

Analysis III

Universität Mannheim

Dozent: H.-J. Bartels

Überarbeitung einer Ausarbeitung von P. Höpner durch D. Boldin

Inhaltsverzeichnis

1	Gewöhnliche Differentialgleichungen	5
1.1	Einführung, Grundbegriffe und Beispiele von Differentialgleichungen	5
1.2	Differentialgleichungen mit getrennten Veränderlichen	11
1.3	Elementare Lösungsmethoden	17
1.4	Existenz- und Eindeutigkeitssätze	21
1.5	Lösung durch Potenzreihenansatz	31
1.6	Lineare Differentialgleichungen	34
1.7	Lineare Differentialgleichungssysteme mit konstanten Koeffizienten	42
2	Kurven- und Flächenintegrale, Vektoranalysis	50
2.1	Orientierte Flächen	50
2.2	Flächenintegrale	54
2.3	Die Grassmannalgebra eines Vektorraums	59
2.4	Tangentialvektoren, Derivationen und Pfaffsche Formen	62
2.5	Alternierende Differentialformen	66
2.6	Der Stokessche Integralsatz	70
2.7	Das Poincarésche Lemma	73
2.8	Vektoranalysis im dreidimensionalen Raum	74
3	Funktionsräume	78
3.1	Banachräume und Hilberträume	78
3.2	Fourierreihen	82
3.3	Fourier-Integrale	84
3.4	Distributionen	85
A	Integrationstheorie	96
A.1	Der Approximationssatz von Stone-Weierstraß	96
A.2	Das Integral für stetige Funktionen mit kompaktem Träger	106
A.3	Die Ausdehnung des Integrals auf halbstetige Funktionen	114
A.4	Der Daniell-Lebesgue-Prozess, Teil 1: Das Integral für halbstetige Funktionen	115
A.5	Der Daniell-Lebesgue-Prozess, Teil 2: Das äußere Integral	124
A.6	Die integrierbaren Funktionen	130
A.7	Integrierbarkeitskriterien	137
A.8	Nullmengen	140
A.9	Der Satz von Fubini	143
A.10	Die Transformationsformel	147

A.11 Messbarkeit	156
B Integralrechnung von Funktionen mehrerer Veränderlicher	160
B.1 Jordanscher Inhalt und Lebesguesches Maß	160
B.2 Das Lebesguesche Integral	166
B.3 Grenzwertsätze	170
B.4 Der Satz von Fubini und das Cavalierische Prinzip	173
B.5 Transformation von Integralen	179

Einleitung

Im ersten Kapitel geht es um die Behandlung von gewöhnlichen Differentialgleichungen. Zusätzlich zu den allgemeinen Existenz- und Eindeutigkeitsaussagen werden auch spezielle Typen eingehend behandelt. Neben den Differentialgleichungen mit getrennten Veränderlichen und solchen, die sich auf diese zurück führen lassen, gehören vor allem die linearen Differentialgleichungssysteme dazu.

Im zweiten Kapitel werden Kurven- und Flächenintegrale behandelt. Um die Vektoranalysis und den Satz von Stokes in allgemeiner Form behandeln zu können, wird dazu in Kapitel 2.5 der Kalkül der alternierenden Differentialformen behandelt.

Das abschließende Kapitel 3 behandelt Funktionenräume, die für die Fourierttransformation benötigt werden.

Da sowohl im zweiten, wie auch im dritten Kapitel die Integrationstheorie von Funktionen mehrerer Veränderlicher benötigt wird, sind im Anhang zwei Zugänge zur n-dimensionalen Integrationstheorie dargestellt. Hier finden sich insbesondere alle wichtigen Konvergenzsätze, wie auch die zum Rechnen mit mehrfachen Integralen benötigten Regeln (Satz von Fubini, Transformationsformel).

Hinweise auf Druckfehler oder Verbesserungsvorschläge nehmen wir gerne entgegen unter: bartels@math.uni-mannheim.de.

1 Gewöhnliche Differentialgleichungen

1.1 Einführung, Grundbegriffe und Beispiele von Differentialgleichungen

Bei einer **Differentialgleichung** (kurz DGL) handelt es sich um eine Gleichung zwischen einer gesuchten Funktion y , gewissen ihrer Ableitungen und dem bzw. den Argumenten von y , d.h. der oder den unabhängigen Veränderlichen x (bzw. x_1, \dots, x_n). Wir beschränken uns hier auf reelle Differentiation, und demzufolge bezeichnet x (bzw. x_1, \dots, x_n) stets eine **reelle** Veränderliche.

Aus der Analysis I kennt man folgendes Beispiel: Zu einer gegebenen Funktion $f(x)$ wurde eine Funktion $y = y(x)$ gesucht, so dass gilt

$$\frac{dy}{dx} = f(x). \quad (1.1)$$

Der Hauptsatz der Differential- und Integralrechnung liefert dann für stetiges f Lösungen in der Form

$$y(x) = \int_{x_0}^x f(t)dt + c. \quad (1.2)$$

Selbst wenn sich das Integral nicht geschlossen auswerten lässt, betrachtet man (1.1) durch (1.2) als gelöst.

Ein weiteres Beispiel ist: Im Zusammenhang mit exponentiellem Wachstum von Populationen sucht man Funktionen $P(t)$ mit

$$\frac{dP}{dt} = c \cdot P(t). \quad (1.3)$$

Die zur Zeit t vorliegende Wachstumsgeschwindigkeit oder Änderungsrate $\frac{dP}{dt}$ der Population ist proportional zur vorhandenen Populationsgröße $P(t)$. Die mathematische Aufgabe besteht darin, sämtliche Funktionen $P(t)$ zu finden, die der Differentialgleichung (1.3) genügen.

Offenbar sind die Funktionen $P(t) = k \cdot e^{ct}$ mit einer beliebigen reellen Konstanten k Lösungen von (1.3). Unten werden wir zeigen, dass umgekehrt **jede** Lösung von (1.3) diese Form hat. Legt man die Größe der Population zu einem Anfangszeitpunkt - etwa $t_0 = 0$ - als P_0 fest, so ist also zwangsläufig

$$P(t) = P_0 \cdot e^{ct}.$$

Die Anfangsgröße P_0 der Population und ihre Veränderungsrate (1.3) legen also eindeutig ihre Größe zu jedem Zeitpunkt $t \geq 0$ fest. Dies bedeutet: Aus der lokalen Betrachtung folgt das globale Wachstumsgesetz.

Ein weiteres Beispiel liefert die Bewegung eines Massenpunktes der Masse m unter Einwirkung einer Gravitationskraft $g(x)$. Für die Bahn $x(t)$ des Punktes gilt dann:

$$m \cdot \ddot{x} = g(t), \quad (1.4)$$

wobei der Doppelpunkt die zweimalige Ableitung nach der Zeit t bezeichnet.

Treten bei einer Differentialgleichung partielle Ableitungen auf, so spricht man von einer **partiellen Differentialgleichung**. In dieser Vorlesung beschränken wir uns auf Differentialgleichungen für Funktionen von nur **einer** unabhängigen Veränderlichen. Man spricht in diesem Fall auch von **gewöhnlichen** Differentialgleichungen.

Die Ordnung der höchsten auftretenden Ableitung heißt **Ordnung** der DGL. Eine gewöhnliche DGL n -ter Ordnung ist also eine DGL der Form

$$F(x, y, y', \dots, y^{(n)}) = 0. \quad (1.5)$$

Sind mehrere Funktionen $y_1(x), \dots, y_m(x)$ gesucht, so fasse man diese zu einem Vektor $y = (y_1(x), \dots, y_m(x))$ zusammen. Hat man dann mehrere Gleichungen

$$F_i(x, y, y', \dots, y^{(n)}) = 0 \quad \text{für } i = 1, \dots, m$$

vorliegen, so sprechen wir von einem **Differentialgleichungssystem n -ter Ordnung**.

Eine DGL heißt **explizit**, wenn sie nach der Ableitung höchster Ordnung aufgelöst ist:

$$y^{(n)} = f(x, y, y', \dots, y^{(n-1)}). \quad (1.6)$$

Falls eine Gleichung der Form (1.5) vorliegt, spricht man von einer DGL in **impliziter** Form.

Analoge Sprechweisen gelten für Differentialgleichungssysteme.

Man kann nun jede explizite DGL n -ter Ordnung

$$y^{(n)} = f(x, y, y', \dots, y^{(n-1)})$$

in ein DGL-System 1. Ordnung umwandeln. Man setze dazu:

$$\begin{cases} y_0(x) := y(x) \\ y_1(x) := y'(x) \\ \vdots \\ y_{n-1}(x) := y^{(n-1)}(x). \end{cases}$$

Dann ist (1.6) offenbar gleichbedeutend mit:

$$\begin{cases} y'_0 = y_1 \\ y'_1 = y_2 \\ \vdots \\ y'_{n-2} = y_{n-1} \\ y'_{n-1} = f(x, y_0, \dots, y_{n-1}). \end{cases} \quad (1.7)$$

Manchmal kann man DGL-Systeme erster Ordnung einfacher behandeln, als dies bei DGLn n -ter Ordnung der Fall ist.

Explizite DGL-Systeme erster Ordnung lassen eine geometrische Interpretation zu, die auch für explizite Lösungsverfahren später nützlich sein wird:

Eine Gleichung $y' = f(x, y)$ (unter Umständen mit vektorwertigem $y \in \mathbb{R}^n$) bedeutet: In jedem Punkt $(x, y) \in \mathbb{R}^{n+1}$, in dem f definiert ist, wird eine Richtung vorgegeben. Eine Lösung des DGL-Systems entspricht geometrisch einer Kurve, die in jedem Punkt des Definitionsbereichs von f die durch $f(x, y)$ vorgegebene Richtung hat. Die Linien $f(x, y) = c$ gleicher Richtung heißen **Isoklinen**. Das Isoklinenfeld gibt Aufschluss über Symmetrieverhältnisse der Lösungsgesamtheit.

Beispiel 1.1. Sei

$$y' = x \cdot y.$$

Die Isoklinen sind Hyperbeln, die Lösungskurven sind symmetrisch zur y -Achse und zur x -Achse.

Intuitiv würde man in diesem Beispiel erwarten, dass durch jeden Punkt $(x_0, y_0) \in \mathbb{R}^2$ eine Lösungskurve geht. Dies wird später noch im allgemeinen Kontext bewiesen werden.

Unter einem **Anfangswertproblem** versteht man bei einem DGL-System erster Ordnung die Aufgabe, zu vorgegebenem (x_0, y_0) eine Lösung $y = y(x)$ mit $y(x_0) = y_0$ zu finden.

