frac{\partial V_i}{\partial x_i} = \frac{\partial V_1}{\partial x_1} + \frac{\partial V_2}{\partial x_2} + \ldots + \frac{\partial V_n}{\partial x_n}$$

Beweis:

$$\int_M \text{div } V\, dx = \sum_{i=1}^n \int_M \frac{\partial V_i}{\partial x_i}dx \overset{Satz}{=} \sum_{i=1}^n \int_{\partial M} V_i\, \nu_i\, dS = \int_{\partial M} < V, \nu > dS$$

□

Bemerkung: Umgekehrt ist Satz 4.3.4 ein Spezialfall des Satzes von Gauß: Nimm für V das Vektorfeld mit f als i-ter Komponente und allen anderen Komponenten gleich null.

Natürlich gilt auch der Satz von Gauß unter den nach Satz 4.3.4 genannten schwächeren Bedingungen an M und V (analog zu f).

Dieser Satz hat viele Anwendungen. Auf einige werden wir weiter unten zu sprechen kommen. Zunächst wollen wir aber erklären, was die Divergenz eines Vektorfeldes und das Integral auf der rechten Seite zu bedeuten haben. Versteht man dies erst, wird der Satz wie eine Selbstverständlichkeit erscheinen.

Interpretation des Satzes von Gauß

Ein Vektorfeld $V : U \to \mathbb{R}^n$ ($U \subset \mathbb{R}^n$ offen) kann man als *Geschwindigkeitsfeld einer stationären Strömung* in U interpretieren. D.h., $V(x) =$ die Geschwindigkeit eines Teilchens am Ort x. »Stationär« bedeutet, dass auch das Teilchen, das sich zu einem anderen Zeitpunkt am Ort x befindet, dann dieselbe Geschwindigkeit $V(x)$ hat. (Mit anderen Worten: Die Geschwindigkeit eines Teilchens hängt nur von dessen momentanem Ort, nicht vom Zeitpunkt ab. Oder: Für einen äußeren Betrachter ist das Strömungsbild konstant.)

Bedeutung des Integrals eines Vektorfelds über eine orientierte Hyperfläche

Sei (H, ν) eine orientierte Hyperfläche, d.h., ν ist ein Einheitsnormalenfeld auf H. Sei V ein stetiges Vektorfeld, definiert in einer offenen Umgebung von H. Dann ist

$$\int_H \langle V, \nu \rangle\, dS = \text{momentane Durchflussrate von } V \text{ durch } H \text{ in Richtung } \nu$$

Das heißt, man interpretiert V als Geschwindigkeitsfeld einer stationären Strömung und misst, wieviel Flüssigkeit F_τ in einem Zeitintervall der Länge $\tau > 0$ durch die Fläche tritt. Dabei wird positiv gezählt, was die Fläche in Richtung ν durchtritt, der Rest negativ. Dann ist $\int_H \langle V, \nu\rangle \, dS = \lim_{\tau\to 0} \frac{F_\tau}{\tau}$. (Durchfluss*rate* bedeutet Durchfluss *pro Zeit*, und *momentan* bedeutet *kurze* Zeitintervalle, also $\tau \to 0$.)

Heuristische Begründung, der Einfachheit halber für $n = 2$: Sei zunächst H eine Strecke und V konstant. Im Zeitintervall $[0, \tau]$ legt jedes Teilchen den Weg τV zurück. Daher ist die Menge der Teilchen, die in diesem Zeitintervall durch H fließen, ein Parallelogramm mit den Seiten H und τV. Es gilt

$$\begin{aligned}\text{Fläche des Parallelogramms } &= \pm(\text{Länge von } H) \cdot \langle \tau V, \nu\rangle \\ &= \pm\tau \cdot \int_H \langle V, \nu\rangle \, dS\end{aligned}$$

wobei das Vorzeichen dasselbe ist wie das von $\langle V, \nu\rangle$; die Behauptung gilt also in diesem Fall. Approximimiert man nun eine beliebige Kurve (=Hyperfläche in $\mathbb{R}^2$) durch viele kurze Geradenstücke, so ist ein beliebiges V nahe jedem davon nahezu konstant, also folgt die Behauptung im Grenzwert $\tau \to 0$.

Diese Heuristik kann man mit Hilfe des Begriffs des Flusses eines Vektorfelds präzisieren, der weiter unten eingeführt wird.

Interpretation der Divergenz eines Vektorfeldes

Wir brauchen zunächst ein algebraisches Lemma. Erinnerung: Zu jeder $n \times n$-Matrix A kann man ihre Determinante $\det A$ und ihre Spur $\operatorname{tr} A$ (die Summe der Diagonalelemente) betrachten. Dies definiert Abbildungen

$$\det : \mathbb{R}^{n\times n} \to \mathbb{R}, \quad \operatorname{tr} : \mathbb{R}^{n\times n} \to \mathbb{R}.$$

tr ist linear, det aber nicht. Immerhin ist det polynomial (nach Leibnizformel ist $\det A$ eine Summe von Produkten von Einträgen von A).

Die Spur ist hier deshalb von Interesse, weil für ein Vektorfeld V gilt:

$$\operatorname{div} V = \operatorname{tr} DV, \quad \text{denn } DV = \begin{pmatrix} \frac{\partial V_1}{\partial x_1} & \cdots & \frac{\partial V_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial V_n}{\partial x_1} & \cdots & \frac{\partial V_n}{\partial x_n} \end{pmatrix}$$

Im Folgenden ist I die Einheitsmatrix.

4.3.7 Lemma

Determinante und Spur sind durch folgende äquivalente Aussagen verbunden:

(a) Das Differential der Determinante bei der Einheitsmatrix ist die Spur:

$$D \det_{|I} = \operatorname{tr}$$

(b) Sei $A : (-\varepsilon, \varepsilon) \to \mathbb{R}^{n\times n}$ ein differenzierbarer Weg von Matrizen mit $A(0) = I$. Dann gilt

$$\frac{d}{dt}_{|t=0} \det A(t) = \operatorname{tr} A'(0)$$

Beweis: Zunächst sind (a) und (b) äquivalent, denn nach der Kettenregel ist

$$\frac{d}{dt}_{|t=0} \det A(t) = D \det_{|A(0)}(A'(0)) = D \det_I(A'(0)).$$

Um (a) zu beweisen, genügt es also, zu vorgebenem $B \in \mathbb{R}^{n\times n}$ (b) für *einen* Weg mit $A'(0) = B$ zu zeigen. Wir nehmen $A(t) = I + tB$ (vergleiche die Beziehung von Richtungsableitung und Differential einer Funktion). Mit $B = (b_{ij})$ gilt

$$\begin{aligned}\det(I+tB) &= \det\begin{pmatrix} 1+tb_{11} & tb_{12} & \dots & tb_{1n} \\ tb_{21} & 1+tb_{22} & \dots & tb_{2n} \\ \vdots & & \ddots & \vdots \\ tb_{n1} & & \dots & 1+tb_{nn}\end{pmatrix} \\ &= (1+tb_{11})\cdot\ldots\cdot(1+tb_{nn}) + O(t^2) \\ &= 1 + t(b_{11}+\ldots+b_{nn}) + O(t^2)\end{aligned}$$

(wobei $O(t^2)$ bedeutet, dass Terme mit mindestens zweiten Potenzen von t weggelassen wurden) und daher $\frac{d}{dt}_{|t=0}\det(I+tB) = b_{11}+\ldots+b_{nn} = \operatorname{tr} B$, was zu zeigen war. □

Erinnerung: Fluss eines Vektorfeldes
Sei $V : U \to \mathbb{R}^n$ ein C^1 Vektorfeld, dann gibt es zu jedem $x \in U$ eine eindeutige Kurve γ_x mit

$$\gamma_x(0) = x, \quad \gamma_x'(t) = V(\gamma_x(t)) \ \forall t,$$

d.h., γ_x ist die Integralkurve von V, die in x startet. Der **Fluss** von V ist dann definiert durch

$$\Phi(t,x) = \gamma_x(t).$$

Man schreibt auch $\Phi_t(x) = \Phi(t,x)$. Dies enthält natürlich dieselbe Information wie alle γ_x zusammen, aber unter einem anderen Blickwinkel: Jedes Φ_t ist eine Abbildung von U nach U; sie gibt an, wie sich die Teilchen in der Zeit t vorwärtsbewegt haben. Per Definition ist

$$\Phi_0 = \mathrm{id}, \quad \frac{d}{dt}\Phi_t(x) = V(\Phi_t(x)).$$

Bemerkung: Genau genommen ist Φ_t nur auf einer Teilmenge von U definiert, da Integralkurven ja auch U »verlassen« können. Der letzte Satz im folgenden Lemma ist dann so zu verstehen: Sind $U' \subset U$ und $t \in \mathbb{R}$ derart, dass Φ_t auf ganz U' definiert ist, dann ist $\Phi_t : U' \to U'' := \Phi_t(U')$ ein Diffeomorphismus mit Inversem $\Phi_{-t} : U'' \to U'$.
Ist aber z.B. $U = \mathbb{R}^n$ und V beschränkt, so ist Φ_t für alle t auf ganz $\mathbb{R}^n$ definiert, und solche Probleme tauchen nicht auf.

Bemerkung (zu Bezeichnungen): In der Physik wird der Begriff ‚Fluss' anders verwendet, nämlich für das Integral $\int_H \langle V, \nu\rangle\, dS$. Im Englischen gibt es diese Begriffsverwirrung nicht: Das Integral heißt **flux**, der Fluss im mathematischen Sinne heißt **flow**.

4.3.8 Lemma

Sei $V : U \to \mathbb{R}^n$ ein C^1-Vektorfeld und Φ sein Fluss. Dann gilt

$$\Phi_{t+s} = \Phi_t \circ \Phi_s$$

für alle $s, t \in \mathbb{R}$, wo beide Seiten definiert sind. Insbesondere ist Φ_t ein Diffeomorphismus mit Inversem Φ_{-t}.

Beweis: Dies sollte anschaulich klar sein: Sei $x \in U$. Fließt ein Teilchen von x aus erst für die Zeit s und dann weiter für die Zeit t, so kommt es an derselben Stelle an, wie wenn es gleich für die Zeit $s+t$ geflossen wäre.
Dies lässt sich leicht mathematisch fassen: Sei $y = \Phi_s(x) = \gamma_x(s)$. Betrachte die Kurven a, b definiert durch

$$a(\tau) = \gamma_y(\tau), \quad b(\tau) = \gamma_x(s+\tau).$$

Dann sind a und b Integralkurven von V mit demselben Anfangspunkt, denn $a(0) = y = b(0)$, und

$$a'(\tau) = \gamma_y'(\tau) = V(\gamma_y(\tau)) = V(a(\tau)) \quad \text{und}$$
$$b'(\tau) = \gamma_x'(s+\tau) = V(\gamma_x(s+\tau)) = V(b(\tau))$$

für alle τ, d.h. a und b sind Lösungen desselben Anfangswertproblems. Nach dem Eindeutigkeitssatz für Lösungen von Differentialgleichungssystemen muss $a = b$ auf dem Existenzintervall dieser Lösungen sein, insbesondere $a(t) = b(t)$, also $\gamma_y(t) = \gamma_x(s+t)$. Dies ist wegen $y = \gamma_s(x)$ gerade die Behauptung. □

4.3.9 Satz

Sei $U \subset \mathbb{R}^n$ offen und $V : U \to \mathbb{R}^n$ ein C^1 Vektorfeld. Sei Φ der Fluss von V. Dann gilt für alle kompakten Teilmengen $K \subset U$:

$$\frac{d}{dt}_{|t=0} \operatorname{vol} \Phi_t(K) = \int_K \operatorname{div} V$$

In Worten: $\int_K \operatorname{div} V =$ momentane Volumenänderungsrate von K unter dem Fluss von V.

Beweis: Vorbereitende Beobachtung: Sei $K \subset U$ kompakt. Dann gibt es ein $\varepsilon > 0$, so dass Φ_t für $|t| < \varepsilon$ auf ganz K definiert ist. Denn nach dem Existenzsatz für Lösungen von Differentialgleichungssysteme gibt es für jedes $x \in U$ eine offene Umgebung U_x und ein $\varepsilon_x > 0$, so dass Φ_t auf U_x für $|t| < \varepsilon_x$ existiert. Da K kompakt ist, überdecken endlich viele, sagen wir $U_{x_1}, \ldots, U_{x_N}$, ganz K. Dann kann man $\varepsilon = \min\{\varepsilon_{x_1}, \ldots, \varepsilon_{x_N}\}$ nehmen.
Also ist die linke Seite der Behauptung wohldefiniert!

Für $|t| < \varepsilon$ ist

$$\operatorname{vol}(\Phi_t(K)) = \int_{\Phi_t(K)} 1\, dy = \int_K |\det(D\Phi_t)_{|x}|\, dx$$

nach der Transformationsformel. Beachte: Rechts wird nur nach x differenziert, nicht nach t.
Behauptung: Sei $x \in K$ fest. Dann gilt:

(a) $\det(D\Phi_t)_{|x} > 0$ für $|t| < \varepsilon$,

(b) $\frac{\partial}{\partial t}_{|t=0} \det(D\Phi_t)_{|x} = \operatorname{div} V(x)$.

Beweis der Behauptung: Sei $A(t) = (D\Phi_t)_{|x}$. Dann gilt:

- $A(0) = I$, da $\Phi_0 = \mathrm{id}$, also $(D\Phi_0)_{|x} = I$,
- die Abbildung $t \mapsto A(t)$ ist C^1,
- $A(t)$ ist invertierbar für jedes t, denn nach Lemma 4.3.8 ist Φ_t ein Diffeomorphismus.

Für $f(t) = \det A(t)$ folgt daraus $f(0) = 1$, $f(t) \neq 0\ \forall t$, und aus der Stetigkeit von f folgt Behauptung (a). Weiterhin folgt aus Lemma 4.3.7 $f'(0) = \operatorname{tr} A'(0)$, und da man t- und x-Differentiation vertauschen kann, gilt

$$A'(0) = \frac{\partial}{\partial t}_{|t=0} (D\Phi_t)_{|x} = D\left(\frac{\partial}{\partial t}_{|t=0} \Phi_t\right)_{|x} = DV_{|x}$$

und damit (b).

Schließlich erhält man durch Vertauschen von Integration und Differentiation

$$\frac{d}{dt}_{|t=0} \operatorname{vol} \Phi_t(K) \overset{(a)}{=} \frac{d}{dt}_{|t=0} \int_K \det(D\Phi_t)_{|x}\, dx = \int_K \frac{\partial}{\partial t}_{|t=0} (D\Phi_t)_{|x}\, dx$$
$$\overset{(b)}{=} \int_K \operatorname{div} V(x)\, dx,$$

was zu zeigen war. □

Bedeutung des Satzes von Gauß

Sei $M \subset \mathbb{R}^n$ kompakt mit glattem Rand und V ein auf einer Umgebung von M definiertes Vektorfeld. Wir betrachten V als stationäres Geschwindigkeitsfeld einer Flüssigkeit. Offenbar gilt dann:

> Das Volumen der anfangs in M befindlichen Flüssigkeit ändert sich genau um so viel, wie durch den Rand von M heraustritt.

Genau genommen gilt dies für die momentane Volumenänderungsrate und die momentane Ausflussrate. Die Interpretation von $\int_{\partial M} \langle V, \nu \rangle\, dS$ und der Satz 4.3.9 zeigen, dass dies genau der Satz von Gauß ist! Das ergibt aber keinen neuen Beweis, da die Herleitung der Bedeutung von $\int_H \langle V, \nu \rangle\, dS$ nicht rigoros war. Man kann's auch andersherum sehen: Da wir den Satz von Gauß streng bewiesen haben, haben wir hiermit die Interpretation von $\int_H \langle V, \nu \rangle\, dS$ im Nachhinein gerechtfertigt (zumindest für $H = \partial M$).

Schließlich wollen wir noch $\operatorname{div} V(x)$ selbst interpretieren, nicht nur sein Integral über kompakte Mengen: Hierzu mittelt man einfach über kleine Mengen K.

4.3.10 Definition

Sei $M \subset \mathbb{R}^N$ eine n-dimensionale Untermannigfaltigkeit mit endlichem, positiven n-Volumen. Sei $f : M \to \mathbb{R}$ über M integrierbar. Der **Mittelwert von f über M** ist definiert als

$$\fint_M f\, dS := \frac{1}{\operatorname{vol}_n(M)} \int_M f\, dS$$

Der wichtigste Fall ist der offener Mengen, d.h. $n = N$. Wir verwenden folgende einprägsame Notation: Ist $x \in \mathbb{R}^n$ und für jedes $\varepsilon > 0$ eine Menge $M_\varepsilon \subset \mathbb{R}^n$ gegeben, so schreiben wir

$$M_\varepsilon \xrightarrow{\varepsilon \to 0} \{x\},$$

falls zu jedem $r > 0$ ein $\varepsilon_0 > 0$ existiert, so dass für alle $\varepsilon < \varepsilon_0$ gilt: $M_\varepsilon \subset K_r(x)$.

Folgende Tatsache sollte intuitiv klar sein.

4.3.11 Lemma

Sei $U \subset \mathbb{R}^n$ und $f : U \to \mathbb{R}$ stetig. Sei $x \in U$. Seien $M_\varepsilon \subset U$, $\varepsilon > 0$, Untermannigfaltigkeiten endlichen positiven Volumens mit $M_\varepsilon \xrightarrow{\varepsilon \to 0} \{x\}$. Dann gilt

$$f(x) = \lim_{\varepsilon \to 0} \fint_{M_\varepsilon} f\, dS$$

Beweis: Übung. □

Das wichtigste Beispiel ist $M_\varepsilon = K_\varepsilon(x)$. Aus Satz 4.3.9 und Lemma 4.3.11 erhalten wir schließlich:

$$\boxed{\operatorname{div} V(x) = \lim_{\varepsilon \to 0} \frac{\text{momentane Vol.änderung von } K_\varepsilon \text{ unter dem Fluss von } V}{\operatorname{vol}(K_\varepsilon)}}$$

für beliebige Mengen $K_\varepsilon \subset \mathbb{R}^n$ positiven n-Volumens mit $K_\varepsilon \xrightarrow{\varepsilon \to 0} \{x\}$.

Die Volumenformel

Als erste Anwendung des Satzes von Gauß geben wir eine Formel an, die das Volumen eines Körpers durch ein Integral über dessen Rand ausdrückt (!).