Bei DGLn n -ter Ordnung $y^{(n)} = f(x, y, \dots, y^{(n-1)})$ besteht die Aufgabe eines Anfangswertproblems darin, zu vorgegebenen Werten $x_0, y_0, y'_0, \dots, y_0^{(n-1)}$ eine Lösung $y = y(x)$ zu finden, die

$$y(x_0) = y_0, \quad y'(x_0) = y'_0, \quad \dots, \quad y^{(n-1)}(x_0) = y_0^{(n-1)}$$

erfüllt.

Viele geometrische Probleme führen auf gewöhnliche Differentialgleichungen. Hier ein Beispiel:

Beispiel 1.2. Zu einer Kurvenschar im \mathbb{R}^2 werden die **orthogonalen Trajektorien** gesucht, d.h. diejenigen ebenen Kurven, die **jede** Kurve der gegebenen Schar senkrecht schneiden. Sind die Kurven gegeben durch

$$y = u(x) = u(x; p)$$

mit einem Scharparameter p und genügen diese Funktionen u einer DGL

$$F(x, u, u') = 0,$$

so gilt für die orthogonalen Trajektorien

$$y = v(x) = v(x; q)$$

folgende Differentialgleichung

$$F(x, v, -\frac{1}{v'}) = 0. \quad (1.8)$$

Begründung: Die durch u und v definierten Kurven schneiden sich senkrecht in $u(x_0) = v(x_0)$, wenn $u'(x_0) \cdot v'(x_0) = -1$ gilt. Genau in diesem Fall haben nämlich die Tangentenvektoren $(1, u'(x_0))$ und $(1, v'(x_0))$ das Skalarprodukt Null.

Die gesuchten orthogonalen Trajektorien ergeben sich als Lösungen der Differentialgleichungen (1.8).

Konkretes Beispiel 1.3. Gesucht sind die orthogonalen Trajektorien zu der Schar konfokaler Ellipsen mit den festen Brennpunkten $(c, 0)$ und $(-c, 0)$. Diese Ellipsen werden beschrieben durch

$$\frac{x^2}{c^2 + p} + \frac{y^2}{p} = 1 \quad \text{mit } p \in \mathbb{R}, p > 0. \quad (1.9)$$

Die Auflösung der Gleichung (1.9) nach y liefert die mit Ausnahme der Punkte auf der x -Achse differenzierbare Funktion

$$y = \pm u(x; p).$$

Differentiation von (1.9) liefert dann:

$$\frac{2x}{c^2 + p} + \frac{2u \cdot u'}{p} = 0 \quad \text{und} \quad \frac{x^2}{c^2 + p} + \frac{u(x, p)^2}{p} = 1.$$

Ist $x \neq 0$ und $y = u(x, p) \neq 0$, so gilt nun

$$\frac{2uu'}{p} = -\frac{2x}{c^2 + p}, \quad (c^2 + p) \cdot 2uu' = -2xp.$$

Damit ist $p \cdot (2uu' + 2x) = -c^2 \cdot 2uu'$, also $p = \frac{-c^2uu'}{uu'+x}$. Einsetzen in (1.9) ergibt

$$\frac{x^2}{c^2 - \frac{c^2uu'}{uu'+x}} + \frac{y^2}{-c^2 \frac{uu'}{uu'+x}} = 1,$$

d.h.

$$\frac{x^2(uu' + x)}{c^2x} - \frac{y^2 \cdot (uu' + x)}{c^2uu'} = 1.$$

Wegen $y = \pm u$ ist $y^2 = u^2$, und daraus ergibt sich

$$x(uu' + x) - \frac{u^2(uu' + x)}{uu'} = c^2,$$

d.h.

$$xuu' + x^2 - u^2 - x \frac{u}{u'} = c^2.$$

Wir haben damit das Ergebnis

$$(x + uu') \left(x - \frac{u}{u'} \right) = c^2. \quad (1.10)$$

Die Ellipsen genügen dieser Differentialgleichung. Ausgenommen sind dabei die Punkte auf den Koordinatenachsen. Für die orthogonalen Trajektorien $y = v(x)$ gilt dann:

$$\left(x - \frac{v}{v'} \right) (x + vv') = c^2.$$

Offenbar ist dies genau dieselbe DGL wie (1.10).

Man hat zwei Kurvenscharen, die dieser DGL genügen. Die erste Schar sind die Ellipsenbögen $y = \pm u(x; p)$ mit $p > 0$. Die zweite Schar erhält man, wenn man den Parameter p die negativen Werte zwischen 0 und $-c^2$ durchlaufen lässt ($-c^2 < p < 0$). Die Kurven dieser Schar werden ebenfalls durch (1.9) beschrieben. Es handelt sich hier um konfokale Hyperbeln; diese sind — wie man nachrechnen kann — auch die gesuchten Trajektorien. Die y -Achse und die Halbgeraden $x < -c$ bzw. $x > c$ sind ebenfalls orthogonale Trajektorien.

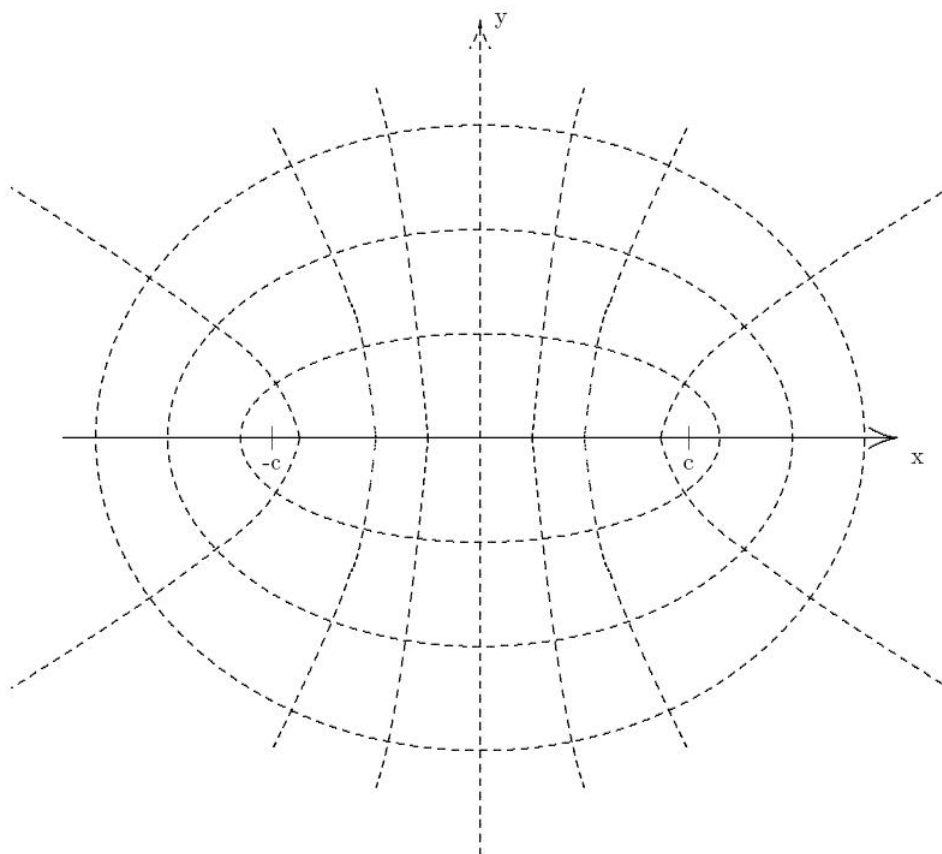


Abbildung 1: Konfokale Ellipsen mit Brennpunkten $(c,0)$ und $(-c,0)$, hier dargestellt zusammen mit einigen der Orthogonal-Trajektorien.

Bevor wir die allgemeine Theorie mit Existenz- und Eindeutigkeitsätzen behandeln, werden in den folgenden Abschnitten Klassen von DGLn untersucht, bei denen explizite Lösungsverfahren einfach angebar sind.

1.2 Differentialgleichungen mit getrennten Veränderlichen

Es geht hier um DGLn des folgenden Typs:

$$y' = f(x) \cdot g(y) \quad (1.11)$$

mit stetigen Funktionen $f: (a, b) \rightarrow \mathbb{R}$ und $g: (c, d) \rightarrow \mathbb{R}$. Es interessieren folgende Fragen:

- Existiert zu vorgegebenem $(x_0, y_0) \in (a, b) \times (c, d)$ stets eine Lösung $y = y(x)$ mit $y(x_0) = y_0$?
- Ist eine solche Lösung eindeutig bestimmt?

Zwei Beispiele vorweg:

Beispiel 1.4. Wir betrachten

$$y' = f(x) \cdot g(y) = y^2,$$

d.h. wir setzen $f(x) \equiv 1$ und $g(y) = y^2$. Gesucht wird eine Lösung $y = y(x)$ der Differentialgleichung mit der Anfangsbedingung $y(0) = c$ ($c \in \mathbb{R}$ fest). Wir werden später sehen, dass dann genau eine Lösung $y = y(x)$ existiert mit $y(0) = c$. Hier unterscheiden wir zunächst drei Fälle:

(i) Sei $c = 0$. Dann ist offenbar $y(x) \equiv 0$ eine Lösung der gegebenen DGL mit $y(0) = 0$.

(ii) Sei $c > 0$. Ist $y = y(x)$ Lösung von $y' = y^2$ mit $y(0) = c > 0$ auf einem Intervall I , das die 0 enthält, so ist $y(x) > 0$ in einer passenden Umgebung von 0. Dort gilt

$$\frac{y'(x)}{y(x)^2} \equiv 1,$$

also

$$\int_0^x \frac{y'(t)}{y(t)^2} dt = x$$

und

$$\int_c^y \frac{dy}{y^2} = \left[-\frac{1}{y} \right]_c^y = \frac{1}{c} - \frac{1}{y} =: G(y).$$

Es ist $G(\mathbb{R}_{>0}) = (-\infty, \frac{1}{c})$. Für eine Lösung $y = y(x)$ mit $y(0) = c$ gilt also zunächst lokal: $y^{-1} = \frac{1}{c} - x$, d.h.

$$y = y(x) = \frac{c}{1 - cx} \quad \text{für alle } x < \frac{1}{c}.$$

(iii) Sei $c < 0$. Man erhält analog eine Lösung $y : (\frac{1}{c}, \infty) \rightarrow \mathbb{R}$ mit

$$y = y(x) = \frac{c}{1 - cx} \quad \text{für } x > \frac{1}{c}.$$

Bemerkung: Obgleich die Differentialgleichung auf der ganzen Ebene definiert ist, lassen sich die Lösungen für $c \neq 0$ nicht auf ganz \mathbb{R} fortsetzen.