Betrachte das Vektorfeld $V(x) = x$, dann ist $V_i(x) = x_i$, also $\frac{\partial V_i}{\partial x_i} = 1$ und damit div $V = n$. Wendet man den Satz von Gauß auf V an und teilt durch n, erhält man

$$\boxed{\text{vol}(M) = \frac{1}{n} \int\limits_{\partial M} \langle x, \nu(x) \rangle \, dS(x)}$$

In zwei Dimensionen ergibt sich:
Sei $M \subset \mathbb{R}^2$ ein Gebiet mit stückweise C^1 Rand. Angenommen, ∂M besteht aus *einer* Kurve. $\gamma : I \to \mathbb{R}^2$ sei eine Parametrisierung von ∂M, entgegen dem Uhrzeigersinn durchlaufen. Dies bedeutet genauer: Das Gebiet liegt immer links von der Kurve. Ist $\gamma(t) = (x(t), y(t))$, so ist (x', y') ein Tangentialvektor (bei Punkten, die keine »Ecken« von M sind), also ist $(y', -x')$ (also die Tangente um 90 Grad im Uhrzeigersinn gedreht) ein äußerer Normalenvektor. Die äußere Einheitsnormale ist also $\frac{(y', -x')}{\|(y', x')\|} = \frac{(y', -x')}{\|\gamma'\|}$ und da $dS = \|\gamma'\| \, dt$ (in offensichtlicher Kurznotation), folgt

$$\begin{aligned} \text{Fläche von } M &= \frac{1}{2} \int\limits_I \langle (x, y), \frac{(y', -x')}{\|(y', x')\|} \rangle \, \|\gamma'(t)\| \, dt \\ &= \frac{1}{2} \int\limits_I (xy' - x'y) dt \end{aligned}$$

Beispiel: Fläche des Einheitskreises in der Ebene: $x(t) = \cos t$, $y(t) = \sin t$, $I = [0, 2\pi]$. Mit $x(t)y'(t) - x'(t)y(t) = \cos t \sin t - (-\sin t) \sin t = 1$ ergibt sich als Fläche $\frac{1}{2} 2\pi = \pi$.

Übrigens erhält man andere (weniger symmetrische) Flächen- bzw. Volumenformeln, wenn man z.B. $V(x, y) = (x, 0)$ oder $V(x, y) = (0, y)$. Man erhält

$$\text{Fläche von } M = \int\limits_I xy' \, dt = -\int\limits_I x'y \, dt$$

Der Satz von Stokes in $\mathbb{R}^2$

Ist $M \subset \mathbb{R}^2$ ein kompaktes Gebiet mit stückweise glattem Rand, so ist ∂M eine kompakte eindimensionale Untermannigfaltigkeit. Daraus folgt (Übung), dass $\partial M = \bigcup_{i=1}^k \text{Bild}(\gamma_i)$ für ein $k \in \mathbb{N}$, wobei die γ_i paarweise disjunkte einfach geschlossene Kurven sind. Ihre Bilder bilden die Wegzusammenhangskomponenten von ∂M. Wir sagen, γ_i werde bezüglich M **positiv durchlaufen**, wenn M immer links liegt, wenn also der um $\pi/2$ gegen den Uhrzeigersinn gedrehte Vektor $\gamma_i'(t)$ immer in das Innere von M zeigt.

4.3.12 Definition

Für ein C^1-Vektorfeld V auf $U \subset \mathbb{R}^2$ ist die **Rotation** definiert durch

$$\textbf{rot} V := \frac{\partial V_2}{\partial x} - \frac{\partial V_1}{\partial y}$$

rot V ist eine Funktion $U \to \mathbb{R}$.

Bemerkung: Die Rotation misst also, zu welchem Ausmass V nicht die Integrabilitätsbedingung $\frac{\partial V_2}{\partial x} = \frac{\partial V_1}{\partial y}$ erfüllt. In Dimension drei muss man für denselben Zweck drei Funktionen bilden, daher ist dort die Rotation als das Vektorfeld

$$\operatorname{rot} V = \left(\frac{\partial V_3}{\partial x_2} - \frac{\partial V_2}{\partial x_3}, \frac{\partial V_1}{\partial x_3} - \frac{\partial V_3}{\partial x_1}, \frac{\partial V_2}{\partial x_1} - \frac{\partial V_1}{\partial x_2} \right)$$

definiert. In n Dimensionen hat die Rotation $\binom{n}{2} = \frac{n(n-1)}{2}$ Komponenten und bildet eine sogenannte 2-Form (oder Differentialform vom Grad 2).

Warum heißt das ‚Rotation'? Ist V das Geschwindigkeitsfeld einer Strömung, so gibt $\operatorname{rot} V$ die doppelte Winkelgeschwindigkeit der Drehung eines mitschwimmenden Körpers an.

4.3.13 Satz

Sei $M \subset \mathbb{R}^2$ ein kompaktes Gebiet mit stückweise glattem Rand und V ein C^1-Vektorfeld auf M. Es sei $\partial M = \bigcup_{i=1}^{k} \operatorname{Bild}(\gamma_i)$ für paarweise disjunkte, bezüglich M positiv durchlaufene geschlossene Kurven γ_i. Dann gilt

$$\sum_{i=1}^{k} \int_{\gamma_i} V = \int_M \operatorname{rot} V \, dx$$

Beweis: Man wende den Satz von Gauss auf das Vektorfeld $F = (V_2, -V_1)$ an. Es ist $\operatorname{div} F = \operatorname{rot} V$. Details als Übung. □

Beispiel: Sei $V(x,y) = \left(\frac{-y}{x^2+y^2}, \frac{x}{x^2+y^2} \right)$ das Vektorfeld auf $\mathbb{R}^2 \setminus 0$, das wir in Abschnitt 4.1 betrachtet haben. Die Kurve γ_0 durchlaufe den Einheitskreis um 0 und die Kurve γ_1 ein Rechteck, das diesen Kreis im Innern enthält, jeweils gegen den Uhrzeigersinn. Sei M das Gebiet zwischen den beiden Kurven. Bezüglich M wird γ_1 positiv und γ_0 negativ durchlaufen. Der Satz ergibt also

$$\int_{\gamma_1} V - \int_{\gamma_0} V = \int_M \operatorname{rot} V \, dx = 0,$$

also $\int_{\gamma_1} V = \int_{\gamma_0} V$. Rechnet man diese Integrale direkt aus, erscheint dies wie ein Zufall. Übrigens folgt diese Gleichheit auch aus der Homotopieinvarianz, Korollar 4.1.11, und der darauf folgenden Bemerkung.

Bemerkung: Satz 4.3.13 wird manchmal auch Satz von Green genannt. Mit ‚Satz von Stokes' wird meist außerdem bezeichnet

- eine direkte Verallgemeinerung von Satz 4.3.13 auf Flächen im $\mathbb{R}^3$,
- ein Satz, der sowohl den Satz von Gauß als auch diesen speziellen Satz von Stokes als Spezialfälle enthält; seine Aussage hat die Form $\int_M d\omega = \int_{\partial M} \omega$ für kompakte Mannigfaltigkeiten mit Rand M beliebiger Dimension. $\omega, d\omega$ sind hier Differentialformen, das sind Verallgemeinerungen von Funktionen, Vektorfeldern... Differentialformen sind fundamentale Objekte der Mathematik. Bei erster Bekanntschaft wirken sie etwas kompliziert, ihre Verwendung vereinfacht und vereinheitlicht aber schließlich alles ganz wunderbar!

Wie erkennt man am Rand, ob ein Gebiet den Nullpunkt enthält?

4.3.14 Satz

Sei $M \subset \mathbb{R}^n$ ein kompaktes Gebiet mit glattem Rand. Sei $0 \notin \partial M$. Dann gilt

$$\int\limits_{\partial M} \langle \frac{x}{\|x\|^n}, \nu(x) \rangle \, dS(x) = \begin{cases} 0 & \text{falls } 0 \notin M \\ \omega_n & \text{falls } 0 \in M \end{cases}$$

Hierbei ist $\omega_n = \mathrm{vol}_{n-1}(S^{n-1})$.

Erstaunlich an der Formel ist, dass der Wert des Integrals nicht von der Form von M abhängt, sondern nur davon, ob der Nullpunkt in M liegt oder nicht. Man kann also am Rand von M »ablesen«, ob M den Nullpunkt enthält!

Wie der Satz von Gauß gilt der Satz auch für Gebiete, deren Rand nur bis auf eine $(n-1)$-Nullmenge glatt ist.

Natürlich kann man auch ablesen, ob ein beliebiger Punkt $a \in \mathbb{R}^n$ in M liegt oder nicht: Man verwende das Vektorfeld $\dfrac{x-a}{\|x-a\|^n}$ statt $\dfrac{x}{\|x\|^n}$.

Beweis: Sei $V(x) = \frac{x}{\|x\|^n}$. Dieses Vektorfeld ist für $x \neq 0$ definiert und C^1. Es gilt

$$\mathrm{div}\, V(x) = 0 \quad \forall x \quad \text{(Übung)},$$

also folgt im Fall $0 \notin M$ die Behauptung sofort aus dem Satz von Gauß.

Sei nun $0 \in M$. Dann können wir den Satz von Gauß nicht anwenden, da V nicht auf ganz M definiert ist. Wie geht man damit um? Ein wichtiger Gedanke:

Wenn der Nullpunkt ein Problem ist, schneide ihn aus!

Das heißt, statt M betrachten wir $M_r = M \setminus K_r(0)$, wobei $r > 0$ so klein gewählt sei, dass $K_r(0) \subset M$. Auf M_r ist V definiert und hat Divergenz Null. Der Satz von Gauß ergibt $\int\limits_{\partial M_r} \langle V, \nu \rangle \, dS = 0$. Nun gilt $\partial M_r = \partial M \cup \partial K_r(0)$, wobei die äußere Einheitsnormale bei $x \in \partial K_r(0)$ gleich $-\frac{x}{\|x\|}$ ist, da sie aus M_r herauszeigt. Für $x \in \partial K_r(0)$ folgt $\langle V(x), \nu(x) \rangle = -\|x\|^{-(n-1)} = -r^{-(n-1)}$, also

$$\begin{aligned} 0 = \int\limits_{\partial M_r} \langle V, \nu \rangle \, dS &= \int\limits_{\partial M} \langle V, \nu \rangle \, dS + \int\limits_{\partial K_r(0)} \langle V, \nu \rangle \, dS \\ &= \int\limits_{\partial M} \langle V, \nu \rangle \, dS - \frac{1}{r^{n-1}} \mathrm{vol}_{n-1}(\partial K_r(0)) \end{aligned}$$

und daraus wegen $\mathrm{vol}_{n-1}(\partial K_r(0)) = r^{n-1}\omega_n$ die Behauptung. □

Bemerkung: Sei $n = 2$, und ∂M bestehe nur aus einer einfachen geschlossene Kurve $\gamma = (x, y) : I \to \partial M$, positiv durchlaufen. Dann ist das Integral links in Satz 4.3.14 (vgl. die Rechnung bei der Volumenformel) gleich $\int\limits_I \frac{xy' - yx'}{x^2+y^2} \, dt = \int_\gamma V$ für unser altbekanntes Vektorfeld $V(x,y) = \left(\dfrac{-y}{x^2+y^2}, \dfrac{x}{x^2+y^2} \right)$ aus Abschnitt 4.1. Satz 4.3.14 ergibt

$$\frac{1}{2\pi} \int_\gamma V = \begin{cases} 0 & \text{falls } 0 \notin M \\ 1 & \text{falls } 0 \in M \end{cases}$$

was sich auch mit den Methoden von Abschnitt 4.1 (Homotopieinvarianz) leicht zeigen lässt. Die linke Seite ist für beliebige geschlossene Kurven in $\mathbb{R}^2 \setminus \{0\}$ definiert (nicht notwendig einfache), und es ist nicht schwer zu zeigen, dass immer eine ganze Zahl herauskommt, die **Windungszahl** von γ. Sie gibt an, wie oft sich γ um den Nullpunkt herumwindet. Sie wird uns in Analysis IV wiederbegegnen.

Um dies und eine höherdimensionale Verallgemeinerung (den sogenannten **Abbildungsgrad** einer Abbildung $S^{n-1} \to \mathbb{R}^n \setminus 0$) zu untersuchen und zu verstehen, sind wiederum **Differentialformen** nützlich. Damit kann man dann sehr einfach so hübsche Dinge wie den Brouwerschen Fixpunktsatz zeigen, der besagt, dass jede stetige Abbildung einer Kugel in sich einen Fixpunkt hat. Ein höchst nicht-triviales Ergebnis (außer für $n = 1$, wo Sie es schon kennen sollten).

Partielle Integration in mehreren Dimensionen

Wir erinnern uns, dass die Formel für die partielle Integration in $\mathbb{R}$ sich direkt aus der Produktregel

$$(fg)' = f'g + fg'$$

und dem Hauptsatz der Differential- und Integralrechnung ergibt: Integriert man beide Seiten von a bis b und verwendet links den Haupsatz, folgt

$$f(b)g(b) - f(a)g(a) = \int_a^b f'g + \int_a^b fg'$$

und daraus die übliche Formel.

Kann man auch im $\mathbb{R}^n$ partiell integrieren? Satz 4.3.4 ersetzt den Hauptsatz, und die Produktregel gilt auch für partielle Ableitungen, daher folgt mit exakt demselben Argument:

4.3.15 Satz

Seien f, g C^1-Funktionen auf dem kompakten Gebiet $M \subset \mathbb{R}^n$ mit glattem Rand. Dann gilt für $i = 1, \dots, n$

$$\int_M \frac{\partial f}{\partial x_i} g \, dx = -\int_M f \frac{\partial g}{\partial x_i} \, dx + \int_{\partial M} fg \, \nu_i \, dS$$

Bemerkung: Ein wichtiger Spezialfall ist folgender: Falls $f = 0$ auf ∂M oder $g = 0$ auf ∂M, so gilt

$$\int_M \frac{\partial f}{\partial x_i} g \, dx = -\int_M f \frac{\partial g}{\partial x_i} \, dx$$

Also: Man tut die Ableitung auf die andere Funktion und fügt ein Minuszeichen ein. Das ist trotz seiner Einfachheit eines der wichtigsten Hilfsmittel der Analysis!

Eine oft verwendete Variante hiervon betrifft den **Laplace-Operator Δ**. Dieser ist für C^2-Funktionen f definiert durch

$$\Delta f := \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i^2}$$

Man rechnet direkt nach, dass gilt:

$$\Delta f = \operatorname{div} \nabla f$$

Für eine C^1- Funktion auf einem Gebiet M mit glattem Rand führen wir noch eine Schreibweise ein: Für $x \in \partial M$ sei

$$\frac{\partial f}{\partial n}(x) := \langle \nabla f(x), \nu(x) \rangle.$$

Dies ist die Richtungsableitung von f am Punkt x in Richtung $\nu(x)$, gibt also an, wie steil sich f von innen dem Randpunkt x nähert.

Dadurch wird eine Funktion $\partial f / \partial n$ auf dem Rand ∂M definiert.

4.3.16 Satz (Greensche Formeln)

Seien f, g C^2-Funktionen auf dem kompakten Gebiet $M \subset \mathbb{R}^n$ mit glattem Rand. Dann gilt:

(1) $\displaystyle\int_M f \cdot \Delta g \, dx = -\int_M \langle \nabla f, \nabla g \rangle \, dx + \int_{\partial M} f \cdot \frac{\partial g}{\partial n} \, dS$

(2) $\displaystyle\int_M (\Delta f \cdot g - f \cdot \Delta g) \, dx = \int_{\partial M} \left(\frac{\partial f}{\partial n} \cdot g - f \cdot \frac{\partial g}{\partial n} \right) dS$

Beweis: (1) folgt durch Integration der für Funktionen f und Vektorfelder V gültigen Identität (Übung: nachrechnen)

$$\operatorname{div}(f \cdot V) = \langle \nabla f, V \rangle + f \cdot \operatorname{div} V,$$

angewendet mit $V = \nabla g$, wenn man auf $\int_M \operatorname{div}(fV)$ den Satz von Gauß anwendet. (2) folgt durch Anwenden von (1) auf $g\Delta f$ und $f\Delta g$ und Subtraktion. □

Bemerkung: Wie merkt man sich das? Zunächst die im Beweis verwendete Identität: Dies ist einfach die Produktregel, wobei an jeder Stelle der einfachste Ableitungs- bzw. Produktbegriff verwendet wird: Bei Funktionen der Gradient, bei Vektorfeldern die Divergenz, beim Produkt zweier Vektorfelder das Skalarprodukt.

(1) sagt dann, in konsequenter Fortführung dieser Sicht: Von den in Δg vorkommenden zwei Ableitungen kann man eine auf f herüberschaufeln, für den Preis eines Minuszeichens und eines Randterms, der aus dem Randintegral über die nicht abgeleitete Funktion f und die nur einmal abgeleitete Funktion ∇g (verknüpft mit dem immer auftretenden Normalenvektor) besteht.

(2) ist eine hübsche, weil sehr symmetrische, Konsequenz, die häufig angewendet wird.

5 Laplace-Gleichung, Faltung und Fourieranalysis

In diesem Kapitel werden wir einige weiterführende Themen behandeln, die die bisher erarbeitete Theorie anwenden: Zunächst beschäftigen wir uns mit der Laplace-Gleichung, die als Prototyp einer partiellen Differentialgleichungen in vielen Bereichen der Mathematik, der Physik und der Anwendungen auftritt. Dann führen wir die Faltung ein, die unter anderem die kompakte Darstellung einer Lösung der Laplace-Gleichung im $\mathbb{R}^n$ erlaubt. Schließlich wenden wir uns der Fourieranalysis zu: Zunächst den Fourier-Reihen und schließlich der Fouriertransformation, die wiederum einen ganz neuen Zugang zur Laplace-Gleichung eröffnet.

5.1 Die Laplace-Gleichung

Wir untersuchen die Gleichung

$$\Delta u = f.$$

Hierbei ist f eine gegebene Funktion auf $\mathbb{R}^n$ (oder einer Teilmenge des $\mathbb{R}^n$), und die Funktion u ist gesucht. Δ ist der Laplace-Operator. Dies ist eines der einfachsten Beispiele einer *partiellen Differentialgleichung* und zugleich eines der wichtigsten. Wir werden sehen, dass bei der Untersuchung die Greenschen Formeln wesentlich sind.

Physikalische Motivation

Bezeichnet $u(x)$ die Temperatur eines Körpers am Punkt x, so bedeutet $\Delta u = 0$, dass sich die Temperaturverteilung nicht mit der Zeit ändert. Denn Wärme fließt innerhalb des Körpers in der Richtung und Stärke des steilsten Abfalls von u, also wird der Wärmefluss durch $-\nabla u$ beschrieben. Ändert sich die Temperaturverteilung nicht, so muss pro Zeiteinheit in jeden Teilbereich des Körpers genauso viel Wärme hereinwie aus ihm herausfließen. Nach dem Satz von Gauß folgt $\operatorname{div}\operatorname{grad} u = 0$.

Bedeutung von $\Delta u = f$: Beschreibt f eine Verteilung elektrischer Ladungen im Raum (d.h. in jedem Raumbereich K sitzt die Ladung $\int_K f\,dx$), so ist u das Potential des durch diese Ladung erzeugten elektrischen Feldes. Daraus kann man das Feld selbst als $-\nabla u$ berechnen, d.h. auf ein Elektron am Ort x wirkt die Kraft ∇u. Physikalische Konstanten und Einheiten sind hierbei natürlich vernachlässigt!

Harmonische Funktionen

Wir untersuchen zunächst die homogene Form der Laplace-Gleichung.