Beispiel 1.5. Sei

$$y' = \sqrt{|y|}.$$

Für $y \neq 0$ hat man $\frac{y'}{\sqrt{|y|}} = 1$, also ist

$$\pm 2\sqrt{|y|} = \int \frac{dy}{\sqrt{|y|}} = x + c \quad \text{mit einer Konstanten } c,$$

wobei das Pluszeichen im Fall $y > 0$ und das Minuszeichen im Fall $y < 0$ gilt. Es ist also

$$y = \begin{cases} \left(\frac{x+c}{2}\right)^2 & \text{für } x > -c \\ -\left(\frac{x+c}{2}\right)^2 & \text{für } x < -c. \end{cases} \quad (1.12)$$

Diese Parabelzweige können mit Teilstücken der x -Achse, die auch Lösungskurve ist, kombiniert werden. Durch jeden Anfangswert der Ebene gehen daher unendlich viele Lösungskurven.

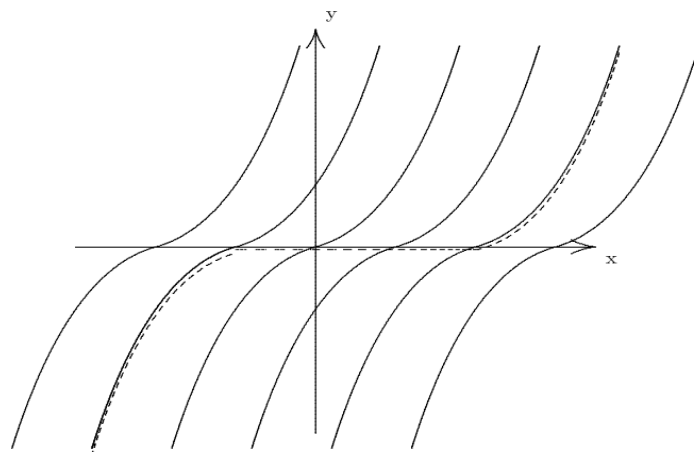


Abbildung 2: Die Zweige der Funktion (1.12), die $y' = \sqrt{|y|}$ löst, können mit Teilstücken der x -Achse kombiniert werden, wie es hier z.B. die gestrichelte Linie andeutet.

Nun zurück zur allgemeinen Theorie der DGLn mit getrennten Veränderlichen. Es werden zwei Fälle unterschieden:

1. Fall: Sei $g(y) \neq 0$ für alle $y \in (c, d)$. Dann folgt: Ist $y = y(x)$ Lösung von (1.11), so gilt

$$\frac{y'(x)}{g(y(x))} = f(x).$$

Mit $F(x) := \int_{x_0}^x f(t)dt$, $G(y) := \int_{y_0}^y \frac{dt}{g(t)}$ folgt dann also

$$G(y(x)) = F(x) + k$$

mit einer reellen Konstanten k . Ist $y(x_0) = y_0$, so ist hierbei $k = 0$.

Es ist $G(y)$ stetig und streng monoton (da g in (c, d) stetig und $\neq 0$ ist, ist g also stets > 0 oder < 0). Für die zu G existierende Umkehrfunktion H gilt also

$$y(x) = H(F(x) + k)$$

für alle x , für die $F(x) + k$ im Wertebereich von G liegt. Jede Lösung hat notwendig diese Gestalt, und umgekehrt zeigt man durch Differentiation, dass hierdurch wirklich eine Lösung gegeben ist.

Insgesamt haben wir damit den

Satz 1.6. $f: (a, b) \rightarrow \mathbb{R}$, $g: (c, d) \rightarrow \mathbb{R}$ seien stetige Funktionen, es gelte $g(y) \neq 0$ auf (c, d) . Dann geht durch jeden Punkt

$$(x_0, y_0) \in \{(x, y) : a < x < b, c < y < d\}$$

genau eine Lösungskurve $y(x)$ von (1.11). Es ist

$$y(x) = H\left(\int_{x_0}^x f(t)dt\right),$$

wenn H die Umkehrfunktion von $G(y) := \int_{y_0}^y \frac{dt}{g(t)}$ bezeichnet. $y(x)$ ist definiert für alle x , für die $F(x) = \int_{x_0}^x f(t)dt$ im Wertebereich B von G liegt. B ist das offene Intervall mit den Endpunkten

$$\int_{y_0}^c \frac{dt}{g(t)} \quad \text{und} \quad \int_{y_0}^d \frac{dt}{g(t)}.$$

Dabei wird $\int_{t_0}^{t_1} \varphi(t)dt = +\infty$ bzw. $-\infty$ gesetzt, falls dieses uneigentliche Integral gegen $+\infty$ bzw. $-\infty$ divergiert.

2. Fall: $g(y)$ habe in (c, d) mindestens eine Nullstelle.

Durch jeden Punkt (x_0, y_0) geht mindestens eine Lösungskurve von (1.11), denn: Ist $g(y_0) = 0$, so ist $y(x) = y_0$ offenbar Lösung. Ist $g(y_0) \neq 0$, so existiert wegen der Stetigkeit von g ein Teilintervall $(e, f) \subseteq (c, d)$ mit $y_0 \in (e, f) \subseteq (c, d)$ mit $g(y) \neq 0$ für $y \in (e, f)$. In diesem Teilintervall lässt sich Satz (1.6) anwenden, und es folgt der

Satz 1.7. $f: (a, b) \rightarrow \mathbb{R}$, $g: (c, d) \rightarrow \mathbb{R}$ seien stetige Funktionen. Durch jeden Punkt $(x_0, y_0) \in \{(x, y) : a < x < b, c < y < d\}$ geht mindestens eine Lösungskurve y von (1.11). Die Lösungskurven sind genau dann **nicht** für jeden Punkt eindeutig bestimmt, wenn $f \not\equiv 0$ ist und es ein auf einer Seite durch die Nullstelle $y_1 \in (c, d)$ von g begrenztes offenes Teilintervall I von (c, d) gibt, so dass $g(y) \neq 0$ in I gilt, und

$$\int_y^{y_1} \frac{dt}{g(t)}$$

für $y \in I$ als bei y_1 uneigentliches Integral konvergiert.

Beweis: Es ist nur noch die letzte Aussage über die Eindeutigkeit der Lösungskurven zu beweisen: Im Fall $f \equiv 0$ sind die Lösungen von (1.11) offensichtlich eindeutig festgelegt.

Im Fall $f \not\equiv 0$ folgt aus Satz (1.6): Die Eindeutigkeit ist nur in dem Fall fraglich, in dem nichtkonstante Lösungen existieren mit Werten, an denen g verschwindet.

Sei $y = y(x)$ eine solche Lösung. Es sei $g(y(x_0)) \neq 0$ sowie $g(y(x_1)) = 0$ und oBdA $x_0 < x_1$. Dabei wählen wir x_1 noch als den kleinsten Wert $> x_0$, für den $g(y(x_1)) = 0$ gilt, d.h. es sei $g(y(x)) \neq 0$ für $x_0 \leq x < x_1$. Wir nehmen noch $y_0 = y(x_0) < y_1 = y(x_1)$ an (sonst verläuft der Beweis analog).

Sei $\varepsilon > 0$ passend, so dass $g(y) \neq 0$ für $y_0 - \varepsilon < y < y_1$ gilt. Nach Satz (1.6) ist dann $y(x)$ als Lösung durch (x_0, y_0) eindeutig bestimmt in dem Rechteck $\{(x, y) : a < x < b, y_0 - \varepsilon < y < y_1\}$. Es gilt

$$G(y(x)) = F(x). \tag{1.13}$$

F ist in $[x_0, x_1]$ beschränkt, und daher ist wegen Gleichung (1.13) auch $G(y) = \int_{y_0}^y \frac{dt}{g(t)}$ in $[y_0, y_1]$ beschränkt. Da G in diesem Intervall eine monotone Funktion ist, existiert $\int_{y_0}^{y_1} \frac{dt}{g(t)}$ als uneigentliches Integral.

Umgekehrt gelte nun die Bedingung des Satzes, d.h. es sei $g(y_1) = 0$ und $g(y) \neq 0$ für $y_0 \leq y < y_1$. Zudem konvergiere das uneigentliche Integral $\int_{y_0}^{y_1} \frac{dt}{g(t)}$. Wegen $f(x) \neq 0$ existiert dann ein $x_1 \in (a, b)$ mit $f(x_1) \neq 0$. OBdA sei etwa $f(x_1) > 0$ und auch $g(y) > 0$ für $y_0 \leq y < y_1$.

Ist nun $g = y(x)$ Lösung in $\{(x, y) : a < x < b, y_0 < y < y_1\}$ mit $\lim_{x \rightarrow x_1^-} y(x) = y_1$, so gilt nach Satz (1.6)

$$G(y(x)) = F(x) + k \quad \text{mit } k = -F(x_1) + \int_{y_0}^{y_1} \frac{dt}{g(t)}. \quad (1.14)$$

Diese Gleichung ist in $(x_1 - \varepsilon, x_1)$ nach y auflösbar, wenn $\varepsilon > 0$ klein genug gewählt wird, so dass erstens

$$f(x) > 0 \quad \text{für } x \in (x_1 - \varepsilon, x_1)$$

gilt (dann ist nämlich $F(x)$ streng monoton wachsend in diesem Intervall) und dass zweitens $F(x_1 - \varepsilon) + k$ noch im Wertebereich von G liegt (vgl. Gleichung (1.14)).

Damit hat man eine Lösung $y(x)$ in $(x_1 - \varepsilon, x_1)$, die nicht konstant ist und für die $\lim_{x \rightarrow x_1^-} y(x) = y_1$ ist. Denn es gilt

$$\lim_{y \rightarrow y_1^-} G(y) = \lim_{x \rightarrow x_1^-} G(y(x)),$$

und bezeichnet z_1 diesen Grenzwert, so ist

$$\lim_{z \rightarrow z_1} H(z) = y_1,$$

wobei $H(z)$ die stetige Umkehrfunktion von G ist. Also gilt

$$\lim_{x \rightarrow x_1^-} y(x) = \lim_{x \rightarrow x_1^-} H(G(y(x))) = y_1.$$

Setzt man $y(x) := y_1$ für $x_1 \leq x < b$, dann ist $y(x)$ für $x_1 - \varepsilon < x < b$ Lösung der DGL (1.11), d.h. durch (x_1, y_1) gehen zwei verschiedene Lösungskurven: die konstante Funktion $\bar{y}(x) \equiv y_1$ und die oben angegebene Lösung. \square

Es ist noch zu zeigen, dass die oben angegebene Lösung an der „Nahtstelle“ x_1 differenzierbar ist. Dazu gilt allgemeiner Folgendes

Lemma 1.8. *Sind die Funktionen $y = y(x)$ in einer Umgebung von x_1 und $f(x, y)$ in einer Umgebung von $(x_1, y(x_1))$ definiert und stetig, und ist dort $y(x)$ für $x \neq x_1$ differenzierbar mit $y'(x) = f(x, y(x))$, dann ist $y(x)$ auch an der Stelle x_1 differenzierbar mit $y'(x_1) = f(x_1, y(x_1))$.*

Zum Beweis: Aufgrund der Stetigkeit von f und der Stetigkeit von $f(x, y)$ und $y(x)$ gilt

$$\lim_{x \rightarrow x_1} y'(x) = \lim_{x \rightarrow x_1} f(x, y(x)) = f(x_1, y(x_1)).$$

Nach der Regel von de l'Hospital gilt also $\lim_{x \rightarrow x_1} \frac{y(x) - y(x_1)}{x - x_1} = \lim_{x \rightarrow x_1} y'(x)$. D.h. $y'(x_1)$ existiert und ist gleich $f(x_1, y(x_1))$, und dies beendet den Beweis von (1.8). \square

Wendet man die Sätze (1.6) bzw. (1.7) auf die Beispiele am Anfang des Paragraphen an, so sieht man: Die DGL

$$y' = y^2$$

ist für jeden Anfangswert $y(x_0) = c$ eindeutig lösbar, nicht dagegen $y' = \sqrt{|y|}$: sobald man mit einer Lösungskurve die x -Achse erreicht, kann man unendlich viele verschiedene Lösungen angeben.

1.3 Elementare Lösungsmethoden

Wir betrachten zunächst die beiden speziellen DGLn mit getrennten Variablen

$$y' = f(x) \quad (1.15)$$

und

$$y' = g(y). \quad (1.16)$$

Bei (1.15) hängt das zugehörige Richtungsfeld nur von der x -Koordinate ab und ist daher invariant gegenüber Verschiebungen (=Translationen) in Richtung der y -Achse.

Die DGL (1.15) und damit auch ihre Lösungsgesamtheit ist invariant gegenüber Transformationen $x \mapsto x$, $y \mapsto y + a$ ($a \in \mathbb{R}$). Analog ist (1.16) invariant unter Transformationen $x \mapsto x + a$, $y \mapsto y$.

Manchmal lassen sich DGLn durch Koordinatentransformationen auf eine einfachere Gestalt bringen. Dabei wird man die Symmetrieverhältnisse ausnutzen und etwa bei rotationssymmetrischen Problemen z.B. Polarkoordinaten einführen. Dazu betrachten wir im Folgenden einige Beispiele.

Gegeben sei

$$y' = f(ax + by + c). \quad (1.17)$$

Die Geraden $ax + by + c = \text{const}$ sind Isoklinen dieser DGL. Wir führen folgende Koordinatentransformation durch:

$$x \mapsto x, \quad y \mapsto z := ax + by + c.$$

Damit wird

$$z' = a + b \cdot f(z),$$

und es liegt eine DGL mit getrennten Veränderlichen vor, die mit den Methoden vom Kapitel 1.2 behandelt werden kann.

Nun betrachten wir

$$y' = f\left(\frac{y}{x}\right), \quad (1.18)$$

die sog. **homogene** DGL. Wir transformieren

$$x \mapsto x, \quad y \mapsto z = \frac{y}{x}.$$

Dann ist $z' = \frac{x \cdot y' - y}{x^2} = \frac{1}{x}(f(z) - z)$. Die transformierte DGL ist wieder eine Differentialgleichung mit getrennten Veränderlichen.

Gegeben sei nun

$$F(x, y, y') = 0, \quad (1.19)$$

und F sei homogen vom Grad m in x und y , d.h. es gelte

$$F(tx, ty, y') = t^m \cdot F(x, y, y') \quad (t \in \mathbb{R}).$$

Dann existiert eine Funktion G , so dass $F(x, y, y') = x^m \cdot G(\frac{y}{x}, y')$ gilt. Lässt sich diese DGL explizit nach y' auflösen, so hat man eine homogene DGL vom Typ (1.18) vorliegen.

$$y' = f\left(\frac{ax + by + c}{dx + ey + g}\right) \quad \text{mit } D = \det \begin{pmatrix} a & b \\ d & e \end{pmatrix} \neq 0. \quad (1.20)$$

Wir transformieren

$$u = x + r, \quad v = y + s$$

mit

$$r := \frac{1}{D}(ce - bg), \quad s = \frac{1}{D}(ag - cd).$$

Dann rechnet man nach

$$\frac{au + bv}{du + ev} = \frac{ax + by + c}{dx + ey + g},$$

und es ist $\frac{dy}{dx} = \frac{dv}{du}$, so dass man als transformierte DGL erhält:

$$\frac{dv}{du} = f\left(\frac{au + bv}{du + ev}\right) = f\left(\frac{a + b\frac{v}{u}}{d + e\frac{v}{u}}\right) = F\left(\frac{v}{u}\right),$$

und wieder ist dies eine homogene DGL.

Als nächstes betrachten wir die **linearen Differentialgleichungen erster Ordnung**:

$$y' = a(x)y + b(x) \quad (1.21)$$

mit einem Intervall $I \subseteq \mathbb{R}$ und mit stetigen Funktionen $a, b : I \rightarrow \mathbb{R}$.

Ist $b(x) \equiv 0$, so nennt man die DGL (1.21) **homogen linear** (nicht zu verwechseln mit der homogenen DGL, die wir unter (1.18) betrachtet haben).

Satz 1.9. *Sind $x_0 \in I$ und $c \in \mathbb{R}$, und ist $b(x) \equiv 0$, so existiert genau eine Lösung $y : I \rightarrow \mathbb{R}$ der DGL (1.21) mit $y(x_0) = c$, und es gilt*

$$y(x) = c \cdot \exp\left(\int_{x_0}^x a(t)dt\right).$$

Beweis: (vgl. Kapitel 1.2): Die Form der Lösung ergibt sich aus dem Ansatz $\frac{y'}{y} = a(x)$ für $y \neq 0$, und die Eindeutigkeit folgt mit Satz (1.7). \square

Beispiel 1.10. Die DGL $y' = k \cdot y$ mit dem Anfangswert $y(x_0) = c$ wird eindeutig gelöst durch $y(x) = c \cdot e^{k(x-x_0)}$.

Wir gehen nun über zur inhomogenen linearen DGL

$$y' = a(x)y + b(x). \quad (1.22)$$

Hierfür gibt es ein Lösungsverfahren, das in der Literatur als **Variation der Konstanten** bekannt ist.

Ist $u : I \rightarrow \mathbb{R}$ eine nichtverschwindende Lösung der zugehörigen homogenen DGL, d.h. gilt

$$u'(x) = a(x)u(x) \quad (x \in I),$$

so machen wir den Ansatz

$$y(x) = z(x)u(x) \quad (1.23)$$

(Variation der bei u eingehenden Konstanten zu einer nicht mehr notwendig konstanten Funktion — daher „Variation der Konstanten“).

Wann ist $y(x)$ in der Form (1.23) Lösung von (1.22)? Es muss gelten

$$\begin{aligned} y'(x) &= z'(x)u(x) + z(x)u'(x) = \\ &= z'(x)u(x) + z(x)a(x)u(x) = \\ &= (z'(x) + a(x)z(x))u(x) \stackrel{!}{=} a(x)y(x) + b(x), \end{aligned}$$

und die letzte Gleichung gilt genau dann, wenn $z'(x)u(x) = b(x)$ gilt, d.h.

$$z'(x) = \frac{b(x)}{u(x)}.$$

Die letzte DGL hat die allgemeine Lösung

$$z(x) = \int_{x_0}^x \frac{b(t)}{u(t)} dt + k,$$

und man hat insgesamt den

Satz 1.11. Sind $I \subseteq \mathbb{R}$ ein Intervall, $a, b : I \rightarrow \mathbb{R}$ stetige Funktionen, dann gibt es zu beliebigem $x_0 \in I$, $c \in \mathbb{R}$ genau eine Lösung der DGL

$$y'(x) = a(x)y(x) + b(x)$$

mit der Anfangsbedingung $y(x_0) = c$. Es ist

$$y(x) = \exp\left(\int_{x_0}^x a(t)dt\right) \cdot \left(c + \int_{x_0}^x b(t) \cdot \exp\left(-\int_{x_0}^t a(u)du\right) dt\right).$$

Zum **Beweis** ist noch die Eindeutigkeit nachzuweisen. Dazu zeigen wir den

Satz 1.12. Ist \tilde{y} eine spezielle Lösung der inhomogenen DGL (1.21), so erhält man mit $y = \tilde{y} + z$ sämtliche Lösungen der inhomogenen Gleichung, wenn z sämtliche Lösungen der homogenen DGL $z'(x) = a(x)z(x)$ durchläuft.

Beweis von (1.12): Sind y und \tilde{y} Lösungen von (1.21), so ist

$$z := y - \tilde{y}$$

eine Lösung von $z'(x) = y'(x) - \tilde{y}'(x) = a(x)(y(x) - \tilde{y}(x)) = a(x)z(x)$. \square

Beweis von (1.11): Dies folgt direkt mit (1.12) aus den Ergebnissen vom Kapitel 1.2. \square

Lineare Differentialgleichungen höherer Ordnung sind i.A. etwas komplizierter zu behandeln. Eine DGL n -ter Ordnung der Form

$$y^{(n)} = a_{n-1}y^{(n-1)} + \dots + a_0y + b \tag{1.24}$$

mit stetigen Funktionen $a_i, b : I \rightarrow \mathbb{R}$ heißt **lineare DGL n -ter Ordnung**. Sie heißt **homogene** lineare DGL n -ter Ordnung, wenn $b(x) \equiv 0$ gilt. Ist $b(x) \not\equiv 0$, so heißt (1.24) **inhomogen**.

Man zeigt leicht: Man kann alle Lösungen von (1.24) angeben, wenn man alle Lösungen der zugehörigen homogenen DGL kennt sowie eine partikuläre Lösung der inhomogenen DGL.