5.1.1 Definition

Eine C^2-Funktion u auf einer offenen Menge $U \subset \mathbb{R}^n$ heißt **harmonisch** auf U, falls

$$\Delta u = 0.$$

Für $n = 1$ sind harmonische Funktionen genau die Funktionen u mit $u'' = 0$, also die (affin) linearen Funktionen $u(x) = ax + b$. Für $n \geq 2$ gibt es sehr viel mehr harmonische Funktionen, genauer ist der

Vektorraum der harmonischen Funktionen (Übung: das ist ein Vektorraum) unendlich-dimensional. Z.B. sind für $n = 2$ die Funktionen

$$xy, x^2 - y^2, \mathrm{Re}(x+iy)^m \ (m \in \mathbb{N}_0),\ e^x \sin y, e^{x^2-y^2} \cos(2xy)$$

harmonisch (woher diese Beispiele kommen und wie man das fast ohne Rechnen sofort einsieht, werden Sie in Analysis IV verstehen).

Was bedeutet die Gleichung $\Delta u = 0$, wie stellt man sich das vor? Eine ungefähre Idee, z.B. für $n = 2$: Die Gleichung $\partial_x^2 u + \partial_y^2 u = 0$ bedeutet insbesondere, dass $\partial_x^2 u$ und $\partial_y^2 u$ verschiedene Vorzeichen haben (oder beide null sind). D.h., wenn der Graph von u an einem Punkt in der x-Richtung konvex nach oben gebogen ist (d.h. $\partial_x^2 u > 0$), so muss er in der y-Richtung konkav sein. Das heißt, der Graph muss bei jedem Punkt wie ein Sattel aussehen (oder »flach« sein).

Dies wird in dem folgenden Satz sehr genau quantifiziert.

5.1.2 Satz (Mittelwerteigenschaft harmonischer Funktionen)

Sei $U \subset \mathbb{R}^n$ offen und $u : U \to \mathbb{R}$ harmonisch. Dann gilt für jede Kugel $K_r(x) \subset U$

$$\begin{aligned} u(x) &= \fint_{K_r(x)} u(y)\, dy \\ &= \fint_{\partial K_r(x)} u(y)\, dS(y) \end{aligned}$$

Der Funktionswert bei x ist also gleich dem Mittelwert der Funktionswerte auf einer Kugel oder Sphäre mit Mittelpunkt x.

Beweis: Wir beweisen zunächst die zweite Aussage, über Sphären. Schreibe

$$M(r) := \fint_{\partial K_r(x)} u(y)\, dS(y) = \fint_{S^{n-1}} u(x + r\omega)\, dS(\omega)$$

Die Gleichheit rechts beweist man so, dass man mit der Substitution $y = x + z$ zunächst das Integral $\fint_{\partial K_r(0)} u(x+z)\, dS(z)$ erhält und dann $z = r\omega$ substituiert, vgl. die Bemerkung nach Lemma 5.1.3. Das zweite Gleichheitszeichen sieht man z.B. so, dass man zunächst im ersten Integral $y = x + z$ substituiert, dann ergibt sich $\frac{r^{n-1}\omega_n}{\int_{\partial K_r(0)}} u(x+z)\, dS(z)$,... Wir leiten dies nach r ab. Da im letzten Integral das Integrationsgebiet nicht von r abhängt und kompakt ist, und da der Integrand C^1 ist, darf man die Ableitung in das Integral ziehen und erhält mit der Kettenregel

$$M'(r) = \fint_{S^{n-1}} \frac{\partial}{\partial r} u(x + r\omega)\, dS(\omega) = \frac{1}{\omega_n} \int_{S^{n-1}} \langle (\nabla u)(x + r\omega), \omega \rangle\, dS(\omega).$$

Dies ist genau $\frac{1}{\omega_n} \int_{S^{n-1}} \langle V(\omega), \nu(\omega) \rangle\, dS(\omega)$ für das Vektorfeld $y \mapsto V(y) = (\nabla u)(x + ry)$. Wendet man nun den Gaußschen Integralsatz an, so folgt wegen $\operatorname{div} F(y) = r(\operatorname{div} \nabla u)(x + ry) = 0$, dass $M'(r) = 0$. Also ist M eine konstante Funktion von r. Da außerdem $\lim_{r \to 0} M(r) = u(x)$ gilt, siehe Lemma 4.3.11, folgt $M(r) = u(x)$ für alle r. Damit ist die zweite Behauptung des Satzes bewiesen. Die erste folgt aus der zweiten mittels Integration über Kugelschalen, siehe das folgende Lemma. □

Einschub: Integration über Kugelschalen

5.1.3 Lemma (Integration über Kugelschalen)

Sei f über eine Kugel $K_R(0)$ integrierbar. Dann gilt

$$\fint_{K_R(0)} f(x)\, dx = \int_0^R \fint_{S^{n-1}} f(r\omega)\, dS(\omega)\, r^{n-1}\, dr$$

Dies ist im Sinne des Satzes von Fubini zu verstehen: Das innere Integral rechts existiert für fast alle r und das Resultat ist, nach Multiplikation mit r^{n-1}, bzgl. r integrierbar. Das Analogon des Satzes von Tonelli gilt auch: Existiert das iterierte Integral rechts für $|f|$, so ist f über $K_R(0)$ integrierbar.

Beweis: Für eine lokale Karte $\varphi : U \to V \subset S^{n-1}$ mit $U \subset \mathbb{R}^{n-1}$ sei

$$\Phi : U' := U \times (0,R) \to K_R(0), \quad (u,r) \mapsto r\varphi(u).$$

Dies ist eine lokale Karte für die offene Teilmenge $V' := \{r\omega : r \in (0,R),\, \omega \in V\}$ von $K_R(0)$: Bijektivität sieht man leicht, und das Differential ist

$$\left(\frac{\partial\Phi}{\partial u_1}, \dots, \frac{\partial\Phi}{\partial u_{n-1}}, \frac{\partial\Phi}{\partial r}\right) = \left(r\frac{\partial\varphi}{\partial u_1}, \dots, r\frac{\partial\varphi}{\partial u_{n-1}}, \varphi\right)$$

Für jedes $u \in U$ sind die Vektoren $\frac{\partial\varphi}{\partial u_i}(u)$ tangential on die Sphäre im Punkt $\varphi(u)$ und daher orthogonal zu $\varphi(u)$, und dies ist ein Einheitsvektor. Daher ist die Gram-Matrix $G_\Phi = \begin{pmatrix} r^2 G_\varphi & 0 \\ 0 & 1 \end{pmatrix}$ und daher die Gramsche Determinante $g_\Phi = \det G_\Phi = r^{2(n-1)} g_\varphi$. Insbesondere ist $g_\Phi \neq 0$, also $D\Phi$ bijektiv. Für eine integrierbare Funktion f auf V' folgt damit aus der Definition des Integrals und Fubini

$$\int\limits_{V'} f(x)\,dx = \int\limits_{U\times(0,R)} f(\Phi(u,r))\sqrt{g_\Phi(u,r)}\,du dr = \int\limits_0^R \int\limits_U f(r\varphi(u)) r^{n-1}\sqrt{g_\varphi(u)}\,du\,dr = \int\limits_0^R \int\limits_V f(r\omega)\,dS(\omega)\,r^{n-1}\,dr$$

Da eine Zerlegung von S^{n-1} in Kartengebiete (bis auf $(n-1)$-Nullmengen) zu einer Zerlegung von $K_R(0)$ in Kartengebiete (bis auf n-Nullmengen) führt, folgt die Behauptung. □

Bemerkung: Hier ist eine andere Beweisidee, die besser zeigt, woher der Faktor r^{n-1} kommt:
1. Schritt: Zeige, dass gilt

$$\int\limits_{K_R(0)} f(x)\,dx = \int_0^R Z_r\,dr \quad \text{mit } Z_r = \int\limits_{S_r^{n-1}} f(\alpha)\,dS(\alpha)$$

wobei S_r^{n-1} die Sphäre vom Radius r ist. In Worten: Statt f über die Kugel zu integrieren, integriert man f über jede ‚Zwiebelschale' S_r^{n-1} und integriert dann die Ergebnisse über r.
2. Schritt: Zeige, dass gilt

$$\int\limits_{S_r^{n-1}} f(\alpha)\,dS = r^{n-1} \int\limits_{S^{n-1}} f(r\omega)\,dS(\omega)$$

Dies zeigt man ähnlich wie im Beweis oben – es ist eine Verallgemeinerung der Tatsache, dass das Volumen einer $(n-1)$-dimensionalen Untermannigfaltigkeit bei Streckung um den Faktor r mit dem Faktor r^{n-1} multipliziert wird.

Die Behauptung im 1. Schritt sieht zwar recht offensichtlich aus, ist aber mit Vorsicht zu genießen: Sie stimmt nur in dieser Form, da für die Funktion $r(x) = \|x\|$ gilt, dass $\|\nabla r(x)\| = 1$ für alle $x \neq 0$. Denn es gilt die allgemeinere sogenannte Koflächen-Formel (coarea formula): Sei $U \subset \mathbb{R}^n$ offen und $h : U \to \mathbb{R}$ eine C^1-Funktion, deren Gradient fast überall ungleich Null ist. Sei $M_t = h^{-1}(t)$ die Niveaumenge. Dann gilt:

$$\int\limits_U f(x)\,dx = \int\limits_{\mathbb{R}} Z_t\,dt \quad \text{mit } Z_t = \int\limits_{M_t} f(\alpha)\|\nabla h(\alpha)\|^{-1}\,dS(\alpha)$$

Genauer: M_t ist für fast alle t eine Untermannigfaltigkeit der Dimension $n-1$, und falls f über U integrierbar ist, existiert das innere Integral rechts für fast alle t, und die Formel gilt.
Hier tritt also ein zusätzlicher Faktor $\|\nabla h(\alpha)\|^{-1}$ auf, der im Fall $h = r$ gleich Eins ist.

Folgende Folgerung von Lemma 5.1.3 werden wir später benötigen. Sie ist für viele Zwecke sehr grundlegend.

5.1.4 Lemma

Sei $a \in \mathbb{R}$ und $R > 0$. Die Funktion $x \mapsto \|x\|^{-a}$ ist genau dann über eine Kugel $K_R(0)$ integrierbar, wenn $a < n$.

Beweis:

$$\int\limits_{K_R(0)} \|x\|^{-a}\, dx = \int_0^R \int\limits_{S^{n-1}} r^{-a}\, dS(\omega)\, r^{n-1}\, dr = \omega_n \int_0^R r^{n-1-a}\, dr,$$

und wir wissen bereits, dass dies genau dann endlich ist, wenn $n - 1 - a > -1$, also $n > a$ ist. □

Bemerkung: Wie kommt es, dass in höheren Dimensionen höhere negative Potenzen lokal integrierbar sind? Beispiel: Die Fläche unter der Hyperbel $y = \frac{1}{\|x\|}$, $0 < x < 1$, ist unendlich; rotiert man aber diese Fläche um die y-Achse, so ist das Volumen des entstehenden Körpers $\{(x_1, x_2, y) : 0 < y < \frac{1}{\|x\|}, \|x\| < 1\}$ endlich!

Eine einfache Erklärung:
Der Querschnitt auf ‚Höhe' y ist ein Interval der Länge y^{-1} bzw. eine Kreisscheibe vom Radius y^{-1}, hat also $\begin{cases} \text{die Länge } y^{-1} & \text{für die Fläche unter der Hyperbel bzw.} \\ \text{die Fläche } \pi y^{-2} & \text{für den Rotationskörper} \end{cases}$
Um die Fläche unter der Hyperbel zu erhalten, muss man y^{-1} über $y \in (1, \infty)$ integrieren, das ist unendlich. Für das Volumen des Rotationskörpers muss man πy^{-2} über $y \in (1, \infty)$ integrieren, das ist endlich.

Der Unterschied rührt daher, dass für große y die Zahl y^{-1} viel größer ist als ihr Quadrat y^{-2}.

Aus der Mittelwerteigenschaft folgt insbesondere, dass nicht alle Funktionswerte der harmonischen Funktion u auf dem Kreis $K_r(x)$ echt kleiner als $u(x)$ sein können; entweder sie sind alle gleich $u(x)$, oder manche sind größer und manche kleiner. Daraus ergibt sich leicht:

5.1.5 Satz (Maximumsprinzip für harmonische Funktionen)

Sei $U \subset \mathbb{R}^n$ offen und $u : U \to \mathbb{R}$ harmonisch. Angenommen, U ist wegzusammenhängend. Falls u nicht konstant ist, dann kann u auf U weder sein Maximum noch sein Minimum annehmen.

Wegzusammenhang wird vorausgesetzt, da sonst z.B. u auf einer Wegzusammenhangskomponente konstant gleich eins und auf einer anderen konstant gleich null sein könnte.

Beweis: Angenommen, u nimmt in $x_0 \in U$ sein Maximum M an. Für jede Kugel $K_r(x_0) \subset U$ gilt dann

$$M = u(x_0) = \fint_{K_r(x_0)} u(y)\, dy \le \fint_{K_r(x_0)} M\, dy = M.$$

Daher muss die Ungleichung in Wirklichkeit eine Gleichung sein und daher $u(y) = M$ für alle $y \in K_r(x_0)$ gelten.

Dies zeigt, dass die Menge $u^{-1}(M)$ offen ist. Gleichzeitig ist sie abgeschlossen, da u stetig und $\{M\}$ abgeschlossen ist, und nicht leer. Damit folgt die Behauptung aus dem nächsten Lemma, angewendet auf $U = X$ mit der euklidischen Metrik und $A = u^{-1}(M)$. Der Beweis für das Minimum verläuft analog. □

5.1.6 Lemma

Sei (X, d) ein wegzusammenhängender metrischer Raum. Sei $A \subset X$ eine Teilmenge, die gleichzeitig offen und abgeschlossen ist. Dann ist $A = \emptyset$ oder $A = X$.

Bemerkung: Einen metrischen Raum mit der Eigenschaft, dass nur die Mengen $\emptyset$ und X offen und abgeschlossen sind, nennt man **zusammenhängend**. Die Aussage des Lemmas ist also:

X wegzusammenhängend $\Rightarrow$ X zusammenhängend

Die Umkehrung gilt nicht. Für Beispiele siehe beliebige Bücher zur (mengentheoretischen) Topologie.

Beispiel: Die Menge $X = [0,1] \cup [2,3]$ mit der von $\mathbb{R}$ induzierten Metrik ist offenbar nicht wegzusammenhängend. Sie ist sogar nicht zusammenhängend, denn $A = [0,1]$ ist, als Teilmenge von X, sowohl offen als auch abgeschlossen. Dasselbe gilt für $A = [2,3]$.

Beweis: Angenommen, $A \neq \emptyset$. Dann existiert $p \in A$. Wir müssen zeigen, dass dann $A = X$ folgt. Sei also $q \in X$ beliebig, dann müssen wir $q \in A$ zeigen.

Da X wegzusammenhängend ist, gibt es eine stetige Kurve $\gamma : [a,b] \to X$ mit $\gamma(a) = p$, $\gamma(b) = q$. Sei $A' = \gamma^{-1}(A)$, also die Menge der ‚Zeitpunkte', zu denen sich γ in A aufhält. Da γ stetig und A offen und abgeschlossen ist, ist A' ebenfalls offen und abgeschlossen (in $[a,b]$). Es ist $a \in A'$ wegen $\gamma(a) = p \in A$. Wir müssen $b \in A'$ zeigen. Sei dazu $t_0 = \sup A'$. Da A' abgeschlossen ist, ist $t_0 \in A'$. Wäre $t_0 < b$, so würde ein $\varepsilon > 0$ existieren mit $[t_0, t_0 + \varepsilon) \subset A'$, da A' offen ist. Dies wäre ein Widerspruch zu $t_0 = \sup A'$. Also muss $t_0 = b$ sein, was zu zeigen war. □

Das Maximumsprinzip lässt sich auch so formulieren (diese Aussage ist etwas schwächer):

5.1.7 Korollar (Maximumsprinzip, 2. Form)

Ist $U \subset \mathbb{R}^n$ offen und beschränkt und $u \in C^2(U) \cap C^0(\overline{U})$ harmonisch in U, so gilt

$$\max_{\overline{U}} u = \max_{\partial U} u$$

$$\min_{\overline{U}} u = \min_{\partial U} u$$

Kurz: Eine harmonische Funktion nimmt ihr Maximum und Minimum am Rand an.

Beweis: Da $\overline{U}$ abgeschlossen und beschränkt ist, ist es kompakt, also nimmt u darauf sein Maximum an. Falls das in einem Randpunkt geschieht, sind wir fertig. Falls das Maximum in einem inneren Punkt x_0 angenommen wird, so ist u nach dem Maximumsprinzip auf der Wegzusammenhangskomponente von U, die x_0 enthält, konstant, hat also wegen der Stetigkeit von u auch am Rand dieser Komponente denselben maximalen Wert. Analog für's Minimum. □

Das Dirichlet-Problem

Sei $U \subset \mathbb{R}^n$ offen. Das **Dirichlet-Problem** auf U besteht darin, zu einer gegebenen stetigen Funktion f auf dem Rand ∂U eine Funktion u auf $\overline{U}$ zu finden, die auf U harmonisch ist und am Rand die gegebenen Funktionswerte von f annimmt. Genauer:

$$u \in C^2(U) \cap C^0(\overline{U})$$

$$\Delta u = 0 \quad \text{in } U$$

$$u_{|\partial U} = f$$

Physikalische Bedeutung: Setzt man einen Körper der Form U an seinem Rand der Temperaturverteilung f aus, so beschreibt die Lösung u die Temperaturverteilung im gesamten Körper, die sich nach einiger Zeit einstellt.

Bemerkung: Dass wir hier nicht $u \in C^2(\overline{U})$ fordern, sondern nur die Stetigkeit am Rand, ist nicht einem unsinnigen Streben nach den schwächsten möglichen Voraussetzungen entsprungen, sondern hat reale Bedeutung: Hat nämlich U Ecken, so sind Lösungen u des Dirichletproblems selbst für $f = 0$ typischerweise nicht C^2 am Rand. Ein konkretes Beispiel ist folgendes: Ist $0 < \alpha \leq 2\pi$ und $U = \{(x,y) = (r\cos\varphi, r\sin\varphi) :$

$r > 0, 0 < \varphi < \alpha\}$ der Sektor mit Öffnungswinkel α, so ist $u(r,\varphi) = r^{\pi/\alpha} \sin \frac{\pi}{\alpha}\varphi$ eine harmonische Funktion auf U mit Randwerten null, die im Nullpunkt für $\alpha > \pi$ gar nicht und für $\frac{\pi}{2} < \alpha < \pi$ nur einmal differenzierbar ist. Dies hat in der Numerik große Bedeutung: Numerische Verfahren zur Berechnung einer Lösung u sind aus diesem Grund in der Nähe von Ecken mit großem Innenwinkel meist sehr ungenau.

Andererseits kann man zeigen, dass eine Lösung des Dirichlet-Problems auf einem Gebiet mit glattem Rand automatisch auch am Rand C^2 ist, falls f es ist.

5.1.8 Korollar

Ist $U \subset \mathbb{R}^n$ beschränkt, so hat das Dirichlet-Problem auf U für jede Randwertfunktion f höchstens eine Lösung.

Beweis: Sind u_1, u_2 harmonische Funktionen mit denselben Randwerten, so ist $u = u_1 - u_2$ harmonisch und hat Randwerte null, also ist nach Korollar 5.1.7 Maximum und Minimum von u gleich null, also ist u konstant gleich null, also $u_1 = u_2$. □

Hier ist ein anderer Beweis: Nach der Greenschen Formel ist $\int_U \|\nabla u\|^2\, dx = -\int_U u\Delta u\, dx + \int_{\partial U} u \frac{\partial u}{\partial \nu}\, dS$. Ist $\Delta u = 0$ in U und $u = 0$ am Rand, so ist die rechte Seite gleich Null, also folgt $\nabla u \equiv 0$, also ist u konstant, also gleich Null (weil Null am Rand). Dieser Beweis funktioniert allerdings nur, wenn U (stückweise) glatten Rand hat.

Bemerkung: Für unbeschränktes U gilt die Eindeutigkeit im Allgemeinen nicht, z.B. sind auf $U = \{y > 0\}$ die Funktionen $u(x,y) = y$ und $u(x,y) = 0$ beide harmonisch und verschwinden am Rand. Eindeutigkeit auf unbeschränkten Gebieten erhält man, wenn man zusätzlich zu Randwerten Wachstumsbedingungen an $u(x)$ für $\|x\| \to \infty$ stellt. Als Beispiel sei genannt: Falls $\Delta u = 0$ auf $\mathbb{R}^n$ und $u(x) \to 0$ für $\|x\| \to \infty$, dann folgt $u \equiv 0$. Beweis als Übung (Hinweis: Verwende die Mittelwerteigenschaft über Sphären vom Radius R und betrachte $R \to \infty$).

Bemerkung: Man kann auch zeigen, dass (für beschränktes U) für jedes stetige f wirklich eine Lösung des Dirichlet-Problems existiert, falls der Rand von U nicht allzu »wild« ist. Dies ist weit schwieriger. Es ist Gegenstand der sogenannten Potentialtheorie oder auch der allgemeineren Theorie partieller Differentialgleichungen.

Wir werden noch weitere Techniken zur Lösung der Gleichung $\Delta u = f$ im $\mathbb{R}^n$ kennenlernen. Für diese werden die Faltung und die Fouriertransformation benötigt.

5.2 Die Faltung im $\mathbb{R}^n$ und Anwendungen

Wir werden sehen, dass die Gleichung $\Delta u = f$ im $\mathbb{R}^3$ durch die Funktion $u(x) = \frac{1}{4\pi}\int \frac{1}{\|x-y\|} f(y)\, dy$ gelöst wird (unter gewissen Bedingungen an f). Daher untersuchen wir zunächst Integrale dieser Form.

Definition und grundlegende Eigenschaften der Faltung

5.2.1 Definition

Seien $f, g \in \mathcal{L}^1(\mathbb{R}^n)$. Die **Faltung** (englisch: convolution) von f und g ist die Funktion $f * g$ definiert durch

$$(f * g)(x) = \int_{\mathbb{R}^n} f(x-y)g(y)\, dy, \quad x \in \mathbb{R}^n$$

Beispiel: $g = \frac{1}{2\varepsilon}\chi_{[-\varepsilon,\varepsilon]}$ für ein $\varepsilon > 0$. Dann ist $(f * g)(x) = \frac{1}{2\varepsilon}\int_{-\varepsilon}^{\varepsilon} f(x-y)$, das ist der Mittelwert von f über das Intervall $[x-\varepsilon, x+\varepsilon]$.

Das Beispiel legt eine allgemeine Interpretation der Faltung nahe: Für jedes y ist $x \mapsto f(x-y)$ die um y verschobene Funktion f (d.h. der Graph wird um den Vektor y verschoben). $f * g$ ist also eine Art kontinuierlicher Summe über alle möglichen Verschiebungen von f, wobei die Verschiebung um y mit dem Gewicht $g(y)$ gezählt wird.

Bemerkung: Falls $\int g = 1$ und $g \geq 0$ ist, kann man $(f * g)(x)$ als **gewichteten Mittelwert** der Funktionswerte $f(x-y)$ auffassen, wobei $g(y)$ das Gewicht angibt.

Erklärung zum Begriff des gewichteten Mittelwerts: Zunächst ist der Mittelwert zweier Zahlen a, b bekanntlich $(a+b)/2 = \frac{1}{2}a + \frac{1}{2}b$ (auch arithmetisches Mittel genannt). a und b tragen gleich stark (mit gleichem »Gewicht«) zu diesem Wert bei. Will man, dass b doppelt so stark zu dem Resultat beiträgt wie a, muss man $\frac{1}{3}a + \frac{2}{3}b$ betrachten. Allgemeiner hat man für zwei beliebige Zahlen $g_1, g_2 \geq 0$ mit $g_1 + g_2 = 1$ den mit g_1, g_2 gewichteten Mittelwert $g_1 a + g_2 b$.

Warum fordert man $g_1 + g_2 = 1$? Damit für $a = b$ auch wieder $g_1 a + g_2 b = (g_1 + g_2)a = a$ herauskommt, sonst würde man es wohl nicht »Mittelwert« nennen.

Warum fordert man $g_1, g_2 \geq 0$? Damit $g_1 a + g_2 b$ zwischen a und b liegt, wiederum würde man es sonst wohl nicht »Mittelwert« nennen können. (Nachprüfen: Falls $a \leq b$, so $a = g_1 a + g_2 a \leq g_1 a + g_2 b \leq g_1 b + g_2 b = b$.)

Für endlich viele Zahlen $a_1, \ldots, a_n$ hat man analog das (»gleichgewichtete«) arithemetische Mittel $(a_1 + \cdots + a_n)/n$ und für beliebige $g_1, \ldots, g_n$ mit $g_i \geq 0 \; \forall i$, $\sum_{i=1}^{n} g_i = 1$ das gewichtete Mittel

$$g_1 a_1 + \cdots + g_n a_n = \sum_{i=1}^{n} g_i a_i.$$

Dieses liegt immer zwischen der größten und der kleinsten der Zahlen a_i.

Die analoge Begriffsbildung für ein Kontinuum von Zahlen (also die Funktionswerte einer Funktion a) ist dann das mit einer Gewichtsfunktion g gebildete Mittel $\int g(y)a(y)\,dy$. Damit dies den Namen »Mittelwert« verdient, muss g die Bedingungen

- $\int g = 1$ (damit für $a =$ konstant dieselbe Konstante als Mittelwert herauskommt), und
- $g \geq 0$ (damit $\int ga$ zwischen dem Minimum und dem Maximum von a liegt)

erfüllen.

In dieser Interpretation spielen f und g verschiedene Rollen. Der folgende Satz zeigt, dass sie trotzdem vertauscht werden können.

5.2.2 Satz

Die Faltung hat folgende Eigenschaften. Seien $f, g \in \mathcal{L}^1(\mathbb{R}^n)$.

(a) $(f * g)(x)$ ist für fast alle x definiert.

(b) $f * g$ ist integrierbar, und es gilt $\int f * g = (\int f)\,(\int g)$ (alle Integrale über $\mathbb{R}^n$).

(c) $f * g = g * f$

Die Integralformel in (b) ist so einfach wie bemerkenswert: Der Traum jedes Erstsemesters, dass man das Integral eines Produkts zweier Funktionen als Produkt der Integrale schreiben kann – was ja leider nicht stimmt –, wird wahr, wenn man »Produkt zweier Funktionen« durch »Faltung zweier Funktionen« ersetzt. Die Faltung ist eine Art kompliziertes, verschränktes Produkt.

Beweis: Wir rechnen:

$$\begin{aligned}\int (f*g)(x)\,dx &\overset{\text{def}}{=} \int\int f(x-y)g(y)\,dy\,dx \overset{\text{Fubini}}{=} \int \left(\int f(x-y)g(y)\,dx\right) dy \\ &= \int g(y)\left(\int f(x-y)\,dx\right) dy \\ &= \int g(y)\left(\int f(z)\,dz\right) dy = \int f(z)\,dz \int g(y)\,dy\end{aligned}$$

wobei beim vorletzten Gleichheitszeichen im inneren Integral x durch $z = x - y$ substituiert wurde. Dies würde (b) beweisen – außer dass wir noch gar nicht wissen, ob $(f*g)(x)$ überhaupt definiert ist und ob Fubini anwendbar ist. Um dies zu rechtfertigen, setze $u(x,y) = f(x-y)g(y)$ und $v(z,y) = f(z)g(y)$. Nach dem Satz von Tonelli und der letzten Zeile der Rechnung oben ist $v \in \mathcal{L}^1(\mathbb{R}^{2n})$. Es gilt $u = v \circ T$ mit der Variablentransformation $T(x,y) = (x-y,y)$. T ist linear und invertierbar mit $|\det DT| = |\det T| = 1$, also folgt aus der Transformationsformel, dass $u \in \mathcal{L}^1(\mathbb{R}^{2n})$ und $\int_{\mathbb{R}^{2n}} u(x,y)\,dx\,dy = \int_{\mathbb{R}^{2n}} v(z,y)\,dz\,dy$. Schließlich folgt aus Fubini, dass $u(x,\cdot)$ für fast jedes x bzgl. y integrierbar ist, das ist Behauptung (a), und dass die Rechnung oben gerechtfertigt ist, also folgt (b).

Behauptung (c) folgt mit der Variablentransformation $y \mapsto z = x - y$ (für festes x), dies ist affin linear mit Betrag der Determinante gleich eins, also $\int f(x-y)g(y)\,dy = \int f(z)g(x-z)\,dz$. □

Bemerkung: $(f*g)(x)$ braucht nicht für alle x definiert zu sein. Denn obwohl mit f auch $y \mapsto f(x-y)$ für jedes x integrierbar ist, muss die Funktion $y \mapsto f(x-y)g(y)$ nicht integrierbar sein. Zum Beispiel: $f(x) = g(x) = |x|^{-1/2}\chi_{[-1,1]}(x)$, dann ist der Integrand von $(f*g)(0)$ gleich $f(-y)g(y) = |y|^{-1}\chi_{[-1,1]}(y)$, das ist nicht integrierbar (Exponent -1 im $\mathbb{R}^1$, siehe Lemma 5.1.4).

Im Unterschied zum Produkt zweier Funktionen ist die Faltung so regulär wie *mindestens* einer der Faktoren, und es gilt eine sehr einfache Ableitungsregel:

5.2.3 Satz

Sei $f \in \mathcal{L}^1(\mathbb{R}^n)$ und $g \in C^1(\mathbb{R}^n)$ beschränkt mit beschränktem Gradienten. Dann gilt $f*g \in C^1(\mathbb{R}^n)$ und für $i = 1,\dots,n$

$$\frac{\partial}{\partial x_i}(f*g) = f * \frac{\partial}{\partial x_i} g$$

Natürlich gelten analoge Aussagen, wenn die Rollen von f und g vertauscht sind, und auch für höhere Ableitungen.

Beweis: Mit (c) aus dem vorigen Satz schreibe $\frac{\partial}{\partial x_i}(f*g)(x) = \frac{\partial}{\partial x_i}\int f(y)g(x-y)\,dy$. Nach Satz 1.5.5 darf man die Ableitung ins Integral ziehen, falls

- ▷ die Funktion $x \mapsto f(y)g(x-y)$ für jedes y nach x_i differenzierbar ist (das stimmt hier, da $g \in C^1$),
- ▷ die Funktion $y \mapsto f(y)g(x-y)$ für jedes x integrierbar ist (das stimmt hier, da f integrierbar und g beschränkt ist), und
- ▷ es eine Majorante für die Ableitung gibt, also eine integrierbare Funktion M mit $\left|\frac{\partial}{\partial x_i} f(y)g(x-y)\right| \le M(y)$ für alle x,y (dies stimmt hier, da $\frac{\partial}{\partial x_i} g$ beschränkt ist, etwa durch C, dann kann man $M = C|f|$ nehmen).

Damit ist Satz 1.5.5 anwendbar und gibt direkt die Behauptung. □

Bemerkung: Wir hatten die Faltung zunächst nur für $f, g \in \mathcal{L}^1(\mathbb{R}^n)$ definiert. Alles funktioniert (d.h. sie ist definiert und Sätze 5.2.2 und 5.2.3 gelten entsprechend) auch unter gewissen anderen Voraussetzungen, wie man sich leicht überzeugt, z.B.:

- f integrierbar, g messbar und beschränkt (beachte: eine beschränkte Funktion braucht nicht integrierbar zu sein, z.B. $g \equiv 1$), oder
- f lokal integrierbar, , g integrierbar mit kompaktem Träger.
  Hierbei heißt f **lokal integrierbar**, in Zeichen $f \in \mathcal{L}^1_{\mathrm{loc}}(\mathbb{R}^n)$, falls $\int_K |f| < \infty$ für alle kompakten Mengen $K \subset \mathbb{R}^n$. Zum Beispiel ist $f(x) = \|x\|^{-a}$ mit $a < n$ lokal integrierbar, aber nicht integrierbar – Beweis als Übung mittels Integration über Kugelschalen. Das werden wir in Satz 5.2.5 brauchen.

Die Lösung der Laplace-Gleichung

5.2.4 Definition

Das **Newton-Potential** ist die Funktion auf $\mathbb{R}^n$ definiert durch

$$N(x) := \begin{cases} \frac{1}{2\pi} \log \|x\|, & \text{falls } n = 2 \\ \frac{1}{(2-n)\omega_n} \frac{1}{\|x\|^{n-2}}, & \text{falls } n \neq 2 \end{cases}$$

Hierbei ist $\omega_n = \mathrm{vol}_{n-1}(S^{n-1})$.

Zum Beispiel ist $N(x) = -\frac{1}{4\pi} \frac{1}{\|x\|}$ für $n = 3$. Eine einfache Rechnung zeigt, dass

$$\nabla N(x) = \frac{1}{\omega_n} \frac{x}{\|x\|^n}, \quad \text{also} \quad \Delta N(x) = 0, \quad \text{für } x \neq 0$$

für *alle* n gilt. Der Grund für die bei $n = 2$ andere Formel für N liegt darin, dass $\log \|x\|$ die einzige radiale Funktion auf $\mathbb{R}^2 \setminus \{0\}$ mit Gradient $\frac{x}{\|x\|^2}$ ist. (Das läuft darauf hinaus, dass die Stammfunktion von x^a für $a \neq -1$ eine Potenz, für $a = -1$ aber der Logarithmus ist – rechnen Sie es nach!)

N ist für alle n lokal über $\mathbb{R}^n$ integrierbar. Für $n \neq 2$ sahen wir dies bereits, für $n = 2$ ist es ebenfalls einfach nachzuprüfen (Übung).

5.2.5 Satz (Lösung der Laplace-Gleichung)

Sei $f \in C^2_0(\mathbb{R}^n)$ (das heißt, C^2 mit kompaktem Träger).

a) (Lösungsformel) Dann definiert

$$u(x) = (N * f)(x) = \int_{\mathbb{R}^n} N(x-y) f(y)\, dy$$

eine C^2-Funktion, welche die Gleichung $\Delta u = f$ erfüllt.

b) (Verhalten für große x) Es gibt eine Konstante C, so dass für alle x mit $\|x\| \geq 2$ gilt

$$|u(x)| \leq C \|x\|^{2-n}$$

(falls $n \neq 2$, und $|u(x)| \leq C \log \|x\|$ für $n = 2$).

c) (Eindeutigkeit) Falls $n \geq 3$, so ist $u = N * f$ die einzige Lösung von $\Delta u = f$ mit der Eigenschaft b). Für $n = 2$ ist es die einzige Lösung bis auf Addition von Konstanten.

Beachte, dass der Träger von u nicht kompakt zu sein braucht, auch wenn der von f es ist. Die Eindeutigkeit in c) gilt nicht ohne eine Annahme über das Verhalten von $u(x)$ für große x, da man zu einer Lösung eine beliebige harmonische Funktion, z.B. x_1, addieren kann und wieder eine Lösung erhält.

Beweis: Die Faltung $u = N * f$ ist definiert, da N lokal integrierbar ist (wegen $n-2 < n$ für $n \neq 2$ und wegen $\log r \leq Cr^{-1} + r$ für $n = 2$) und f kompakten Träger hat.

Zweimalige Anwendung von Satz 5.2.3 für jedes i und Summation über i zeigt $u \in C^2(\mathbb{R}^n)$ und $\Delta u = N * (\Delta f)$.

Sei nun x fest gewählt. Wir haben also $(\Delta u)(x) = \int N(y)(\Delta f)(x-y)\,dy$. Nun ist $\partial_{y_i}(f(x-y)) = -(\partial_{x_i} f)(x-y)$ nach der Kettenregel, und nochmalige Anwendung liefert $\partial^2_{y_i}(f(x-y)) = (\partial^2_{x_i} f)(x-y)$, also erhält man nach Summation über i: $(\Delta f)(x-y) = \Delta_y(f(x-y))$ (hierbei bedeutet Δ_y, dass man die Ableitungen bezüglich y nimmt). Wähle R so groß, dass $f(x-y) = 0$ für $\|y\| > R-1$. Um partiell integrieren zu können – das heißt, die zweite Greensche Formel anzuwenden –, schneiden wir eine kleine Kugel um 0 aus und erhalten:

$$\begin{aligned}(\Delta u)(x) &= \lim_{\varepsilon\to 0} \int\limits_{\varepsilon<\|y\|<R} N(y)\Delta_y(f(x-y))\,dy \\ &= \lim_{\varepsilon\to 0}\left(\int\limits_{\varepsilon<\|y\|<R} \Delta_y N(y) f(x-y)\,dy + \int\limits_{\|y\|=\varepsilon}\left(N(y)\frac{\partial}{\partial n_y}(f(x-y)) - \left[\frac{\partial}{\partial n_y}N(y)\right] f(x-y)\right) dS(y)\right)\end{aligned}$$

plus ein ähnlicher Randterm für $\|y\| = R$, der aber gleich null ist, da f in einer Umgebung davon verschwindet. Wir untersuchen nun die drei Terme im letzten Ausdruck, im Grenzwert $\varepsilon \to 0$.

- ▷ Der erste Term verschwindet, weil $\Delta N(y) = 0$ für $y \neq 0$.
- ▷ Im zweiten Term ist $\|y\| = \varepsilon$, also $N(y) = C\varepsilon^{2-n}$ (bzw. $C\log\varepsilon$ für $n = 2$), und wegen $\mathrm{vol}_{n-1}(\{\|y\| = \varepsilon\}) = \omega_n \varepsilon^{n-1}$ und der Beschränktheit von $\partial f/\partial n$ gibt es eine Konstante C', so dass
$$\left|\int\limits_{\|y\|=\varepsilon} N(y)\frac{\partial}{\partial n_y}(f(x-y))\,dS(y)\right| \leq C'\varepsilon \quad \text{(bzw. } C'\varepsilon\log\varepsilon \text{ für } n=2),$$
also ist der Grenzwert für $\varepsilon \to 0$ des zweiten Terms gleich null.
- ▷ Im dritten Term ist ebenfalls $\|y\| = \varepsilon$, also $\frac{\partial N}{\partial n}(y) = \langle \nabla N(y), -\frac{y}{\|y\|}\rangle = -\frac{1}{\omega_n\|y\|^{n-1}} = -\frac{1}{\omega_n\varepsilon^{n-1}}$, also ist
$$-\int\limits_{\|y\|=\varepsilon}\left[\frac{\partial}{\partial n_y}N(y)\right] f(x-y)\,dS(y) = \frac{1}{\omega_n\varepsilon^{n-1}}\int\limits_{\|y\|=\varepsilon} f(x-y)\,dS(y) = \fint_{\partial K_\varepsilon(x)} f(z)\,dS(z)$$
(Substitution $z = x-y$), und das strebt für $\varepsilon \to 0$ gegen $f(x)$ (siehe Lemma 4.3.11).