Im Allgemeinen ist für $n \geq 2$ die homogene Gleichung nicht so einfach wie für $n = 1$ zu behandeln. Später werden wir noch Spezialfälle im Rahmen der Theorie linearer DGLn behandeln.

1.4 Existenz- und Eindeutigkeitsätze

Zugrunde liegt ein explizites Differentialgleichungssystem

$$y' = f(x, y), \quad (1.25)$$

mit $y = (y_1, \dots, y_n)$ als ein System von n Funktionen. Dabei wird $f : D \rightarrow \mathbb{R}^n$, $D \subseteq \mathbb{R}^{n+1}$ als stetige Abbildung vorausgesetzt.

Als Anfangsbedingung wird für einen vorgegebenen Punkt $(x_0, y_0) \in D$ verlangt, dass für die gesuchte Lösung $y = y(x)$ die Eigenschaft

$$y(x_0) = y_0$$

zu gelten hat.

Äquivalent zu $y' = f(x, y)$, $y(x_0) = y_0$ ist das folgende System von Integralgleichungen

$$y(x) = y_0 + \int_{x_0}^x f(t, y(t)) dt. \quad (1.26)$$

Ist $y = y(x)$ eine Lösung des Anfangswertproblems $y' = f(x, y)$ mit $y(x_0) = y_0$, so folgt (1.26) durch Integration. Umgekehrt folgt mit dem Hauptsatz der Differential- und Integralrechnung, dass jede Lösung von (1.26) auch (1.25) erfüllt.

Der **Existenzsatz von Peano** sagt aus, dass allein aus der Stetigkeit von f die Lösbarkeit jedes Anfangswertproblems folgt (bei genügend harmlosem Definitionsbereich von f). Allerdings wird man eine solche Lösung u.U. nur lokal finden können, d.h. $y = y(x)$ muss nicht für alle x definiert sein, wie das Beispiel

$$y' = y^2 \quad \text{mit der Lösung } y = -\frac{1}{x+c}$$

bereits gezeigt hat.

Es gilt folgender Satz (nach **Peano**, 1858-1932, Universität Turin, Militärakademie):

Satz 1.13 (Existenzsatz von Peano). *Sei $f(x, y)$ definiert auf $K := \{(x, y) \in \mathbb{R}^{n+1} : |x - x_0| \leq a, |y - y_0| \leq b\}$ und dort stetig. Es sei $|f(x, y)| \leq c$ für $(x, y) \in K$. Dann besitzt das Anfangswertproblem (1.25) mit $y(x_0) = y_0$ mindestens eine Lösung $y = y(x)$, und diese ist (wenigstens) für alle x mit $|x - x_0| \leq \min(a, \frac{b}{c})$ definiert.*

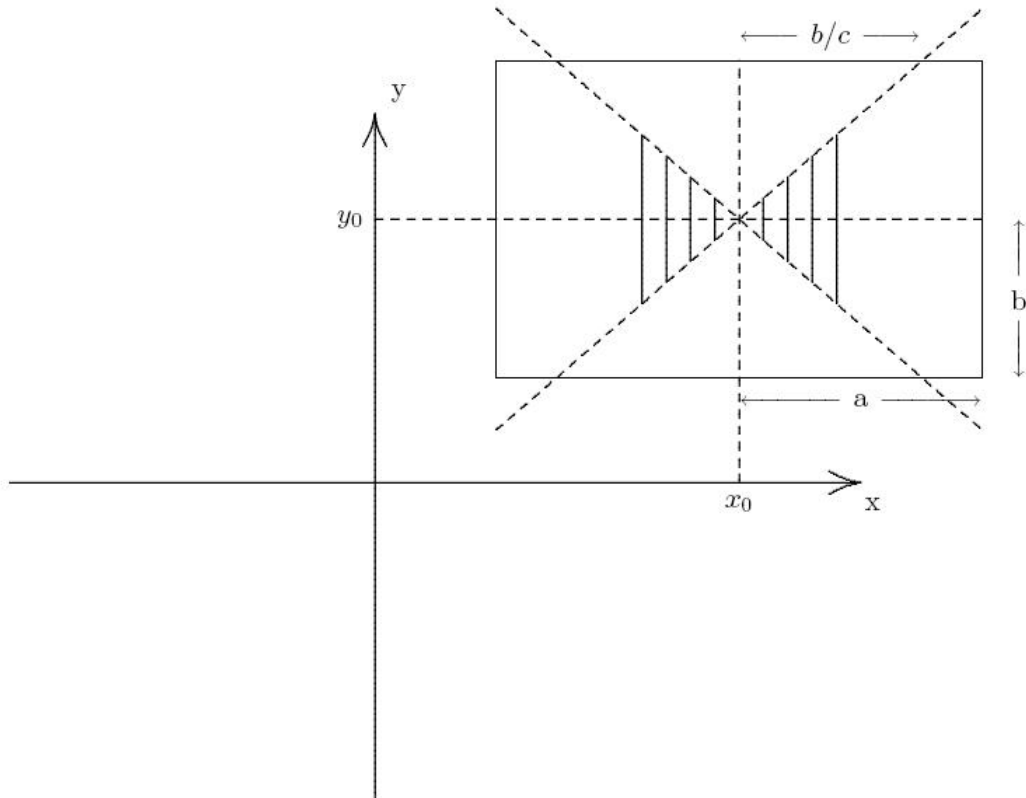


Abbildung 3: Der Doppelkegel bzw. Winkelbereich als Definitionsbereich in Satz 1.13.

Wie kommt die Bedingung $|x - x_0| \leq \min\left(a, \frac{b}{c}\right)$ zustande? Hierzu überlegt man sich Folgendes:

Wegen $|f(x, y)| \leq c$ ist die Steigung einer Lösungskurve jedenfalls $\leq c$, und damit verläuft eine Lösungskurve durch (x_0, y_0) stets in dem durch

$$|y - y_0| \leq c|x - x_0|$$

gegebenen Doppelkegel. Die Lösungskurve ist daher mindestens für $|x - x_0| < \min\left(a, \frac{b}{c}\right)$ definiert.

Die Idee des Beweises für den Existenzsatz von Peano ist ganz einfach und liegt im übrigen auch manchen numerischen Verfahren zur näherungsweise Lösung von DGLn zugrunde:

Man definiert approximierende Streckenzüge wie folgt: Mit Hilfe des durch f definierten Richtungsfeldes schreite man von (x_0, y_0) ein kleines Stück auf

der zu diesem Punkt gehörigen Geraden des Feldes fort bis zu einem Punkt (x_1, y_1) . Von (x_1, y_1) schreite man auf der hierzu gehörigen Geraden des Feldes wieder ein kleines Stück fort bis zu einem Punkt (x_2, y_2) und so weiter. Man erhält einen Streckenzug und kann hoffen, dass die bei Verkleinerung der Schrittweite entstehende Folge von Streckenzügen gegen eine Lösungskurve konvergiert. Das gilt in der Tat, der ausführliche Beweis ist z.B. in [8] oder [11] dargestellt.

Unter etwas stärkeren Voraussetzungen lässt sich nicht nur die Existenz einer Lösung, sondern auch deren Eindeutigkeit nachweisen. Es gilt:

Satz 1.14 (von Picard-Lindelöf). Sei $f(x, y)$ in $B = \{(x, y) : |x - x_0| \leq a, |y - y_0| \leq b\}$ stetig und erfülle auf B eine Lipschitzbedingung bezüglich y ; d.h. es existiere eine Konstante L , so dass

$$|f(x, y_2) - f(x, y_1)| \leq L|y_2 - y_1|$$

für alle $(x, y_i) \in B$ gilt. Dann hat das Anfangswertproblem

$$y' = f(x, y), \quad y(x_0) = y_0$$

genau eine Lösung $y(x)$, und diese ist auf $\{x : |x - x_0| < \min(a, \frac{b}{c})\}$ definiert. Dabei ist c eine Schranke für $|f(x, y)|$ auf B .

Beweis: Wir benutzen die zum Anfangswertproblem äquivalente Integralgleichung (1.26). Sei $y_0(x)$ eine beliebige in $I = \{x : |x - x_0| < \min(a, \frac{b}{c})\}$ stetige Vektorfunktion, so dass $|y_0(x) - y_0| < b$ gilt. (Man kann etwa $y_0(x) \equiv y_0$ wählen.) Dann kann man induktiv eine Folge von Funktionen y_m folgendermaßen definieren:

$$y_{m+1}(x) := y_0 + \int_{x_0}^x f(t, y_m(t)) dt \quad \text{für } x \in I, \quad m \in \mathbb{N} \cup \{0\}. \quad (1.27)$$

Die Rekursion (1.27) ist wohldefiniert. Denn ist $y_m(x)$ bereits definiert, und gilt $|y_m(x) - y_0| < b$ in I , so ist auch $f(x, y_m(x))$ in I definiert und damit auch $y_{m+1}(x)$. Weiter ist

$$|y_{m+1}(x) - y_0| \leq \left| \int_{x_0}^x |f(t, y_m(t))| dt \right| \leq c \cdot |x - x_0| < b.$$

Wegen der Lipschitz-Bedingung folgt:

$$|y_{m+1}(x) - y_m(x)| \leq \left| \int_{x_0}^x |f(t, y_m(t)) - f(t, y_{m-1}(t))| dt \right| \leq$$

$$\leq L \cdot \left| \int_{x_0}^x |y_m(t) - y_{m-1}(t)| dt \right|.$$

Ist $|y_1(x) - y_0(x)| < 2b$, so folgt per Induktion

$$|y_{m+1}(x) - y_m(x)| \leq 2b \cdot L^m \cdot \frac{|x - x_0|^m}{m!} \leq$$

$$\leq 2b \cdot \frac{(a \cdot L)^m}{m!} \quad \text{für beliebige } x \in I.$$

Die Reihe

$$\sum_{m=0}^{\infty} |y_{m+1}(x) - y_m(x)|$$

hat daher in I die konvergente Majorante $\sum_{m=0}^{\infty} 2b \frac{(ak)^m}{m!} = 2be^{ak}$. Somit konvergiert die Vektorreihe

$$\sum_{m=0}^{\infty} (y_{m+1}(x) - y_m(x))$$

gleichmäßig auf I , und wegen $\sum_{m=0}^M (y_{m+1}(x) - y_m(x)) = y_{M+1}(x) - y_0$ folgt: $(y_m(x))$ konvergiert auf I gleichmäßig gegen eine Grenzfunktion $y(x)$. Diese Grenzfunktion ist als gleichmäßiger Limes einer Folge stetiger Funktionen wieder stetig.