Insgesamt erhält man also $(\Delta u)(x) = 0 + 0 + f(x)$. Damit folgt a). b) und c) als Übung. □

Bemerkung: Warum kann man nicht einfach $(\Delta u)(x) = \Delta_x \int N(x-y)f(y)\,dy = \int \Delta_x N(x-y)f(y)\,dy = \int 0 = 0$ rechnen, da der Integrand außer bei dem Punkt $y = x$, also außerhalb einer Nullmenge, gleich Null ist? Das kann nicht stimmen! Der Beweis zeigt ja, dass $f(x)$ herauskommt, nicht null.

Wo liegt der Fehler? Die Voraussetzungen des Satzes über das Vertauschen von Ableiten und Integrieren sind nicht erfüllt! Denn man findet keine Majorante für die Ableitungen von $N(x-y)f(y)$, und, schlimmer noch, die zweiten partiellen x-Ableitungen sind von der Form $1/\|x-y\|^n$ plus ein ähnlicher Term, und dies ist nicht einmal über kompakte Mengen integrierbar, siehe Lemma 5.1.4.

Dies zeigt in eindrucksvoller Klarheit, was Sie schon immer wussten, aber nie so recht wahrhaben wollten:

> Man muss darauf achten, dass die Voraussetzungen der Sätze, die man verwendet, erfüllt sind. Sonst kommt vielleicht etwas Falsches heraus!

Der Haupttrick des Beweises liegt darin, mit der Variablensubstitution $z = x - y$ die x-Abhängigkeit in dem Integral, das u definiert, auf f zu übertragen (vgl. die Symmetrie der Faltung). Da man trotzdem irgendwie ausnutzen will, dass $\Delta N(y) = 0$ für $y \neq 0$ ist (das ist alles, was wir in der Hand haben), verwenden wir dann die Greensche Formel nach Ausschneiden des Problempunktes $y = 0$, um Δ von f auf N herüberzuschieben, und die auftretenden Randterme liefern genau das Resultat!

Approximationen der Identität

Wir haben gesehen, dass die Faltung eine Art Produktoperation auf Funktionen darstellt: Sie macht aus zwei Funktionen eine neue, ist bilinear (z.B. gilt das Assoziativgesetz $(f_1 + f_2) * g = f_1 * g + f_2 * g$) und kommutativ. Es liegt nahe, zu fragen:

Gibt es ein neutrales Element für das Faltungsprodukt, d.h. eine Funktion $\delta \in \mathcal{L}^1(\mathbb{R}^n)$, die $f * \delta = f$ für alle $f \in \mathcal{L}^1(\mathbb{R}^n)$ erfüllt?

Man sieht leicht, dass die Antwort NEIN ist: Zu gegebenem δ kann man immer ein f finden, so dass diese Gleichung nicht stimmt. Z.B. setze man $f(x) = \delta(-x)$ für $x \neq 0$ und $f(0) = -1$. Dann ist $(f * \delta)(0) = \int f(-y)\delta(y)\, dy = \int (\delta(y))^2\, dy > 0$, aber $f(0) < 0$. Man sieht auch leicht, dass es selbst unter Stetigkeitsvoraussetzungen an f nicht funktionieren kann.

Es gibt aber einen Ersatz, nämlich Familien δ_ε von Funktionen mit $f * \delta_\varepsilon \xrightarrow{\varepsilon \to 0} f$ für alle $f \in \mathcal{L}^1(\mathbb{R}^n)$. Diese werden im Zusammenhang mit Fourierreihen eine große Rolle spielen, daher zeigen wir nun, wie man solche erhält.

5.2.6 Satz (Approximation der Identität)

Seien δ_ε für $\varepsilon > 0$ integrierbare Funktionen auf $\mathbb{R}^n$ mit den folgenden Eigenschaften:

(1) $\int \delta_\varepsilon = 1 \quad \forall \varepsilon$,

(2) $\delta_\varepsilon \geq 0 \quad \forall \varepsilon$,

(3) Für jedes $r > 0$ gilt: $\int\limits_{\|x\| > r} \delta_\varepsilon(x)\, dx \to 0$ für $\varepsilon \to 0$.

Dann gilt $f * \delta_\varepsilon \to f$ für $\varepsilon \to 0$, und zwar

- bzgl. der L^1-Norm, falls f integrierbar ist
- punktweise, falls f stetig und beschränkt ist,
- gleichmäßig, falls f gleichmäßig stetig und beschränkt ist.

Eine Familie $(\delta_\varepsilon)_\varepsilon$ mit diesen Eigenschaften nennt man eine **Approximation der Identität**. (Manchmal hat man auch eine Folge (δ_k) mit analogen Eigenschaften für $k \to \infty$ statt $\varepsilon \to 0$ gegeben, dann spricht man auch von einer **Dirac-Folge**.)

Die ersten beiden Bedingungen an δ_ε sagen, dass $f(x)$ ein mit $\delta_\varepsilon(y)$ gewichteter Mittelwert der Funktionswerte $f(x - y)$ ist (vgl. die Bemerkung nach der Definition der Faltung). Die dritte Bedingung sagt, dass sich δ_ε für $\varepsilon \to 0$ immer näher am Nullpunkt konzentriert. Damit sollte die Behauptung zumindest für stetiges f anschaulich klar sein.

Das ist von der Idee her nichts anderes als Lemma 4.3.11, welches für offene Mengen K_ε der Spezialfall $\delta_\varepsilon = \frac{1}{\mathrm{vol}(K_\varepsilon)} \chi_{K_\varepsilon}$ von Satz 5.2.6 ist.

Beispiel: Ist φ eine integrierbare Funktion auf $\mathbb{R}^n$ mit $\int \varphi = 1$, $\varphi \geq 0$, so bilden die φ_ε definiert durch

$$\varphi_\varepsilon(x) = \frac{1}{\varepsilon^n}\varphi\left(\frac{x}{\varepsilon}\right)$$

eine Approximation der Identität (Übung).

Den Spezialfall mit $\varphi(x) = \frac{1}{2}\chi_{[-1,1]}$ haben wir bereits nach der Definition 5.2.1 der Faltung kennengelernt.

Weitere wichtige Beispiele werden wir im Zusammenhang mit Fourierreihen und der Fouriertransformation kennenlernen.

Beweis (von Satz 5.2.6): Wir behandeln zuerst den Fall, wo f stetig und beschränkt ist, da dabei die Hauptidee am besten sichtbar wird. Hieraus erhalten wir die Aussage für Treppenfunktionen f mittels Approximation durch stetige, beschränkte Funktionen und dann für integrierbare Funktionen mittels Approximation durch Treppenfunktionen.

Fall 1: f ist stetig und beschränkt. Wegen $\int \delta_\varepsilon = 1$ ist $f(x) = \int f(x)\delta_\varepsilon(y)\,dy$ und damit

$$f(x) - (f * \delta_\varepsilon)(x) = \int f(x)\delta_\varepsilon(y)\,dy - \int f(x-y)\delta_\varepsilon(y)\,dy = \int [f(x) - f(x-y)]\,\delta_\varepsilon(y)\,dy,$$

also $|f(x) - (f*\delta_\varepsilon)(x)| \leq \int |[f(x) - f(x-y)]\,\delta_\varepsilon(y)|\,dy = \int |f(x) - f(x-y)|\,\delta_\varepsilon(y)\,dy$ (wegen $\delta_\varepsilon \geq 0$). Warum wird das klein für $\varepsilon \to 0$? Idee: Falls $\|y\|$ groß ist, so ist $\delta_\varepsilon(y)$ klein (im Mittel) nach Bedingung (3), und falls $\|y\|$ klein ist, so ist $f(x) - f(x-y)$ klein wegen der Stetigkeit von f. Wir teilen also für ein $r > 0$ (wird später gewählt) das Integral in die Teile $\|y\| < r$ und $\|y\| > r$ auf und erhalten

$$|f(x) - (f*\delta_\varepsilon)(x)| \leq I_{r,\varepsilon} + J_{r,\varepsilon},$$

$$I_{r,\varepsilon} = \int\limits_{\|y\|<r} |f(x) - f(x-y)|\,\delta_\varepsilon(y)\,dy, \quad J_{r,\varepsilon} = \int\limits_{\|y\|>r} |f(x) - f(x-y)|\,\delta_\varepsilon(y)\,dy.$$

Sei M eine obere Schranke für $|f|$. Dann ist $|f(x) - f(x-y)| \leq 2M$ für alle x, y. Sei $\alpha > 0$ gegeben (das ersetzt das übliche ε, das hier schon anders verwendet wird). Wegen der Stetigkeit von f im Punkt x gibt es ein $r > 0$, so dass $|f(x) - f(x-y)| < \alpha$ für $\|y\| < r$ gilt, also

$$I_{r,\varepsilon} \leq \alpha \int \delta_\varepsilon = \alpha$$

(für beliebiges ε). Wegen Bedingung (3) gibt es zu diesem r ein ε_0, so dass für alle $\varepsilon < \varepsilon_0$ gilt $\int\limits_{\|y\|>r} \delta_\varepsilon(y)\,dy < \alpha$, also

$$J_{r,\varepsilon} \leq 2M \int\limits_{\|y\|>r} \delta_\varepsilon(y)\,dy < 2M\alpha.$$

Insgesamt erhält man $|f(x) - (f*\delta_\varepsilon)(x)| < (2M+1)\alpha$ für $\varepsilon < \varepsilon_0$. Da $\alpha > 0$ beliebig war, folgt $(f*\delta_\varepsilon)(x) \to f(x)$ für $\varepsilon \to 0$, also die punktweise Konvergenz.

Fall 1a: f ist gleichmäßig stetig und beschränkt. Dann kann r und damit ε_0 unabhängig von x gewählt werden, also erhält man die gleichmäßige Konvergenz.

Für das Folgende brauchen wir noch die L^1-Konvergenz im Fall, dass f stetig ist und kompakten Träger hat. Dann ist f gleichmäßig stetig und beschränkt. Sei $\operatorname{supp} f \subset K_{R-1}(0)$ und $K = K_R(0)$. Wir teilen $\int |f - f*\delta_\varepsilon| = \int_K |f - f*\delta_\varepsilon| + \int_{K^c} |f - f*\delta_\varepsilon|$ auf und zeigen, dass beide Integrale für $\varepsilon \to 0$ gegen Null gehen.

Wegen der gezeigten gleichmäßigen Konvergenz ist $\sup_K |f - f*\delta_\varepsilon| \to 0$ für $\varepsilon \to 0$, also folgt $\int_K |f - f*\delta_\varepsilon| \to 0$, da K endliches Maß hat.

Für $x \in K^c$ ist $f(x) = 0$, also $\int_{K^c} |f - f * \delta_\varepsilon| = \int_{K^c} |(f * \delta_\varepsilon)(x)|\, dx \leq \int_{K^c} \int_{K_{R-1}(0)} |f(y)\delta_\varepsilon(x-y)|\, dy dx = \int_{K_{R-1}(0)} |f(y)| \int_{K^c} |\delta_\varepsilon(x-y)|\, dx\, dy \leq \int_{K_{R-1}(0)} |f(y)|\, dy \int\limits_{\|z\| \geq 1} |\delta_\varepsilon(z)|\, dz$, wobei die letzte Ungleichung daraus folgt, dass $\|x - y\| \geq 1$ für $x \in K^c$, $y \in K_{R-1}(0)$ gilt. Wegen Bedingung (3) geht nun das letzte Integral für $\varepsilon \to 0$ gegen Null, und die L^1-Konvergenz ist bewiesen.

Wir wollen nun die L^1-Konvergenz für L^1-Funktionen zeigen. Dies geschieht mittels Approximation.

Erinnerung: Für eine Teilmenge $\mathcal{M} \subset \mathcal{L}^1(\mathbb{R}^n)$ ist der Abschluss $\overline{\mathcal{M}}$ die Menge der Funktionen, die sich bzgl. der L^1-Norm durch Elemente von $\mathcal{M}$ approxmieren lassen:

$$\overline{\mathcal{M}} = \{f \in \mathcal{L}^1(\mathbb{R}^n) : \exists (f_k) \subset \mathcal{M} \text{ mit } f_k \xrightarrow{k\to\infty} f\}$$

wobei $f_k \xrightarrow{k\to\infty} f :\Leftrightarrow \|f - f_k\|_1 \xrightarrow{k\to\infty} 0$.

Zentral ist nun folgendes:

Approximationsprinzip: Angenommen, wir wissen für eine Klasse $\mathcal{M} \subset \mathcal{L}^1(\mathbb{R}^n)$ von Funktionen bereits, dass $\|f - f * \delta_\varepsilon\|_1 \xrightarrow{\varepsilon\to 0} 0$ für alle $f \in \mathcal{M}$ gilt. Dann gilt dies auch für alle Funktionen im Abschluss von $\mathcal{M}$.

Beweis (des Approximationsprinzips): Der zentrale Punkt ist die Stetigkeit der Faltung bzgl. der L^1-Norm:

$$\|f * g\|_1 \leq \|f\|_1 \|g\|_1.$$

Dies folgt aus $|(f * g)(x)| = |\int f(x-y)g(y)\, dy| \leq \int |f(x-y)g(y)|\, dy = (|f| * |g|)(x)$, da nach Satz 5.2.2(b) $\| |f| * |g| \|_1 = \int |f| * |g| = \int |f| \cdot \int |g| = \|f\|_1 \|g\|_1$ gilt.

Ist nun $f \in \overline{\mathcal{M}}$ und (f_k) eine Folge in $\mathcal{M}$ mit $f_k \xrightarrow{k\to\infty} f$, so schreiben wir $f - f * \delta_\varepsilon = (f - f_k) + (f_k - f_k * \delta_\varepsilon) + (f_k - f) * \delta_\varepsilon$, woraus mit der Dreiecksungleichung

$$\|f - f * \delta_\varepsilon\|_1 \leq \|f - f_k\|_1 + \|f_k - f_k * \delta_\varepsilon\|_1 + \|(f_k - f) * \delta_\varepsilon\|_1$$

folgt. Zu gegebenem $\alpha > 0$ wählen wir nun k so groß, dass $\|f - f_k\|_1 < \alpha$ ist, und dann ε_0 so, dass für dieses k $\|f_k - f_k * \delta_\varepsilon\|_1 < \alpha$ für $\varepsilon < \varepsilon_0$ ist. Der dritte Summand erfüllt $\|(f_k - f) * \delta_\varepsilon\|_1 \leq \|f_k - f\|_1 \|\delta_\varepsilon\|_1 = \|f_k - f\|_1 < \alpha$ (denn $\|\delta_\varepsilon\|_1 = 1$ nach Bedingungen (1) und (2) im Satz), also folgt

$$\|f - f * \delta_\varepsilon\|_1 < 3\alpha \quad \text{für } \varepsilon < \varepsilon_0,$$

was zu zeigen war. □

Fall 2: f ist eine Treppenfunktion. Dann gilt $\|f - f * \delta_\varepsilon\|_1 \to 0$. Dies folgt aus dem Approximationsprinzip und Fall 1, da jede Treppenfunktion L^1-Grenzwert einer Folge stetiger Funktionen mit kompaktem Träger ist. (Die stetigen Funktionen erhält man durch »Glättung« der Sprungstellen: Betrachte zunächst die charakteristische Funktion eines Quaders, z.B. in einer Dimension $\chi_{[0,1]}$. Dies ist der L^1-Grenzwert der Funktionen $f_k(x) = \begin{cases} kx, & 0 \leq x \leq \frac{1}{k} \\ 1, & 1/k \leq x \leq 1 - \frac{1}{k} \\ k(1-x), & 1 - \frac{1}{k} \leq x \leq 1 \\ 0 & \text{sonst} \end{cases}$. Durch Bilden von Linearkombinationen erhält man stetige Approximationen beliebiger Treppenfunktionen. Details als Übung.)

Fall 3: $f \in \mathcal{L}^1(\mathbb{R}^n)$: Da nach der Definition der Integrierbarkeit $\mathcal{L}^1(\mathbb{R}^n)$ der Abschluss der Menge der Treppenfunktionen ist, folgt die Behauptung aus dem Approximationsprinzip. □

Wir notieren eine Folgerung, die in vielen Überlegungen die Argumentation stark erleichtert.

5.2.7 Korollar

$C_0^\infty(\mathbb{R}^n)$, die Menge der **glatten Funktionen mit kompaktem Träger**, ist dicht in $\mathcal{L}^1(\mathbb{R}^n)$.

»Dicht« bedeutet, dass sich jede $\mathcal{L}^1$-Funktion durch C_0^∞-Funktionen bzgl. der L^1-Norm beliebig gut approximieren lässt, also $\overline{C_0^\infty} = \mathcal{L}^1$.

Beweis: Alle Funktionen im Folgenden sind auf $\mathbb{R}^n$ definiert, daher lassen wir $\mathbb{R}^n$ in den Bezeichnungen weg.

Wähle ein $\varphi \in C_0^\infty$ mit $\varphi \geq 0$ und $\int \varphi = 1$. Hierfür wähle ρ wie in Lemma 4.3.2 mit $U = \mathbb{R}^n$ und x_0 beliebig und setze $\varphi = \rho / \int \rho$.

Dann bildet die Familie $(\varphi_\varepsilon)_{\varepsilon>0}$ mit $\varphi_\varepsilon = \varepsilon^{-n}\varphi(\frac{x}{\varepsilon})$ eine Approximation der Identität.

Wir zeigen zunächst, dass C_0^∞ in $\mathcal{L}_0^1 := \{f \in \mathcal{L}^1 : \operatorname{supp} f \text{ kompakt}\}$ dicht liegt. $f \in \mathcal{L}^1$ habe also kompakten Träger. Dann ist $f * \varphi_\varepsilon$ glatt nach Satz 5.2.3 und hat kompakten Träger (Übung), für jedes $\varepsilon > 0$. Da $f * \varphi_\varepsilon \xrightarrow{\varepsilon\to 0} f$ bzgl. der L^1-Norm gilt, folgt die Behauptung.

Das Prinzip der Integration durch Ausschöpfung zeigt, dass $\mathcal{L}_0^1$ dicht in $\mathcal{L}^1$ liegt. Also folgt die Behauptung aus dem folgenden Lemma. □

5.2.8 Lemma

Seien (X, d) ein metrischer Raum und $A \subset B \subset X$ Teilmengen. Ist A dicht in B und B dicht in X, so ist A dicht in X.