Die Funktion f ist stetig und auf der kompakten Menge B daher gleichmäßig stetig, daher konvergiert

$$f(x, y_m(x))$$

auch gleichmäßig gegen $f(x, y(x))$.

Aus der Analysis I ist Folgendes bekannt: Sind f_n ($n \in \mathbb{N}$) integrierbare Funktionen und konvergiert f_n gegen f gleichmäßig, so ist auch f über $[a, b]$ integrierbar mit $\int_a^x f(t) dt = \lim_{n \rightarrow \infty} \int_a^x f_n(t) dt$. In unserem Fall konvergiert daher

$$\int_{x_0}^x f(t, y_m(t)) dt$$

gleichmäßig gegen $\int_{x_0}^x f(t, y(t)) dt$. Wegen (1.27) ist daher $y(x)$ eine stetige Lösung der Integralgleichung (1.26).

Zur Eindeutigkeit: Seien $y(x)$, $z(x)$ zwei Lösungen von (1.26), die auf I nicht in jedem Punkt übereinstimmen. Wir wählen einen Punkt $\bar{x} \in I$ mit

$y(\bar{x}) \neq z(\bar{x})$. OBdA gelte $\bar{x} > x_0$, und es sei $x_1 := \inf\{x > x_0 : y(x) \neq z(x)\}$, d.h. $y(x) = z(x)$ für $x \in [x_0, x_1)$.

Wegen der Lipschitz-Bedingung gilt

$$\begin{aligned} |z(x) - y(x)| &\leq L \cdot \int_{x_0}^x |z(t) - y(t)| dt \leq \\ &\leq L \int_{x_1}^x |z(t) - y(t)| dt \leq \\ &\leq L(x - x_1) \cdot \max_{x_1 \leq t \leq x} |z(t) - y(t)| \end{aligned}$$

für $x \in [x_1, \min(a, \frac{b}{c})$.

Für $0 < q < 1$ und $x \in [x_1, \min(a, \frac{b}{c}, x_1 + \frac{q}{L})]$ gilt

$$\begin{aligned} |z(x) - y(x)| &\leq q \cdot \max_{x_1 \leq t \leq x} |z(t) - y(t)| \\ &\leq \min\left(a, \frac{b}{c}, x_1 + \frac{q}{L}\right) =: q \cdot M. \end{aligned}$$

Damit folgt

$$\max |z(x) - y(x)| \leq q \cdot M,$$

so dass $M = 0$ gilt. Dies aber ist ein Widerspruch zur Definition von x_1 , und damit ist alles gezeigt. \square

Bemerkung: Die Lipschitz-Bedingung ist keine notwendige Bedingung für die Eindeutigkeit. Man kann auch schwächere hinreichende Bedingungen als die Lipschitz-Bedingung angeben. Diese aber zeichnet sich durch ihre Einfachheit aus (vgl. hierzu auch Satz 1.7).

Frage: Wann gilt für eine Funktion (Abbildung) f eine Lipschitz-Bedingung bezüglich y ? Eine mögliche Antwort gibt der folgende

Satz 1.15. Sei $G \subseteq \mathbb{R} \times \mathbb{R}^n$ offen und $f : G \rightarrow \mathbb{R}^n$ eine bezüglich der Variablen y_i in $y = (y_1, \dots, y_n)$ stetig partiell differenzierbare Abbildung. Dann genügt f in G lokal einer Lipschitz-Bedingung.

Bemerkung: Die in Satz (1.15) verwendete Sprechweise

„ f erfüllt in G lokal eine Lipschitzbedingung“

heißt dabei: Zu jedem Punkt $P \in G$ existiert eine Umgebung $K(P, r)$, so dass f in $K(P, r)$ eine Lipschitzbedingung erfüllt mit einer Konstanten L , die von P und r abhängen darf.

Zum Beweis von Satz (1.15): Sei $P = (x_0, y_0) \in G$. Für genügend kleines $r > 0$ liegt dann

$$V := \{(x, y) \in \mathbb{R} \times \mathbb{R}^n : |x - x_0| \leq r, |y - y_0| \leq r\}$$

ganz in G . Die Menge V ist beschränkt und abgeschlossen, also kompakt. Daher existiert

$$L := \sup \left\{ \left| \frac{\partial f}{\partial y}(x, y) \right| : (x, y) \in V \right\},$$

und mit der Taylor-Formel ergibt sich

$$\begin{aligned} |f(x, y_2) - f(x, y_1)| &= \left| \int_0^1 (y_2 - y_1) \cdot \frac{\partial f}{\partial y}(x, y_1 + t \cdot (y_2 - y_1)) dt \right| \\ &\leq L \cdot |y_2 - y_1|. \end{aligned}$$

Dies beendet den Beweis von (1.15). □

Bemerkung: In einfachen Fällen kann man mit dem Picard-Lindelöfschen Verfahren die Lösungen einer DGL explizit berechnen.

Beispiel 1.16. Gegeben sei auf \mathbb{R}^2 die Differentialgleichung

$$y' = 2xy$$

mit der Anfangsbedingung $y(0) = c$. Dann sieht die aus (1.27) bekannte Iterationsvorschrift wie folgt aus:

$$y_{m+1}(x) = c + 2 \cdot \int_0^x t y_m(t) dt.$$

Wir starten mit der konstanten Funktion $y_0(x) \equiv c$ und erhalten

$$y_1(x) = c + 2c \int_0^x t dt = c(1 + x^2),$$

$$y_2(x) = c + 2c \int_0^x t(1 + t^2) dt = c \left(1 + x^2 + \frac{x^4}{2} \right)$$

und per Induktion

$$y_m(x) = c \left(1 + x^2 + \frac{x^4}{2} + \frac{x^6}{3!} + \dots + \frac{x^{2m}}{m!} \right).$$

Also ist der Limes

$$y(x) = \lim_{m \rightarrow \infty} y_m(x) = c \cdot \sum_{k=0}^{\infty} \frac{x^{2k}}{k!} = c \cdot e^{x^2}$$

eine Lösung der in (1.16) angegebenen DGL mit $y(0) = c$.

Fragen:

1. Was ist der maximale Definitionsbereich einer durch den Punkt (x_0, y_0) gehenden Lösung?
2. Wann ist eine Lösung in ganz G eindeutig bestimmt?

Zu 2. Die Funktion f genüge auf G lokal einer Lipschitzbedingung. Dann lässt sich 2. positiv beantworten:

Satz 1.17. Sei $f(x, y)$ stetig in dem offenen Bereich $G \subseteq \mathbb{R}^{n+1}$. f erfülle in G lokal eine Lipschitzbedingung. Sind $(x_0, y_0) \in G$, $I = [x_0, x_0 + a]$ bzw. $[x_0, x_0 + a)$ und y_1, y_2 über I definierte Lösungen von $y' = f(x, y)$ mit $y_0 = y_i(x_0)$ für $i = 1, 2$, so gilt

$$y_1 \equiv y_2 \quad \text{auf } I.$$

Beweis: Wir nehmen an, $N := \{x \in I : y_1(x) \neq y_2(x)\}$ sei nicht leer. Dann existiert $x_1 := \inf N$ und liegt in I . Wegen der Stetigkeit der Lösungen gilt

$$y_1 = y_1(x_1) = \lim_{x \rightarrow x_1^-} y_1(x) = \lim_{x \rightarrow x_1^-} y_2(x) = y_2(x_1).$$

Zu (x_1, y_1) existiert eine Umgebung $K = K((x_1, y_1), r)$, in der f eine Lipschitzbedingung erfüllt. Die Lösung durch (x_1, y_1) ist also in einer Umgebung dieses Punktes eindeutig bestimmt. In einem kleinen Intervall um x_1 gilt deshalb $y_1 \equiv y_2$ im Widerspruch zur Konstruktion von x_1 . \square

Zur Frage 1. Mit den bewiesenen Sätzen findet man für stetiges f auf $G \subseteq \mathbb{R}^{n+1}$, falls f lokal einer Lipschitzbedingung genügt, stets eine Lösung in einem gewissen Intervall I um x_0 . Wie weit lässt sich y über das Intervall I hinaus fortsetzen?

Dabei heißt $\tilde{y}(x)$ mit einem zusammenhängenden Definitionsbereich $D_{\tilde{y}}$ eine **echte** Fortsetzung von $y(x)$ mit Definitionsbereich D_y , wenn D_y echte Teilmenge von $D_{\tilde{y}}$ ist und $y(x) \equiv \tilde{y}(x)$ auf D_y gilt.

Satz 1.18. Sei $f(x, y)$ in dem offenen Bereich $G \subseteq \mathbb{R}^{n+1}$ definiert und stetig. Sei $(x_0, y_0) \in D$, und f genüge in G lokal einer Lipschitzbedingung. Ist b maximal mit der Eigenschaft, dass es ein über $I = [x_0, x_0 + b)$ definierte Lösung $y = y(x)$ von $y' = f(x, y)$ mit $y(x_0) = y_0$ gibt, dann ist $b = \infty$ oder $b < \infty$ und

$$\overline{\{(x, y(x)) : x \in I\}} \cap \{(x, y) : x = x_0 + b\} \cap G = \emptyset,$$

d.h. y lässt sich „nach rechts“ nicht echt fortsetzen.

Bemerkung: Die gleichen Überlegungen kann man für die Intervalle der Form $[x_0 - a, x_0]$ anstellen. Man findet dann ein maximales b_1 mit $0 < b_1 \leq +\infty$, so dass über $I^* = (x_0 - b_1, x_0]$ eine Lösung $y(x)$ von $y' = f(x, y)$ mit $y(x_0) = y_0$ existiert. (Für $b_1 = \infty$ ist dann $I^* = (-\infty, x_0]$ zu setzen.)

Zum Beweis von (1.18): Wir nehmen an, dass $b < \infty$ gilt, aber der fragliche Durchschnitt nicht leer ist. Sei

$$(x_1, y_1) \in \overline{\{(x, y(x)) : x \in I\}} \cap \{(x, y) : x = x_0 + b\} \cap G.$$

Dann ist $x_1 = x_0 + b$ und $(x_1, y_1) \in G$. G ist offen, also existiert ein $\varepsilon > 0$ mit $\overline{K}((x_1, y_1), \varepsilon) \subseteq G$.