Beweis: A dicht in B bedeutet $\overline{A} \supset B$, und daraus folgt $\overline{\overline{A}} \supset \overline{B} = X$. Da $\overline{A}$ bereits abgeschlossen ist, ist $\overline{\overline{A}} = \overline{A}$, also folgt $\overline{A} = X$, d.h. A ist dicht in X. □

Bemerkung: Ganz ähnlich zeigt man, dass für beliebiges offenes $U \subset \mathbb{R}^n$ die Menge $C_0^\infty(U)$ von glatten Funktionen auf U mit kompaktem Träger dicht in $\mathcal{L}^1(U)$ liegt.

5.3 Grundbegriffe über Fourierreihen

Vorbemerkung: Die Kapitel über Fourierreihen sind noch in der 'Rohfassung' und müssen noch überarbeitet werden!!

Im Folgenden betrachten wir Funktionen der Form $f : S^1 \to \mathbb{C}$, wobei S^1 der Einheitskreis in der Ebene bzw. den komplexen Zahlen ist:

$$S^1 = \{z \in \mathbb{C} : |z| = 1\} = \{e^{i\varphi} : \varphi \in \mathbb{R}\}$$

Es gibt aber Sachverhalte, die sich mit Funktionen auf anderen Definitionsbereichen besser interpretieren lassen. Daher werden hier folgende Funktionen als äquivalent zu den obigen auf S^1 definierten angesehen:

- 2π-periodische Funktionen $g : \mathbb{R} \to \mathbb{C}$
- Auf einem halboffenen Intervall I der Länge 2π definierte Funktionen $h : I \to \mathbb{C}$, z.B. mit $I = [-\pi, \pi)$ oder $I = [0, 2\pi)$

Wir können diese Funktionsklassen als äquivalent betrachten, weil wir sie z.B. wie folgt ineinander überführen können: Mit $f, g, h, \varphi \in \mathbb{R}$ wie oben und $I = [-\pi, \pi)$ ist

$$g(\varphi) = f(e^{i\varphi}) \quad \text{und} \quad g(\varphi + 2\pi) = f(e^{i\varphi + 2\pi i}) = f(e^{i\varphi} \cdot e^{2\pi i}) = f(e^{i\varphi}) = g(\varphi)$$

und

$$g(\varphi) = h\left(\varphi - 2\pi - \left\lfloor \frac{\varphi - \pi}{2\pi} \right\rfloor \cdot 2\pi\right).$$

Um all diese Schreibweisen zusammenzufassen schreiben wir einfach

$$f \in \mathcal{L}^1(S^1) \quad \text{oder} \quad f \in \mathcal{L}^1([-\pi, \pi)).$$

Bemerkung: Ist $I = [a, a+2\pi)$ (mit $a \in \mathbb{R}$), so gilt:

$$f : S^1 \to \mathbb{C} \text{ ist stetig} \quad \Leftrightarrow \quad h : I \to \mathbb{C} \text{ ist stetig und hat eine stetige Fortsetzung auf } [a, a+2\pi] \text{ mit } h(a) = h(a+2\pi)$$

Es ergeben sich dadurch kleine Ungenauigkeiten beim Schreiben, die wir aber hinnehmen, da aus dem Kontext direkt ersichtlich ist was gemeint ist:

Beispiel: f mit $f(x) = e^{ikx}$ ist eine Funktion auf (und nach) S^1. Sie ist aber nur 2π-periodisch und stetig falls $k \in \mathbb{Z}$:

$$e^{ik(x+2\pi)} = e^{ikx} \cdot e^{2\pi ki} = e^{ikx} \quad \Leftrightarrow \quad e^{2\pi ki} = 1 \quad \Leftrightarrow \quad k \in \mathbb{Z}$$

Dies gilt auch für $f(x) = \sin(kx)$ oder $f(x) = \cos(kx)$.

Die zentrale Frage um die es bzw. deren Beantwortung in diesem Kapitel geht ist die folgende:

Frage: Können beliebige Funktionen auf S^1 als (unendliche) Linearkombinationen verschiedener e^{ikx} ($k \in \mathbb{Z}$) geschrieben werden?

5.3.1 Definition

Sei $f \in \mathcal{L}^1(S^1)$. Der ***k*-te Fourierkoeffizient** ($k \in \mathbb{Z}$) von f ist

$$\hat{f}(k) := \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-ikx}\, dx$$

5.3.2 Lemma

Für $k, l \in \mathbb{Z}$ ist

$$\int_{-\pi}^{\pi} e^{ikx} \cdot e^{-ilx}\, dx = \begin{cases} 0 & , k \neq l \\ 1 & , k = 2\pi \end{cases}$$

Beweis:

▷ Ist $k \neq l$ so gilt:

$$\int_{-\pi}^{\pi} e^{ikx} \cdot e^{-ilx}\, dx = \int_{-\pi}^{\pi} e^{ix(k-l)}\, dx = \frac{1}{k-l} e^{ix(k-l)} \Big|_{-\pi}^{\pi} = 0$$

▷ Ist $k = l$ so gilt:

$$\int_{-\pi}^{\pi} e^{ikx} \cdot e^{-ilx}\, dx = \int_{-\pi}^{\pi} e^{ix \cdot 0}\, dx = \int_{-\pi}^{\pi} 1\, dx = 2\pi$$

□

Für die sogenannten **trigonometrischen Polynome** $f \in \mathcal{T}$, das sind die mit $f(x) = \sum_{k=-N}^{N} c_k e^{ikx}$ für ein $N \in \mathbb{N}, c_k \in \mathbb{C}$ lässt sich mit den gerade gewonnen Einsichten die Frage schon beantworten:

5.3.3 Lemma

Ist $f \in \mathcal{T}$, so gilt:

$$\sum_{k=-N}^{N} \hat{f}(k) e^{ikx}$$

Das heißt: Der k-te Koeffizient eines trigonometrischen Polynoms ist dessen k-ter Fourierkoeffizient.

Beweis: Sei $k_0 \in \{-N, \dots, N\}$. Es gilt dann:

$$\int_{-\pi}^{\pi} f(x)e^{-ik_0x}\,dx = \int_{-\pi}^{\pi} \sum_{k=-N}^{N} c_k e^{ikx}e^{-ik_0x}\,dx = \sum_{k=-N}^{N} c_k \cdot \underbrace{\int_{-\pi}^{\pi} e^{ikx}e^{-ik_0x}\,dx}_{\begin{array}{ll} = 0 & \forall k \neq k_0 \\ = 2\pi & k = k_0 \end{array}} = c_{k_0} \cdot 2\pi$$

Damit folgt also:

$$c_{k_0} = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-ik_0x}\,dx = \hat{f}(k_0)$$

□

Schreibweise: Für $f \in \mathcal{L}(S^1)$ schreibe

$$f \sim \sum_{k=-\infty}^{\infty} \hat{f}(k)e^{ikx}$$

Die rechte Seite nennt man die **Fourierreihe** von f.

Es gibt noch eine andere Darstellungsform für die Fourierreihe, zu der man gelangt, wenn man die bekannte Eulersche Identität $e^{ikx} = \cos(kx) + i \cdot \sin(kx)$ verwendet:

$$f \sim \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cos(kx) \; + \; \sum_{k=1}^{\infty} b_k \sin(kx)$$

wobei die a_k, b_k wie folgt definiert sind:

$$a_k := \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos(kx)\,dx \quad \text{und} \quad b_k := \frac{1}{\pi}\int_{-\pi} \pi f(x)\sin(kx)\,dx$$

Frage: Ist eine Funktion *immer* gleich ihrer Fourierreihe?

Nein, aber oft. In den folgenden Kapiteln werden wir dies näher untersuchen.

5.4 Konvergenz von Fourierreihen; der Satz von Fejér

Frage: Wann gilt $f(x) = \sum_{k=-\infty}^{\infty} \hat{f}(k)e^{ikx}$? Schreibt man für $f \in \mathcal{L}(S^1)$ und $N \in \mathbb{N}$

$$(S_N f)(x) := \sum_{k=-N}^{N} \hat{f}(k)e^{ikx}$$

so könnte man auch fragen: Wann gilt $S_N f \xrightarrow{N\to\infty} f$? Und bezüglich welcher (Halb-)Norm gilt diese Konvergenz?

Bevor wir diese Fragen beantworten können brauchen wir noch einige Hilfsmittel.

Zunächst stellt man fest, dass man $S_N f$ noch etwas anders darstellen kann:

$$\begin{aligned}(S_N f)(x) &= \sum_{k=-N}^{N} \hat{f}(k)e^{ikx} = \sum_{k=-N}^{N} \frac{1}{2\pi}\int_{-\pi}^{\pi} f(y)e^{-iky}\,dy e^{ikx} \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \underbrace{\left(\sum_{k=-N}^{N} e^{ik(x-y)}\right)}_{=D_N(x-y)} \cdot f(y)\,dy = \frac{1}{2\pi}\int_{-\pi}^{\pi} D_N(x-y)f(y)\,dy\end{aligned}$$

Die letzte Form sieht der schon bekannten Faltung sehr ähnlich, lediglich die Grenzen des Integrals und der Faktor vor diesem sind neu. Man bezeichnet D_N als den **Dirichlet-Kern n-ten Grades**:

$$D_N(x) := \sum_{k=-N}^{N} e^{ikx}$$

Weil es zweckmäßig ist und keiner bisherigen Definition direkt widerspricht führt man speziell für Funktionen aus $\mathcal{L}(S^1)$ einen modifizierten Faltungsbegriff ein:

5.4.1 Definition

Für $f, g \in \mathcal{L}(S^1)$ sei die **Faltung** $f * g$ von f und g gegeben durch

$$(f * g)(x) := \frac{1}{2\pi} f(x-y)g(y)\,dy$$

Bemerkung: Statt $(-\pi, \pi)$ kann auch jedes andere Intervall der Länge 2π für das Integral verwendet werden. Der Faktor $\frac{1}{2\pi}$ wird in einigen Lehrbüchern weggelassen um eine universelle Definition der Faltung über beliebige Mengen geben zu können. Dieser Faktor tritt dann aber im Bereich der Fourierreihen natürlich immer mal wieder zwischendurch auf.

5.4.2 Lemma

Zu $\varepsilon > 0$ sei $g_\varepsilon \in \mathcal{L}^1(S^1)$ mit

(1) $\int g_\varepsilon = 1$

(2) $g_\varepsilon \geq 0$

(3) $\forall r > 0 : \int_{\pi > \|x\| > r} g_\varepsilon \xrightarrow{\varepsilon \to 0} 0$

Es gilt dann: $f * g_\varepsilon \xrightarrow{\varepsilon \to 0} f$ für alle $f \in \mathcal{L}^1(S^1)$,

- bezüglich der $\|\cdot\|_1$-Norm, falls $f \in \mathcal{L}^1$ ist
- bezüglich der $\|\cdot\|_\infty$-Norm ($\hat{=}$gleichmäßig), falls $f \in \mathcal{C}^0$ ist

Beweis: Der Beweis läuft genauso wie der zu 5.2.6. □

5.4.3 Lemma

Es gilt für $f \in \mathcal{L}(S^1)$ und $N \in \mathbb{N}$:

$$S_N f = D_N * f \quad \text{und} \quad D_N(x) = \sum_{k=-N}^{N} e^{ikx} = \begin{cases} \frac{\sin\left((N+\frac{1}{2})x\right)}{\sin\left(\frac{1}{2}x\right)} & , x \neq 0 \\ 2N+1 & , x = 0 \end{cases}$$

Beweis: Dass $S_N f = D_N * f$ ist, wurde schon am Anfang des Kapitels hergeleitet. Es bleibt zu zeigen, dass sich der N-te Dirichletkern wie oben angegeben darstellen lässt. Es gilt

$$D_N(x) = \sum_{k=-N}^{N} e^{ikx} = \sum_{k=-N}^{N} (e^{ix})^k$$

Substituiert man nun $t = e^{ix}$ erkennt man hier eine geometrische Summe:

$$\sum_{k=-N}^{N} (e^{ix})^k = \sum_{k=-N}^{N} t^k = t^{-N} \cdot \sum_{k=0}^{2N} t^k = \begin{cases} t^{-N} \cdot \frac{t^{2N+1}-1}{t-1} & , t \neq 1 \Leftrightarrow x \neq 0 \\ 2N+1 & , t = 1 \Leftrightarrow x = 0 \end{cases}$$

Der im Fall $t \neq 1$ entstehende Bruch wird nun weiter umgeformt:

$$t^{-N} \cdot \frac{t^{2N+1}-1}{t-1} = \frac{t^{N+1}-t^{-N}}{t-1} = \frac{t^{N+\frac{1}{2}}-t^{-(N+\frac{1}{2})}}{t^{\frac{1}{2}}-t^{-\frac{1}{2}}}$$

Mit der Eulerschen Identität zur komplexen Exponentialfunktion folgt außerdem:

$$t^{\frac{1}{2}}-t^{-\frac{1}{2}} = e^{i\frac{1}{2}x}-e^{-i\frac{1}{2}x} = 2i\sin\left(\frac{1}{2}x\right)$$

Per vollständiger Induktion lässt sich so nun weiter zeigen:

$$t^{N+\frac{1}{2}}-t^{-(N+\frac{1}{2})} = 2i\sin\left((N+\frac{1}{2})x\right)$$

Damit folgt also insgesamt im Fall $x \neq 0$:

$$\frac{t^{N+\frac{1}{2}}-t^{-(N+\frac{1}{2})}}{t^{\frac{1}{2}}-t^{-\frac{1}{2}}} = \frac{2i\sin\left((N+\frac{1}{2})x\right)}{2i\sin\left(\frac{1}{2}x\right)} = \frac{\sin\left((N+\frac{1}{2})x\right)}{\sin\left(\frac{1}{2}x\right)}$$

□

Bemerkung: Es gilt für beliebiges $N \in \mathbb{N}$:

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} D_N(x)\,dx = 1 \quad \text{denn} \quad \int_{-\pi}^{\pi} D_N(x)\,dx = \sum_{k=-N}^{-1}\int_{-\pi}^{\pi} e^{ikx}\,dx + \int_{-\pi}^{\pi} 1\,dx + \sum_{k=1}^{N}\int_{-\pi}^{\pi} e^{ikx}\,dx = 2\pi$$

Es gilt zudem

$$|D_N(x)| \leq \frac{1}{\left|\sin\frac{1}{2}x\right|} \leq \frac{4}{x}$$

Mit $g_N := \frac{1}{2\pi}D_N$ ist $\int g_N = 1$, aber nicht $g_N \geq 0$. Das heißt $D_N * f$ konvergiert (im Sinne von 5.4.2) vielleicht gar nicht gegen f!

Anmerkung: Ist f *stetig*, so konvergiert dessen Fourierreihe fast überall gegen f.
Auswege:

- ▷ Stärkere Bedingungen an f
- ▷ Schwächere Konvergenzbegriffe

Aus der Analysis 1 weiß man:

$$(S_n)_n \in \mathbb{R}^{\mathbb{N}} : s_n \xrightarrow{n\to\infty} s \quad \Longrightarrow \quad \sigma_n := \frac{1}{n}\sum_{k=0}^{n-1} s_k \xrightarrow{n\to\infty} s\,,$$

aber nicht umgekehrt (z.B. gilt für $s_n = (-1)^n$: $\sigma_n \xrightarrow{n\to\infty} 0$, aber nicht $s_n \xrightarrow{n\to\infty} 0$).
Hier schreiben wir nun also formal: $s_N = S_N$ und $\sigma_N := \frac{1}{N}\sum_{k=0}^{N-1} S_k$ bzw. $\sigma_N f = \frac{1}{N}\sum_{k=0}^{N-1} S_k f$.

5.4.4 Satz (von Fejér)

(1) $f \in \mathcal{L}^1(S^1) \quad \Longrightarrow \quad \sigma_N f \xrightarrow{N\to\infty} f$ bzgl. $\|\cdot\|_1$

(2) $f \in \mathcal{C}^0(S^1) \quad \Longrightarrow \quad \sigma_N f \xrightarrow{N\to\infty} f$ bzgl. $\|\cdot\|_\infty$

(3) Falls die Grenzwerte $f(x_0^-)$ und $f(x_0^+)$ für ein $x_0 \in S^1$ existieren, so gilt:

$$(\sigma_N f)(x_0) \xrightarrow{N\to\infty} \frac{f(x_0^-)+f(x_0^+)}{2}$$

Beweis: Zum Beweis der ersten beiden Teile wollen wir das Lemma 5.4.2 verwenden. Dazu muss zunächst $\sigma_N f$ als Faltungmit einer Funktion dargestellt werden. Danach zeigen wir, dass diese Funktion eine Approximation der Identität ist.
Durch einfaches ausschreiben und Anwenden des Distributivgesetzes für das Faltungsprodukt gelangt man zu besagter Darstellung:

$$S_N f = D_N * f \Rightarrow \sigma_N f = \frac{1}{N}\sum_{k=0}^{N-1} S_k f = \frac{1}{N}\sum_{k=0}^{N-1} D_k * f = \underbrace{\left(\frac{1}{N}\sum_{k=0}^{N-1} D_k\right)}_{=:F_N} * f = F_N * f$$

Mit Hilfe der Additionstheoreme für den Sinus erhält man eine weitere Form für F_N, die es uns erleichtert zu zeigen, dass $\frac{1}{2\pi}F_N$ (für $n \to \infty$) eine Approximation der Identität ist:

$$F_N = \frac{1}{N}\sum_{k=0}^{N-1} D_k = \ldots = \frac{1}{N\sin\frac{x}{2}}\sum_{k=0}^{N-1}\sin\left(\frac{2k+1}{2}x\right) = \ldots = \frac{1}{N}\left(\frac{\sin\left(\frac{N}{2}x\right)}{\sin\left(\frac{1}{2}x\right)}\right)^2$$

Wie angekündigt zeigen wir nun, dass $g_N := \frac{1}{2\pi}F_N$ für $N \to \infty$ eine Approximation der Identität ist:

(1) Nach der Bemerkung von oben gilt für beliebiges $N \in \mathbb{N}$:

$$\int_{-\pi}^{\pi} g_N = \frac{1}{N}\sum_{k=0}^{N-1}\frac{1}{2\pi}\int_{-\pi}^{\pi} D_k = \frac{1}{N}\sum_{k=0}^{N-1} 1 = \frac{1}{N}N = 1$$

(2) Da $\frac{1}{2\pi}F_N(x) = \frac{1}{2\pi N}\left(\frac{\sin\left(\frac{N}{2}x\right)}{\sin\left(\frac{1}{2}x\right)}\right)^2$ und $x \in \mathbb{R}$ ist $F_N \geq 0$.

(3) Sei $r > 0$. Dann folgt aufgrund der Eigenschaften des Sinus, dass es ein $c > 0$ gibt, sodass für alle x mit $r \leq |x| \leq \pi$ gilt: $\left|\sin\left(\frac{1}{2}x\right)\right| \geq c$. Damit lässt sich F_N abschätzen:

$$F_N(x) = \frac{1}{N}\left(\frac{\sin\left(\frac{N}{2}x\right)}{\sin\left(\frac{1}{2}x\right)}\right)^2 \leq \frac{1}{N}\left(\frac{1}{\sin\left(\frac{1}{2}x\right)}\right)^2 \leq \frac{1}{N}\frac{1}{c^2}$$

Daraus folgt nun schließlich:

$$\int_{r\leq|x|\leq\pi} g_N \leq \frac{1}{2\pi}\int_{r\leq|x|\leq\pi}\frac{1}{c^2} = \frac{c'}{N} \xrightarrow{N\to\infty} 0$$

Mit Lemma 5.4.2 folgen damit die ersten beiden Aussagen des Satzes.

Da $F_N(x) = F_N(-x)$ gilt, folgt insbesondere:

$$\frac{1}{2\pi}\int_0^{\pi} F_N(x)\,dx = \frac{1}{2\pi}\int_{-\pi}^{0} F_N(x)\,dx = \frac{1}{2}$$

Demzufolge und weil $\sigma_N f = F_N * f$ ist, gilt:

$$(\sigma_N f)(x_0) = (F_N * f)(x_0) = \frac{1}{2\pi}\int_{-\pi}^{\pi} F_N(y) f(x_0 - y)\,dy = \underbrace{\frac{1}{2\pi}\int_0^{\pi} F_N(y) f(x_0-y)\,dy}_{\xrightarrow{N\to\infty}\frac{1}{2}f(x_0^-)} + \underbrace{\frac{1}{2\pi}\int_{-\pi}^{0} F_N(y) f(x_0-y)\,dy}_{\xrightarrow{N\to\infty}\frac{1}{2}f(x_0^+)}$$

Die angegebenen Grenzwerte ergeben sich ähnlich wie im Beweis zu Satz 5.2.6. □

Bemerkung: Man bezeichnet F_N als den Fejér-Kern n-ten Grades:

$$\boldsymbol{F_N(x) := \frac{1}{N}\sum_{k=0}^{N-1} D_k(x)}$$

Im Spezialfall, dass f stetig ist ergibt der obige Satz, dass $(\sigma_N f)(x_0) \xrightarrow{N\to\infty} f(x_0)$.

5.4.5 Korollar

Sei $f \in \mathcal{L}^1(S^1)$ und $x_0 \in S^1$. Falls die Fourierreihe von f bei x_0 konvergiert und $f(x_0^-), f(x_0^+)$ existieren, so folgt:

$$\lim_{N\to\infty}(S_N f)(x_0) = \sum_{k=-\infty}^{\infty} \hat{f}(k)e^{ikx} = \frac{f(x_0^-)+f(x_0^+)}{2}$$

Beweis: Per Voraussetzung existiert ein $s \in \mathbb{C}$ so, dass $(S_N f)(x_0) \xrightarrow{N\to\infty} s$ gilt. Dann gilt aber auch $(\sigma_N f)(x_0) \xrightarrow{N\to\infty} s$. Nach Satz 5.4.4 gilt aber gerade $s = fracf(x_0^-) + f(x_0^+)2$. □

Bemerkung: Analoges gilt natürlich für die Sinus/Cosinus-Darstellung, da es sich nur um eine andere Darstellungsform handelt.

Beispiel: Sei $f(x) = |x|$ auf $(-\pi, \pi)$. Da f gerade ist, ist fallen in der Sinus/Cosinus-Darstellung von $S_N f$ die Sinus-Terme weg. Es ist

$$a_k = \frac{1}{\pi}\int_{-\pi}^{\pi} |x|\cos(kx)\,dx = \frac{2}{\pi}\int_0^{\pi} x\cos(kx)\,dx = \underbrace{\frac{2}{\pi}\frac{x}{k}\sin(kx)\Big\|_0^{\pi}}_{=0} - \frac{2}{\pi}\int_0^{\pi}\frac{1}{k}\sin(kx)\,dx = -\frac{2}{\pi}\frac{1}{k^2}\left[1-(-1)^k\right]$$

Für $k = 0$ ergibt sich $a_0 = \pi$. Man erkennt, dass für gerades $k \neq 0$ der Wert von $a_k = 0$ ist. und für alle ungeraden k gilt: $a_k = -\frac{4}{\pi}\frac{1}{k^2}$. Damit ergibt sich:

$$f \sim \frac{\pi}{2} - \frac{4}{\pi}\left(\cos x + \frac{\cos(3x)}{3^2} + \frac{\cos(5x)}{5^2} + \cdots\right) = \frac{\pi}{2} - \frac{4}{\pi}\sum_{k=0}^{\infty}\frac{\cos((2k+1)x)}{(2k+1)^2}$$

Da diese Reihe für alle x konvergiert und f stetig ist gilt hier sogar die Gleichheit, d.h.

$$\forall x \in [-\pi,\pi) : f(x) = |x| = \frac{\pi}{2} - \frac{4}{\pi}\sum_{k=0}^{\infty}\frac{\cos((2k+1)x)}{(2k+1)^2}$$

Als interessantes ›Nebenprodukt ‹ ergibt sich eine Formel, die π beschreibt: Setzt man $x = 0$ und löst nach $\frac{\pi}{2}$ auf so ergibt sich:

$$\frac{\pi}{2} = \frac{4}{\pi}\sum_{k=0}^{\infty}\frac{1}{(2k+1)^2} \quad \text{bzw.} \quad \frac{\pi^2}{8} = \sum_{k=0}^{\infty}\frac{1}{(2k+1)^2}$$

(Vergleiche: $\sum_{k=1}^{\infty}\frac{1}{k^2} = \frac{\pi^2}{6}$)

5.5 Fourierreihen und Orthogonalität

Eín sehr gutes Verständnis von Fourierreihen erhält man, wenn man den Blickwinkel ändert: Man fragt nicht nach der punktweisen oder gleichmäßigen Konvergenz der Fourierreihe, sondern nach der *Konvergenz im quadratischen Mittel*.

Dies versteht man am besten, wenn man zunächst einen abstrakten Standpunkt einnimmt: Wir betrachten Funktionen als Vektoren in einem riesigen normierten Vektorraum.

5.5.1 Definition

Sei $A \subset \mathbb{R}^n$ messbar. Der Raum der über A **quadratintegrierbaren Funktionen** ist

$$\mathcal{L}^2(A) := \{f : A \to \mathbb{C} : f \text{ messbar}, \quad \int_A |f|^2 < \infty\}$$

Das **(L^2)-Skalarprodukt** zweier Funktionen $f, g \in \mathcal{L}^2(A)$ ist definiert als

$$\langle f, g\rangle := \int_A f(x)\overline{g(x)}\,dx$$

und die **L^2-Norm** als

$$\|f\|_2 := \sqrt{\int_A |f(x)|^2\,dx} = \sqrt{\langle f, f\rangle}$$

Hierbei ist $\overline{g(x)}$ das komplex Konjugierte von $g(x)$.

Wir werden gleich nachprüfen, dass das Integral in der Definition von $\langle f, g\rangle$ überhaupt existiert.

Ähnlich wie die L^1-Halbnorm ist auch $\|\cdot\|_2$ eine Halbnorm (wie wir gleich sehen werden). Es ist meist bequemer, mit Normen statt mit Halbnormen zu hantieren, daher definieren wir ähnlich wie früher (Definition 1.4.7)

5.5.2 Definition

Zwei Funktionen $f, g \in \mathcal{L}^2(A)$ heißen äquivalent, wenn sie fast überall gleich sind. Die Menge der Äquivalenzklassen wird mit $L^2(A)$ bezeichnet.

Da es in der Praxis umständlich wäre, immer von Äquivalenzklassen zu reden, denkt man sich Elemente von $L^2(A)$ (wie schon bei $L^1(A)$) als »Funktionen, die nur fast überall definiert sind« und schreibt auch etwas ungenau $f \in L^2(A)$ für Funktionen f.

Bemerkung: Manche Operationen und Begriffe muss man allerdings mit Vorsicht behandeln. Z.B. ist der Wert einer L^1- oder L^2-Funktion an einem Punkt nicht definiert (denn man könnte sie ja an diesem Punkt ändern, ohne ihre Äquivalenzklasse zu ändern). Wohl aber ist $\int_B f$ für messbare Teilmengen $B \subset A$ definiert. $f \cdot g$ und $f + g$ sind fast überall definiert, wenn f und g es sind.

Man sagt, ein Element $f \in L^1(A)$ ist stetig, wenn es einen stetigen Repräsentanten hat (d.h. auf einer Nullmenge so abgeändert werden kann, dass es stetig wird). Ein anderer undefinierter Begriff ist das Supremum $\sup f$. Ein sinnvoller Ersatz ist

$$\operatorname{ess\,sup} f := \inf\{\sup g : g = f \text{ fast überall}\},$$

das **wesentliche Supremum** essential supremum von f. Für stetige Funktionen auf offenen Mengen stimmt dies mit dem üblichen Supremum überein.

5.5.3 Satz

Sei $A \subset \mathbb{R}^n$ messbar.

(1) $L^2(A)$ ist ein (komplexer) Vektorraum.

(2) $\langle \cdot, \cdot \rangle$ definiert ein hermitesches Skalarprodukt auf $L^2(A)$.

(3) Es gilt die **Cauchy-Schwarz Ungleichung**: Für $f, g \in L^2(A)$ ist
$$\langle f, g \rangle \leq \|f\|_2 \cdot \|g\|_2$$

(4) $\| \cdot \|_2$ ist eine Norm auf $L^2(A)$, insbesondere gilt für $f, g \in L^2(A)$:
$$\|f + g\|_2 \leq \|f\|_2 + \|g\|_2$$

(5) Der normierte Raum $(L^2(A), \| \cdot \|_2)$ ist vollständig, das heißt jede Cauchy-Folge konvergiert.

Beweis:

(1) Wir zeigen zunächst, dass $\mathcal{L}^2(A)$ ein Vektorraum ist. Da $\mathcal{L}^2(A)$ eine Teilmenge des Vektorraums aller Funktionen $A \to \mathbb{C}$ ist, brauchen wir nur zu zeigen, dass mit $f, g \in \mathcal{L}^2(A)$ und $c \in \mathbb{C}$ auch cf und $f + g$ in $\mathcal{L}^2(A)$ liegen. Mit f, g sind auch cf und $f + g$ messbar, daher ist nur die Endlichkeit des L^2-Integrals zu zeigen. Für cf ist das offensichtlich. Für $f + g$ verwende, dass für Zahlen $a, b \in \mathbb{C}$ gilt
$$|a + b|^2 \leq 2(|a|^2 + |b|^2) \quad \text{(Beweis: Übung).}$$
Also
$$\int_A |f(x) + g(x)|^2 \, dx \leq \int 2(|f(x)|^2 + |g(x)|^2) \, dx = 2\|f\|_2^2 + 2\|g\|_2^2 < \infty.$$
Schließlich ist $L^2(A) = \mathcal{L}^2(A)/\mathcal{N}$, wobei $\mathcal{N}$ der Untervektorraum der Nullfunktionen auf A ist, also ist auch $L^2(A)$ ein Vektorraum.

(2) Zunächst ist zu zeigen, dass $\int_A f\overline{g}$ für $f, g \in \mathcal{L}^2(A)$ existiert. Da mit f und g auch $f\overline{g}$ messbar ist, ist nur $\int_A |f\overline{g}| < \infty$ zu zeigen, und dies folgt aus der für alle komplexen Zahlen a, b gültigen Ungleichung
$$|a\overline{b}| \leq \frac{|a|^2 + |b|^2}{2} \quad \text{(Beweis: Übung).}$$
Weiterhin ändert sich $\int_A f\overline{g}$ nicht, wenn man f und g auf einer Nullmenge abändert, also ist $\langle f, g \rangle$ auch für $f, g \in L^2(A)$ wohldefiniert.

Es bleiben die Eigenschaften eines Skalarprodukts nachzuprüfen: Die Abbildung $f, g \mapsto \langle f, g \rangle$ ist

- ▷ linear bzgl. des ersten Arguments: $\langle f_1 + f_2, g \rangle = \langle f_1, g \rangle + \langle f_2, g \rangle, \quad \langle cf, g \rangle = c\langle f, g \rangle$
- ▷ antilinear bzgl. des zweiten Arguments: $\langle f, g_1 + g_2 \rangle = \langle f, g_1 \rangle + \langle f, g_2 \rangle, \quad \langle f, cg \rangle = \overline{c}\langle f, g \rangle$
- ▷ hermitesch: $\langle f, g \rangle = \overline{\langle g, f \rangle}$
- ▷ positiv definit: $\langle f, f \rangle \geq 0 \ \forall f$, und aus Gleichheit folgt $f = 0$.

Diese Eigenschaften folgen unmittelbar aus der Definition. Bei der Definitheit verwendet man, dass aus $\int_A |f|^2 = 0$ folgt, dass $|f|^2 = 0$ fast überall gilt.

(3) und

(4) Dies sind allgemeine Eigenschaften von Skalarprodukten.

(5) Siehe z.B. Königsberger II, 8.1 für die analoge Aussage für L^1. Der Beweis für L^2 läuft ähnlich. (Diese Aussage ist auch als Satz von Riesz-Fischer bekannt.) □

Bemerkung: Wir können uns also quadratintegrierbare Funktionen als Vektoren in einem riesigen (unendlichdimensionalen) Raum vorstellen, für die Begriffe wie Orthogonalität und Länge definiert sind. Dabei sagt man natürlich

$$f, g \in L^2(A) \text{ sind orthogonal } :\Leftrightarrow \langle f, g\rangle = 0, \text{ d.h. } \int_A f \cdot \overline{g} = 0.$$

Die Situation ist ähnlich wie beim Beweis der Existenz von Lösungen von Differentialgleichungen mittels des Fixpunktsatzes. Dort hatten wir Funktionen als Punkte in einem riesigen metrischen Raum betrachtet. Dies war der normierte Vektorraum der stetigen beschränkten Funktionen, mit der Supremumsnorm.

Beachten Sie, dass selbst für Funktionen, die sowohl stetig und beschränkt als auch quadratintegrierbar sind, die Konvergenzbegriffe bzgl. der Supremumsnorm und bzgl. der L^2-Norm verschieden sind. D.h., es gibt Folgen f_k, die gleichmäßig gegen ein f konvergieren, aber nicht in $L^2(A)$ (jedenfalls wenn A unendliches Maß hat), und umgekehrt. Der Konvergenzbegriff in $L^1(A)$ ist wieder ein anderer.

Verschiedene Probleme lassen sich mit verschiedenen Konvergenzbegriffen am besten behandeln. Für Fourier-Reihen ist L^2 am besten!

Im Kontext der Fourier-Reihen ist es bequemer, das Skalarprodukt etwas zu modifizieren:
Von jetzt an bezeichnen wir $L^2(-\pi, \pi)$ auch mit $L^2(S^1)$ und verwenden das Skalarprodukt

$$\langle f, g\rangle_F := \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)\overline{g(x)}\, dx$$

Weiter mit Königsberger S. 335 (aber immer mit L^2 statt Regelfunktionen), aber ohne Besselsche Ungleichung; dann Satz Seite 340.

5.5.4 Satz

a) $f \in L^2(S^1) \Longrightarrow \sum_{k=-\infty}^{\infty} |\widehat{f}(k)|^2 < \infty$

b) Seien $c_k \in \mathbb{C}, k \in \mathbb{Z}$ gegeben mit $\sum_{k=-\infty}^{\infty} |c_k|^2 < \infty$. Dann gibt es genau ein $f \in L^2(S^1)$ mit $\widehat{f}(k) = c_k \forall k$

Beweis:

$$f := \sum_{k=-\infty}^{\infty} c_k\, e^{ikx} = \lim_{N\to\infty} \underbrace{\sum_{k=-N}^{N} c_k\, e^{ikx}}_{=:f_N}$$

zu zeigen: $\lim_{N\to\infty} f_N$ existiert.

Für $M > N$ gilt:

$$\|f_M - f_N\|_2^2 = \| \sum_{N<|k|\le M} c_k\, e^{ikx}\|_2^2 \overset{Pythagoras}{=} \sum_{N<|k|\le M} |c_k|^2$$

Wegen $\sum_{N<|k|\le M} |c_k|^2 < \infty$ gibt es zu jedem $\varepsilon > 0$ ein N_0, so dass

$$\forall\, M, N \ge N_0 : \sum_{N<|k|\le M} c_k < \varepsilon$$

Also $\|f_M - f_N\|_2^2 < \varepsilon$ für $M, N \geq N_0$. Das heißt (f_N) ist eine Cauchy-Folge in $L^2(S^1)$. Da $L^2(S^1)$ vollständig ist, hat (f_N) einen Grenzwert $f \in L^2(S^1)$.

Behauptung: Für dieses f ist $\widehat{f}(k) = c_k \forall k$.

Beweis: Aus $\sum c_k\, e^{ikx} = f = \sum \widehat{f}(k)\, e^{ikx}$ folgt

$$0 = \sum (c_k - \widehat{f}(k))\, e^{ikx} \quad \Rightarrow \quad 0 = \|0\|_2^2 = \sum |c_k - \widehat{f}(k)|^2 \quad \Rightarrow \quad c_k = \widehat{f}(k) \forall k$$

Eindeutigkeit so ähnlich. □

Also: Bijektive Abbildung $L^2(S^1) \longrightarrow \underbrace{\left\{(c_k)_{k\in\mathbb{Z}} : \sum_{k=-\infty}^{\infty} |c_k|^2 < \infty\right\}}_{=:l^2(\mathbb{Z})}, f \longmapsto (\widehat{f}(k))_{k\in\mathbb{Z}}$

Allgemeine Frage:

Wie kann man Eigenschaften von $f \in L^2(S^1)$ an ihren Fourierkoeffizienten $(\widehat{f}(k))_{k\in\mathbb{Z}}$ ablesen?

$$f \in L^2(S^1) \Longleftrightarrow (\widehat{f}(k))_{k\in\mathbb{Z}} \in l^2(\mathbb{Z})$$

$$f \in C^\alpha(s^1) \Longleftrightarrow \quad ? \quad \text{Es gibt keine gute Antwort}$$

$$f \in C^\alpha(s^1) \overset{\text{Übung}}{\Longrightarrow} \exists C \forall k : |\widehat{f}(k)| \leq \frac{C}{|k|^\alpha}$$

$$\not\Leftarrow$$

a) A $n \times n$-Matrix. Falls A diagonalisierbar, dann existiert eine Basis $v_1, \cdots, v_n$ von $\mathbb{R}^n$ aus Eigenvektoren.

b) V n-dimensionaler l-Vektorraum, $A : V \longrightarrow V$ lineare Abbildung, A diagonalisierbar, dann existiert eine Basis $v_1, \cdots, v_n$ von V aus Eigenvektoren.

zu a) Falls A symmetrisch ist (im Matrix-Fall), so ist A diagonalisierbar.

zu b) Falls auf V ein Skalarprodukt $\langle \cdot, \cdot \rangle$ gegeben ist und $\langle Av, w\rangle = \langle v, Aw\rangle \forall v, w \in V$ gilt (das heißt A ist selbstadjungiert), so ist A diagonalisierbar.