Man setze

$$c := \sup |f(\overline{K}((x_1, y_1), \varepsilon))| < \infty$$

und

$$\delta := \frac{1}{2} \min \left(\varepsilon, \frac{\varepsilon}{c} \right) > 0.$$

Der Punkt (x_1, y_1) gehört zur abgeschlossenen Hülle des Graphen der Lösung $y(x)$. Also existiert $(x_2, y_2) \in \{(x, y(x)) : x \in I\}$ mit $x_0 \leq x_2 < x_0 + b$ und $|(x_1, y_1) - (x_2, y_2)| < \delta$. Wir setzen jetzt wie in Satz (1.14)

$$a = \delta \quad \text{und} \quad b = \frac{\varepsilon}{2}$$

sowie $B := \{(x, y) : x_2 \leq x \leq x_2 + a, |y - y_2| \leq b\}$. Dann ist $B \subseteq \overline{K}((x_1, y_1), \varepsilon)$.

Ferner ist $c \cdot a = \delta \cdot c \leq \frac{\varepsilon}{2c} \cdot c = \frac{\varepsilon}{2} = b$, d.h. $a \leq \frac{b}{c}$. Nach Satz (1.14) existiert daher auf $I_1 := [x_2, x_2 + a]$ eine Lösung \tilde{y} von $\tilde{y}' = f(x, \tilde{y})$ mit $\tilde{y}(x_2) = y_2$, deren Graph ganz in B liegt.

Setzt man

$$\bar{y}(x) := \begin{cases} y(x) & \text{für } x \in I \\ \tilde{y}(x) & \text{für } x \in I_2, \end{cases}$$

so ist dies eine wohldefinierte Funktion, da $y(x_2) = \tilde{y}(x_2) = y_2$ gilt und damit nach Satz (1.17) y und \tilde{y} auf $I \cap I_2$ übereinstimmen.

\tilde{y} ist Lösung des gegebenen Anfangswertproblems, und es ist $x_2 + a > x_1 - \delta + a = x_1 = x_0 + b$ - ein Widerspruch zur Wahl von b . \square

Ist wie oben $G \subseteq \mathbb{R}^{n+1}$ offen, f in G stetig und erfüllt f lokal eine Lipschitzbedingung in G , so kann man zu jedem Anfangswert (x_0, y_0) das maximale Intervall $I(x_0, y_0)$ bilden, über dem eine Lösung $y = y(x; x_0, y_0)$ der Differentialgleichung $y' = f(x, y)$ definiert ist, die $y(x_0; x_0, y_0) = y_0$ erfüllt.

Die Menge $B := \{(x, x_0, y_0) \in \mathbb{R}^{n+2} : (x_0, y_0) \in G, x \in I(x_0, y_0)\}$ ist Teilmenge des \mathbb{R}^{n+2} , und $y(x, x_0, y_0)$ kann als Abbildung von B nach \mathbb{R}^n aufgefaßt werden. $y(x; x_0, y_0)$ heißt **allgemeine Lösung** der Differentialgleichung $y' = f(x, y)$.

Es liegt nahe, danach zu fragen, wie die Lösungen von den Anfangsbedingungen abhängen. Dazu kann man allgemeiner als bisher Anfangswertprobleme betrachten, bei denen die rechte Seite der DGL noch von einem oder mehreren Parametern abhängt:

$$y' = f(x, y, t), \quad y(x_0) = y_0, \quad (1.28)$$

wobei t ein einzelner Parameter oder ein Parametervektor $t = (t_1, \dots, t_m)$ ist.

Variiert man t und den Startwert y_0 , so hängt die Lösung $y = y(x; x_0, y_0, t)$ von t und y_0 ab. Unter welchen Voraussetzungen hängt die Lösung stetig, differenzierbar usf. von t bzw. y_0 ab?

Hier werden nur die Ergebnisse angegeben (vgl. [8] oder [12]). Man kann z.B. zeigen:

Satz 1.19. *Sei f stetig in dem Bereich*

$$B = \{(x, y, t) : |x - x_0| \leq a, |y - y_0| \leq b, |t - t_0| \leq c\} \subseteq \mathbb{R}^{1+n+m}$$

und genüge in B einer Lipschitzbedingung bezgl. y :

$$|f(x, y, t) - f(x, \bar{y}, t)| \leq L|y - \bar{y}|.$$

Dann ist die Lösung $y = y(x; x_0, y_0, t)$ des Anfangswertproblems stetig als Funktion von y_0 und t .

Satz 1.20. $f = f(x, y, t)$ sei stetig auf der offenen Teilmenge $G \subseteq \mathbb{R}^{1+n+m}$ und stetig nach y und t differenzierbar. Die Lösung $y = y(x; x_0, y_0, t)$ des Anfangswertproblems

$$y' = f(x, y, t), \quad y(x_0) = y(x_0; x_0, y_0, t) = y_0$$

ist partiell nach t und y_0 differenzierbar.

Bemerkung: Man kann zeigen, dass sich auch jede k -malige Differenzierbarkeit von f nach t und y auf die allgemeine Lösung überträgt.

Bemerkung: Zu quantitativen Abschätzungen für Näherungslösungen vgl. auch [8], Kapitel VI.

1.5 Lösung durch Potenzreihenansatz

Ist die Funktion $f(x, y)$ um einen Punkt $(x_0, y_0) \in \mathbb{R}^2$ in eine Potenzreihe nach x und y entwickelbar, so wird man erwarten, dass auch die durch (x_0, y_0) laufende Lösung der DGL $y' = f(x, y)$ um den Punkt x_0 in eine Potenzreihe entwickelbar ist und die Koeffizienten der Potenzreihe einer Lösung durch Koeffizientenvergleich gefunden werden können.

Hierzu vorweg zwei Beispiele:

$$y' = y, \quad y(0) = 1. \quad (1.29)$$

Wir machen den Ansatz $y(x) = \sum_{n=0}^{\infty} a_n x^n$. Falls diese Potenzreihe konvergiert, darf sie gliedweise differenziert werden:

$$y'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1} = a_1 \cdot 1 + 2a_2 x + \dots$$

Wegen $y(0) = 1$ folgt zunächst $a_0 = 1$. Wegen $y'(x) = y(x)$ folgt

$$\sum_{n=1}^{\infty} n a_n x^{n-1} = \sum_{n=0}^{\infty} a_n x^n.$$

Hieraus erhält man durch Koeffizientenvergleich

$$(n+1)a_{n+1} = a_n \quad \text{für } n = 0, 1, 2, \dots$$

und daher $a_{n+1} = \frac{a_n}{n+1}$, $a_0 = 1$.

Induktion zeigt

$$a_n = \frac{1}{n!},$$

und wir haben

$$y(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} = e^x$$

als Lösung der DGL (1.29) mit $y(0) = 1$.

Nun zum zweiten

Beispiel 1.21. Gegeben sei

$$y' = 2xy, \quad y(0) = 1$$

(vgl. auch Beispiel (1.16)). Der Potenzreihenansatz $y = y(x) = \sum_{n=0}^{\infty} a_n x^n$ führt wegen $y(0) = 1$ zu

$$a_0 = 1$$

und

$$y' = \sum_{n=1}^{\infty} n a_n x^{n-1} = \sum_{n=0}^{\infty} (n+1) a_{n+1} \cdot x^n,$$

$$2xy = 2 \sum_{n=0}^{\infty} a_n x^{n+1} = \sum_{n=1}^{\infty} 2a_{n-1} x^n.$$

Koeffizientenvergleich liefert

$$a_0 = 1, \quad a_1 = 0$$

und

$$(n+1)a_{n+1} = 2 \cdot a_{n-1} \quad \text{für } n \in \mathbb{N}. \quad (1.30)$$

Aus der Rekursion (1.30) ergibt sich sofort: Für alle $k \in \mathbb{N}$ gilt

$$\begin{cases} a_{2k+2} = \frac{1}{k+1} a_{2k}, \\ a_{2k+1} = 0 \end{cases}$$

Als Ergebnis erhält man hier die Potenzreihe

$$\sum_{k=0}^{\infty} \frac{1}{k!} x^{2k} = e^{x^2}.$$

Dies ist, wie schon im Beispiel (1.16) gesehen, eine Lösung des gegebenen Anfangswertproblems (1.21).

Allgemeiner kann man zeigen:

Satz 1.22. *Die rechte Seite des DGL-Systems*

$$y' = \sum_{\lambda, \mu} a_{\lambda\mu} x^\lambda y^\mu$$

konvergiere in $B = \{(x, y) : |x| \leq r, |y| \leq t\} \subseteq \mathbb{R}^{n+1}$. Dann existieren ein $s > 0$ und eindeutig bestimmte Koeffizienten $c_\nu = (c_{1\nu}, \dots, c_{n\nu})$, so dass

$$y(x) = \sum_{\nu=1}^{\infty} c_\nu x^\nu$$

für $|x| \leq s$ konvergiert und Lösung des DGL-Systems mit der Anfangsbedingung $y(0) = 0$ ist.

Bemerkung zur Potenzreihe $\sum a_{\lambda\mu} x^\lambda y^\mu$:

Hier werden wieder Multiindizes μ verwendet: y^μ steht für $y_1^{\mu_1} \cdot \dots \cdot y_n^{\mu_n}$, und die $a_{\lambda\mu}$ sind Vektoren $a_{\lambda\mu} = (a_{1\lambda\mu}, \dots, a_{n\lambda\mu})$.

Der prinzipiell nicht schwierige **Beweis** geht im Ansatz von einer konvergen-
ten Potenzreihe $y(x) = \sum c_\nu x^\nu$ aus, die zu bestimmten Koeffizienten c_ν
ergeben sich durch Koeffizientenvergleich aus der Gleichung

$$y'(x) = \sum_{\lambda,\mu} a_{\lambda\mu} x^\lambda y^\mu$$

rekursiv. Man zeigt, dass die so erhaltene Potenzreihe $\sum c_\nu x^\nu$ tatsächlich
konvergiert und eine Lösung des gegebenen DGL-Systems liefert. Zu den
Einzelheiten des Beweises vergleiche z.B. [12] oder [8].

Korollar 1.23. *Lässt sich $f(x, y, y', \dots, y^{(n-1)})$ an der Stelle $(x_0, y_0, y'_0, \dots, y_0^{(n-1)})$
in eine konvergente Potenzreihe entwickeln, dann hat auch die Lösung des
Anfangswertproblems*

$$\left\{ \begin{array}{l} y^{(n)}(x) = f(x, y, y', \dots, y^{(n-1)}) \\ y(x_0) = y_0 \\ y'(x_0) = y'_0 \\ \vdots \\ y^{(n-1)}(x_0) = y_0^{(n-1)} \end{array} \right.$$

eine in einem Intervall um x_0 konvergente Potenzreihenentwicklung $y(x) =$
 $\sum_{\nu=0}^{\infty} (x - x_0)^\nu$. Dabei ist $c_0 = y_0$, $c_1 = y'_0$ usw. bis $c_{n-1} = y_0^{(n-1)}$.
Die Koeffizienten c_ν mit $\nu \geq n$ ergeben sich rekursiv durch formales Einsetzen
in die DGL und Koeffizientenvergleich.