Hier: $V = C^\infty(S^1) \subset L^2(S^1)$, $A = \dfrac{d^2}{dx^2} : C^\infty(S^1) \longrightarrow C^\infty(S^1), f \longrightarrow f''$ lineare Abbildung

A selbstadjungiert:

$$\langle Af, g\rangle = \langle f, Ag\rangle \Longleftrightarrow \langle f'', g\rangle = \langle f, g''\rangle \Longleftrightarrow \int_{-\pi}^{\pi} f''(x)\, \overline{g}(x)\, dx = \int_{-\pi}^{\pi} f(x)\, \overline{g''}(x)\, dx$$

Stimmt, da:

$$\int_{-\pi}^{\pi} f''\, \overline{g} = \underbrace{f'\, \overline{g}\Big|_{-\pi}^{\pi}}_{=0,\ \text{da } f,g\ 2\pi\text{-periodisch}} - \int_{-\pi}^{\pi} f'\, \overline{g}' = -\int_{-\pi}^{\pi} f'\, \overline{g}' = \int_{-\pi}^{\pi} f\, \overline{g}''$$

Bestimmung der Eigenwerte und Eigenvektoren: Allgmein: suche $\lambda \in \mathbb{C}, v \in V$ mit $Av = \lambda v$. Hier $v = f, f'' = \lambda f, f$ 2π-periodisch.

Löse die DGL $f'' - \lambda f = 0$.

Charakteristisches Polynom ist $p(z) = z^2 - \lambda$.

Lösungen der DGL sind gegeben durch $f(x) = ae^{z_1 x} + be^{z_2 x}$, wobei $a, b \in \mathbb{C}$ und z_1, z_2 Nullstellen von p sind.

Fall ($\lambda > 0$): $z_1 = \sqrt{\lambda}$, $z_2 = -\sqrt{\lambda}$ $\Rightarrow$ $f(x) = ae^{\sqrt{\lambda}x} + be^{-\sqrt{\lambda}x}$ nicht 2π-periodisch, falls $(a,b) \neq (0,0)$. (Grenzwerte für $x \to \pm\infty$)

Fall ($\lambda = 0$): $f'' = 0$ $\Rightarrow$ f linear $\Rightarrow$ f konstant $\Rightarrow$ f periodisch

Fall ($\lambda < 0$): Gelte $\lambda = -\mu^2, \mu \in \mathbb{R}$. Dann: $z^2 - \lambda = 0 \Leftrightarrow z^2 + \mu^2 = 0 \Rightarrow z_1 = i\mu, z_2 = -i\mu \Rightarrow f(x) = ae^{i\mu x} + be^{-i\mu x}$. Das ist genau periodisch mit Periode $\frac{2\pi}{\mu}$. Soll f 2π-periodisch sein, so wähle $\mu = 1, 2, 3, \ldots$.

Resultat: $f'' = \lambda f$ hat Lösungen, die 2π-periodisch sind, genau für $\lambda = 0, -1^2, -2^2, -3^2, \ldots$ ($\lambda = -k^2, k \in \mathbb{N}_0$). Lösungen sind e^{ikx}, e^{-ikx} und deren Linearkombinationen für $\lambda = -k^2$.

Also: Die Funktionen e^{ikx} sind »alle« (bis auf Linearkombinationen von $\pm k$) Eigenfunktionen (=Eigenvektoren) des Operators $A = \frac{d^2}{dx^2}$ auf $V = C^\infty(S^1)$.

Unserer Satz über Fourierreihen sagt, dass Lineare-Algebra-Satz auch hier stimmt (und in modifizierter Form: unendliche Summen). Jedes $f \in V$ lässt sich als $f = \sum_{k=-\infty}^{\infty} c_k\, e^{ikx}$ schreiben. (Funktionalanalysis)

zu vorher: $\lambda \in \mathbb{C}, Af = \lambda f$: $\langle Af, g\rangle = \langle f, Ag\rangle \forall f, g \Rightarrow \lambda \in \mathbb{R}$

Beweis:

$f = g$: $\langle \lambda f, f\rangle = \langle f, \lambda f\rangle \Rightarrow \lambda\langle f, f\rangle = \overline{\lambda}\langle f, f\rangle \overset{f\neq 0}{\Longrightarrow} \lambda = \overline{\lambda} \Rightarrow \lambda \in \mathbb{R}$

5.6 Fouriertransformation

5.6.1 Definition

Sei $f \in \mathcal{L}^1(\mathbb{R}^n)$. Die Fouriertransformation von f ist die Funktion $\widehat{f} : \mathbb{R}^n \longrightarrow \mathbb{C}$, definiert durch

$$\widehat{f}(\xi) = \int e^{-ix\cdot\xi} f(x)\, dx$$

Hierbei ist $x \cdot \xi := \sum_{j=1}^n x_j \xi_j$.

Wegen $|e^{-ix\cdot\xi}| = 1$ ist mit f auch $e^{-ix\cdot\xi} f$ integrierbar, also ist das Integral definiert.

Beispiel: Für $f = \chi_{[-1,1]}$ ist

$$\widehat{f}(\xi) = \int e^{-ix\cdot\xi} \chi_{[-1,1]}(x)\, dx = \int_{-1}^{1} e^{-ix\xi}\, dx = \left.\frac{e^{-ix\xi}}{-i\xi}\right|_{x=-1}^{1} = 2\,\frac{\sin(\xi)}{\xi}$$

Beachte: In manchen Büchern ist die Fouriertransformation mit einem Vorfaktor, z.B. als $\frac{1}{(2\pi)^{n/2}} \int e^{-ix\cdot\xi} f(x)\, dx$ definiert. Das bewirkt, dass in einigen der Formeln, die wir zeigen werden, die 2π-Faktoren an anderen Stellen stehen, ist aber ansonsten irrelevant.

Wir leiten nun einige grundlegende Eigenschaften der Fouriertransformation her.

5.6.2 Satz

Für $j = 1, \ldots, n$ gilt

a) $\widehat{(\partial_{x_j} f)}(\xi) = i\xi_j \widehat{f}(\xi)$ (falls $f \in \mathcal{L}^1 \cap C^1, \partial_{x_j} f \in \mathcal{L}^1$)

b) $\widehat{(x_j f)}(\xi) = i\partial_{\xi_j} \widehat{f}(\xi)$ (falls $f, x_j f \in \mathcal{L}^1$)

Kurz: Fouriertransformation überführt Ableitung in Multiplikation mit einer Koordinatenfunktion, und umgekehrt (bis auf Faktoren i).

Beweis: Wir geben den Beweis nur für $n = 1$, für beliebige n geht es analog.
b) $\frac{d}{d\xi}\widehat{f}(\xi) = \frac{d}{d\xi}\int e^{-ix\xi}f(x)\,dx = \int \frac{d}{d\xi}(e^{-ix\xi}f(x))\,dx = \int -ixe^{-ix\xi}f(x)\,dx = -i\cdot\widehat{(xf)}(\xi)$
Vertauschen von Ableitung und Integral ist erlaubt, da der abgeleitete Integrand durch die integrierbare Funktion $|xf(x)|$ majorisiert ist.
a) $\widehat{\left(\frac{d}{dx}f\right)}(\xi) = \int e^{-ix\xi}\frac{d}{dx}f(x)\,dx \overset{\text{partielle Integration}}{=} -\int \underbrace{\frac{d}{dx}(e^{-ix\xi})}_{=-i\xi e^{-ix\xi}} f(x)\,dx + \underbrace{\text{Randterme}}_{=0} = i\xi\int e^{-ix\xi}f(x)\,dx =$
$i\xi\widehat{f}(\xi)$. Warum sind bei der partiellen Integration die Randterme gleich null? Dafür schreiben wir $\int_{-\infty}^{\infty}\ldots = \lim_{N\to\infty}\int_{-N}^{N}\ldots$, die Randterme für festes N sind dann $e^{-ix\xi}f(x)\big|_{x=-N}^{N}$, und dies strebt für $N\to\infty$ gegen Null, denn es gilt: $f\in\mathcal{L}^1(\mathbb{R})\cap C^1(\mathbb{R}), f'\in\mathcal{L}^1(\mathbb{R}) \Rightarrow f(x)\to 0$ für $|x|\to\infty$ (Beweis als Übung). □

5.6.3 Satz
Sei $f(x) = e^{-\frac{\|x\|^2}{2}}$ auf $\mathbb{R}^n$. Dann ist $\widehat{f}(\xi) = (2\pi)^{n/2}e^{-\frac{\|\xi\|^2}{2}}$.

Beweis: Sei zunächst $n = 1$. Wir müssen $\int e^{-ix\xi}\cdot e^{-\frac{x^2}{2}}\,dx$ berechnen. Da wir keine Stammfunktion des Integranden angeben können, verwenden wir folgende elegante indirekte Methode. Es ist

$$f(x) = e^{-\frac{x^2}{2}} \Rightarrow f'(x) = -xe^{-\frac{x^2}{2}} = -xf(x)$$

Also $f'(x) + xf(x) = 0$. Wende Fouriertransformation auf die Gleichung an:

$$\widehat{f'+xf} = \widehat{0} = 0 \quad\Longrightarrow\quad \widehat{f'} + \widehat{xf} = 0 \quad\Longrightarrow\quad i\xi\widehat{f} + i\frac{d}{d\xi}\widehat{f} = 0$$
$$\Longrightarrow \quad \text{die Funktion } g(\xi) := \widehat{f}(\xi) \text{ erfüllt } \xi g(\xi) + g'(\xi) = 0$$

Das ist eine lineare homogene Differentialgleichung erster Ordnung. Also ist der Lösungsraum eindimensional. Eine Lösung ist $e^{-\frac{\xi^2}{2}}$. Also $g(\xi) = c\cdot e^{-\frac{\xi^2}{2}}$ für ein $c\in\mathbb{C}$. Was ist c?

$$c = g(0) = \widehat{f}(0) = \int e^{-ix0}f(x)\,dx = \int f(x)\,dx = \int e^{-\frac{x^2}{2}}\,dx = \sqrt{2\pi}$$

Für beliebiges n ist $e^{-ix\cdot\xi}e^{-\frac{\|x\|^2}{2}} = e^{-ix_1\xi_1-\cdots-ix_n\xi_n}e^{-\frac{x_1^2}{2}-\cdots-\frac{x_n^2}{2}} = e^{-ix_1\xi_1}e^{-\frac{x_1^2}{2}}\cdots e^{-ix_n\xi_n}e^{-\frac{x_n^2}{2}}$, und die Behauptung folgt aus dem Fall $n = 1$ mit Fubini. □

Kann man f aus $\widehat{f}$ zurückerhalten? Ja, mit folgender Formel. Zunächst erstaunlich ist, dass es fast dieselbe Formel ist wie für die Fouriertransformation selbst.

5.6.4 Satz (Umkehrsatz für die Fouriertransformation)
Sei $f\in L^1(\mathbb{R}^n)\cap C^0(\mathbb{R}^n)$ beschränkt, und sei $\widehat{f}\in L^1(\mathbb{R}^n)$. Dann gilt für alle x:

$$f(x) = \frac{1}{(2\pi)^n}\int e^{ix\cdot\xi}\widehat{f}(\xi)\,d\xi$$

Beweis: Sei $x\in\mathbb{R}^n$ fest. Wir würden gerne rechnen:

$$\int e^{ix\cdot\xi}\widehat{f}(\xi)\,d\xi \overset{\text{def}}{=} \int\int e^{ix\cdot\xi}e^{-iy\cdot\xi}f(y)\,dy\,d\xi \overset{\text{Fubini}}{=} \int f(y)\left(\int e^{-i(y-x)\cdot\xi}\,d\xi\right)dy\,,$$

doch leider ist das innere Integral nicht definiert und damit auch Fubini nicht anwendbar, da $|e^{-i(y-x)\cdot\xi}| = 1$ für alle ξ ist. Zur Rettung führen wir einen ‚konvergenzerzeugenden Faktor' ein. Dazu wähle eine beschränkte Funktion $\varphi \in \mathcal{L}^1(\mathbb{R}^n) \cap C^0(\mathbb{R}^n)$ mit $\varphi(0) = 1$ und $\widehat{\varphi} \in \mathcal{L}^1(\mathbb{R}^n)$. Sei $\varepsilon > 0$. Dann ist auch $\xi \mapsto \varphi(\varepsilon\xi)$ integrierbar, und dieselbe Rechnung wie vorher ergibt

$$\int e^{ix\cdot\xi}\widehat{f}(\xi)\varphi(\varepsilon\xi)\,d\xi = \int f(y)\left(\int e^{-i(y-x)\cdot\xi}\varphi(\varepsilon\xi)\,d\xi\right)dy$$

Mit der Variablensubstitution $\eta = \varepsilon\xi$ wird das innere Integral zu $\varepsilon^{-n}\widehat{\varphi}(\frac{y-x}{\varepsilon})$. Nun substituieren wir $z = \frac{y-x}{\varepsilon}$ im y-Integral, also $y = x + \varepsilon z$. Der Variablenwechsel erzeugt einen Faktor ε^n, der das ε^{-n} wegkürzt. Insgesamt folgt

$$\int e^{ix\cdot\xi}\widehat{f}(\xi)\varphi(\varepsilon\xi)\,d\xi = \int f(x+\varepsilon z)\widehat{\varphi}(z)\,dz$$

Dies gilt für jedes $\varepsilon > 0$. Wir bestimmen nun den Grenzwert $\varepsilon \to 0$ beider Seiten mit Hilfe des Satzes über die majorisierte Konvergenz:

$$\begin{array}{ccc} \int e^{ix\cdot\xi}\widehat{f}(\xi)\varphi(\varepsilon\xi)\,d\xi & = & \int f(x+\varepsilon z)\widehat{\varphi}(z)\,dz \\ \downarrow & & \downarrow \\ \int e^{ix\cdot\xi}\widehat{f}(\xi)\,d\xi & & \int f(x)\widehat{\varphi}(z)\,dz \end{array}$$

Denn wegen $\varphi(0) = 1$ konvergiert der linke Integrand für $\varepsilon \to 0$ punktweise gegen $e^{ix\cdot\xi}\widehat{f}(\xi)$, und er ist durch die integrierbare Funktion $C|\widehat{f}|$ majorisiert, wobei C eine obere Schranke für $|\varphi|$ ist. Rechts konvergiert der Integrand ebenfalls punktweise und ist durch die integrierbare Funktion $C'|\widehat{\varphi}|$ majorisiert, wobei C' eine obere Schranke für $|f|$ ist.

Wegen $\int f(x)\widehat{\varphi}(z)\,dz = f(x)\int\widehat{\varphi}(z)\,dz$ folgt der Satz, wenn wir $\int\widehat{\varphi}(z)\,dz = (2\pi)^n$ zeigen können. Hierfür wählen wir konkret $\varphi(\xi) = e^{-\frac{\|\xi\|^2}{2}}$. Dann ist $\varphi(0) = 1$, und die anderen Eigenschaften sind offenbar erfüllt. Es ist $\widehat{\varphi}(z) = (2\pi)^{n/2}e^{-\frac{\|z\|^2}{2}}$ und damit $\int\widehat{\varphi}(z)\,dz = (2\pi)^{n/2}\int e^{-\frac{\|z\|^2}{2}}\,dz = (2\pi)^{n/2}(2\pi)^{n/2} = (2\pi)^n$, was zu zeigen war. □

Die Fouriertransformation verträgt sich in sehr hübscher Weise mit der Faltung:

5.6.5 Satz

Seien $f, g \in \mathcal{L}^1(\mathbb{R}^n)$. Dann ist $\widehat{f * g} = \widehat{f}\,\widehat{g}$.

Fouriertransformation ‚übersetzt' Faltung also in Multiplikation.

Beweis: $\widehat{f*g}(x) = \int e^{-ix\cdot\xi}(f*g)(x)\,dx = \int\int e^{-ix\cdot\xi}f(y)g(x-y)\,dy\,dx = \int\int e^{-i(y+z)\cdot\xi}f(y)g(z)\,dy\,dz = \int e^{-iy\cdot\xi}f(y)\,dy\int e^{-iz\cdot\xi}g(z)\,dz$, wobei die schon bekannte Variablensubstitution $(x,y) \mapsto (z = x-y, y)$ durchgeführt wurde. □

Anwendung der Fouriertransformation auf die Lösung partieller Differentialgleichungen (linear, mit konstanten Koeffizienten)

Da hier nur die Idee gezeigt werden soll, rechnen wir drauflos, ohne uns um Integrierbarkeit etc. zu kümmern.

Als Beispiel betrachten wir die partielle Differentialgleichung $u - \Delta u = f$, wobei u, f Funktionen auf $\mathbb{R}^n$, f gegeben und u gesucht sind.

Trick: Fouriertransformiere beide Seiten:

$$\widehat{\partial_{x_j}u} = i\xi_j\widehat{u}$$

$$\widehat{(\partial_{x_j}^2 u)} = (\widehat{\partial_{x_j}\partial_{x_j} u}) = i\xi_j \widehat{(\partial_{x_j} u)} = (i\xi_j)^2 \widehat{u} \quad \Longrightarrow \quad \widehat{\Delta u} = \widehat{\sum_{j=1}^{n} \partial_{x_j}^2 u} = \sum_{j=1}^{n} (i\xi_j)^2 \widehat{u} = -|\xi|^2 \widehat{u}$$

Also

$$u - \Delta u = f \iff \widehat{u - \Delta u} = \widehat{f} \iff \widehat{u} - \widehat{\Delta u} = \widehat{f}$$
$$\iff \widehat{u} + |\xi|^2 \widehat{u} = \widehat{f} \iff (1 + |\xi|^2)\widehat{u} = \widehat{f} \iff \widehat{u}(\xi) = \frac{1}{1+|\xi|^2} \widehat{f}(\xi)$$

Die Lösung der DGL ist daher

$$u(x) = \frac{1}{(2\pi)^n} \int e^{ix\cdot\xi} \frac{1}{1+|\xi|^2} \widehat{f}(\xi)\, d\xi$$

Mit Hilfe von Satz 5.6.5 und Satz 5.6.4 lässt sich das auch so schreiben:

$$u(x) = K * f,\ K(x) = \frac{1}{(2\pi)^n} \int e^{ix\cdot\xi} \frac{1}{1+|\xi|^2} d\xi$$

Das funktioniert, weil durch Fouriertransformation Ableitung in Multiplikation verwandelt wird!

Analog kann man Formeln für beliebige lineare partielle Differentialgleichungen mit konstanten Koeffizienten hinschreiben. Allerdings gibt es dabei manchmal Probleme mit null im Nenner, z.B. schon bei der uns wohlbekannten Gleichung $\Delta u = f$: Dies ergibt

$$u(x) = \frac{1}{(2\pi)^n} \int e^{i\langle x,\xi\rangle} \frac{1}{-|\xi|^2} \widehat{f}(\xi)\, d\xi$$

Hier sieht man, dass es bei $\xi = 0$ ein Problem gibt. In $n \geq 3$ ist $\|\xi\|^{-2}$ jedoch lokal integrierbar, und alles ist ok. Ein ähnliches Problem entsteht bei der Formel $K(x) = \frac{1}{(2\pi)^n} \int e^{ix\cdot\xi} \frac{1}{1+|\xi|^2} d\xi$ im Beispiel oben: Dies ist für $n \geq 2$ nicht integrierbar, da der Integrand für $\xi \to \infty$ nicht schnell genug gegen Null geht!

All diese Probleme lassen sich in geeignetem Sinn lösen, am elegantesten im Rahmen der Distributionentheorie. Dazu mehr in der Vorlesung zu partiellen Differentialgleichungen.