1.6 Lineare Differentialgleichungen

Wir haben bereits früher gesehen: Eine explizite lineare DGL n -ter Ordnung

$$y^{(n)} = a_0(x)y + a_1(x)y' + \dots + a_{n-1}(x)y^{(n-1)} + b(x)$$

ist äquivalent zu einem linearen DGL-System erster Ordnung

$$\begin{cases} y'_0 = y_1 \\ y'_1 = y_2 \\ \vdots \\ y'_{n-2} = y_{n-1} \\ y'_{n-1} = a_0(x)y_0 + \dots + a_{n-1}(x)y_{n-1} + b(x). \end{cases}$$

Wir betrachten hier ein Gebiet G der Gestalt $G = I \times \mathbb{R}^n$. Dabei sei I ein (nicht notwendig endliches) Intervall. Auf diesem Gebiet betrachten wir ein lineares DGL-System erster Ordnung

$$y'_i = \sum_{k=1}^n a_{ik}(x)y_k + b_i(x) \quad (i = 1, \dots, n) \quad (1.31)$$

mit stetigen Funktionen $a_{ik} : I \rightarrow \mathbb{R}$ für $i, k = 1, \dots, n$ sowie $b_i : I \rightarrow \mathbb{R}$ für $i = 1, \dots, n$. In kürzerer Schreibweise

$$y'(x) = A(x) \cdot y(x) + b(x), \quad (1.32)$$

wobei

$$y(x) = \begin{pmatrix} y_1(x) \\ \vdots \\ y_n(x) \end{pmatrix} \quad \text{und} \quad b(x) = \begin{pmatrix} b_1(x) \\ \vdots \\ b_n(x) \end{pmatrix}$$

Spaltenvektoren sind und $A : I \rightarrow M(n \times n, \mathbb{R})$ die durch $A(x) := (a_{ik}(x))$ definierte Abbildung bezeichnet.

Man bezeichnet (1.32) als ein **inhomogenes lineares Differentialgleichungssystem**. Ist $b(x) \equiv 0$ der Nullvektor, so heißt (1.32) ein **homogenes lineares DGL-System**, bzw.

$$y'(x) = A(x)y(x) \quad (1.33)$$

heißt das (1.32) zugeordnete **homogene lineare DGL-System**.

Erste Bemerkung: Offenbar ist jede Funktion $f(x, y) := A(x)y + b(x)$ stetig nach y differenzierbar, d.h. diese Abbildung genügt lokal einer Lipschitzbedingung.

Es gilt darüber hinaus folgende gegenüber dem allgemeinen Existenzsatz verschärfte Aussage:

Satz 1.24. *Es sei I ein offenes Intervall. A und b seien stetig in I . Ist $(x_0, y_0) \in I \times \mathbb{R}^n$, so ist die durch (x_0, y_0) laufende (eindeutig bestimmte) Lösung von (1.33) $y : I \rightarrow \mathbb{R}^n$ mit $y(x_0) = y_0$ auf dem ganzen Intervall I definiert.*

Bemerkung: Bei einer allgemeinen Differentialgleichung $y' = f(x, y)$ kann i.A. eine Lösung nicht über dem ganzen Intervall $I \subseteq \mathbb{R}$ gefunden werden, selbst wenn f über ganz $I \times \mathbb{R}^n$ definiert ist und dort lokal einer Lipschitz-Bedingung genügt. Das zeigt etwa das Beispiel $y' = y^2$.

Zum Beweis von Satz (1.24): Man setzt $f(x, y) := A(x)y + b(x)$. Ist $J \subseteq I$ ein kompaktes Teilintervall, so genügt f in $J \times \mathbb{R}^n$ einer globalen Lipschitzbedingung. Denn A ist stetig, und daher ist

$$L := \sup\{|A(x)| : x \in J\} < \infty.$$

Dabei ist

$$|A| = |a_{ik}(x)| := n \cdot \sup_{i,k} |a_{ik}(x)|.$$

Dann gilt $|c \cdot A| = |c| \cdot |A|$ für alle $c \in \mathbb{R}$ (oder \mathbb{C}), und für $B = (b_{ik})$ gilt

$$|A + B| = n \cdot \sup_{i,k} |a_{ik} + b_{ik}| \leq n \cdot \sup_{i,k} |a_{ik}| + n \cdot \sup_{i,k} |b_{ik}| = |A| + |B|$$

und

$$\begin{aligned} |AB| &= n \cdot \sup_{i,k} \left| \sum_{\nu=1}^n a_{i\nu} b_{\nu k} \right| \leq \\ &\leq n^2 \sup_{i,k,\nu} |a_{i\nu}| \cdot |b_{\nu k}| \leq n \cdot \sup_{i,k} |a_{ik}| \cdot n \cdot \sup_{i,k} |b_{ik}| = |A| \cdot |B|. \end{aligned}$$

Insbesondere folgt mit vollständiger Induktion

$$|A^m| \leq |A|^m \quad \text{für } m \in \mathbb{N}.$$

Achtung: Im Allgemeinen gilt $|A \cdot B| \neq |B \cdot A|$. Ist beispielsweise $A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ und $B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, so gilt zwar $|AB| = 2$, aber $|BA| = 4$.

Für $x \in J$, $y, \tilde{y} \in \mathbb{R}^n$ folgt dann

$$\begin{aligned} |f(x, y) - f(x, \tilde{y})| &= |A(x)(y - \tilde{y})| = \\ &= \sqrt{\sum_{i=1}^n \left(\sum_{k=1}^n a_{ik}(x)(y_k - \tilde{y}_k) \right)^2}. \end{aligned}$$

Nach der Cauchy-Schwarzschen Ungleichung ist

$$\left(\sum_{k=1}^n a_{ik}(x)(y_k - \tilde{y}_k) \right)^2 \leq \left(\sum_{k=1}^n a_{ik}^2(x) \right) \cdot \left(\sum_{k=1}^n (y_k - \tilde{y}_k)^2 \right).$$

Daher gilt

$$\begin{aligned} \sqrt{\sum_{i=1}^n \left(\sum_{k=1}^n a_{ik}(x)(y_k - \tilde{y}_k) \right)^2} &\leq \sqrt{\sum_{i,k=1}^n a_{ik}^2(x)} \cdot |y - \tilde{y}| \leq \\ &\leq \sqrt{n^2 \cdot \sup |a_{ik}|^2} \cdot |y - \tilde{y}| = n \cdot \sup |a_{ik}(x)| \cdot |y - \tilde{y}| = \\ &= |A(x)| \cdot |y - \tilde{y}| \leq L \cdot |y - \tilde{y}|. \end{aligned}$$

Damit ist die Eindeutigkeit der Lösung gegeben (vgl. Kapitel 1.4, Satz (1.17)). Das Picard-Lindelöfsche Iterationsverfahren kann jetzt wie im Kapitel 1.4, Satz (1.14) durchgeführt werden, um die Existenz einer Lösung auf ganz J sicherzustellen. Dies beendet den Beweis von Satz (1.24). \square

Für die zu (1.32) zugeordnete homogene Gleichung gilt der

Satz 1.25. *Die Lösungen y der homogenen Gleichung $y' = A(x)y$ bilden einen Vektorraum über \mathbb{R} (oder \mathbb{C} , wenn man komplexwertige Lösungen zulässt).*

Beweis: trivial.

Weiter gilt der bemerkenswerte

Satz 1.26. (i) *Seien $y_1(x), \dots, y_l(x)$ Lösungen von $y'(x) = A(x)y(x)$, und sei $x_0 \in I$. Die Vektoren $y_1(x_0), \dots, y_l(x_0)$ sind genau dann linear abhängig, wenn die Vektorfunktionen $y_1(x), \dots, y_l(x)$ über I linear abhängig sind.*

(ii) *Die Dimension des Lösungsraumes von $y'(x) = A(x)y(x)$ ist n .*

Beweis: (i) In der Gleichung $\sum_{i=1}^l c_i y_i(x_0) = 0$ seien nicht alle $c_i = 0$. Dann ist $\sum c_i y_i$ eine Lösung des homogenen DGL-Systems mit Wert 0 an der Stelle x_0 . Wegen des Eindeutigkeitsatzes ist dann $\sum c_i y_i \equiv 0$. Insbesondere folgt, dass die Dimension des Lösungsraumes $\leq n$ ist.

(ii) Sei $x_0 \in I$ beliebig. Sei $e_i = (0, \dots, 0, 1, 0, \dots, 0)$ der i -te Einheitsvektor des \mathbb{R}^n mit der 1 an der i -ten Stelle. Weiter sei y_i eine Lösung mit $y_i(x_0) = e_i$.

Ist $\sum_i c_i y_i \equiv 0$, so folgt $0 = \sum c_i y_i(x_0) = \sum c_i e_i$. Dann ist $c_1 = \dots = c_n = 0$. Der Lösungsraum hat also eine Dimension $\geq n$, und mit (i) ergibt sich, dass der Lösungsraum tatsächlich die genaue Dimension n hat. \square

Definition 1.27. *Eine Basis*

$$\varphi_1(x), \dots, \varphi_n(x)$$

des Vektorraums von Lösungen von $y'(x) = A(x)y(x)$ heißt **Fundamentalsystem** (oder **Hauptsystem**) von Lösungen. Die n Vektorfunktionen eines Fundamentalsystems von

$$y'(x) = A(x)y(x)$$

kann man zu einer n -reihigen quadratischen Matrix $\Phi = (\varphi_1, \dots, \varphi_n)$ zusammenfassen. Diese Matrix heißt **Fundamentalmatrix** von (1.33). Ist $\Phi(x)$ Fundamentalmatrix, so ist $\det \Phi(x)$ auf ganz I von Null verschieden.

Wäre nämlich $\det \Phi(x_0) = 0$ für ein $x_0 \in I$, so wären $\varphi_1(x_0), \dots, \varphi_n(x_0)$ linear abhängig. In diesem Fall wären auch $\varphi_1, \dots, \varphi_n$ linear abhängig und damit kein Fundamentalsystem.

Satz 1.28. *Ist Φ eine Fundamentalmatrix von (1.33), so ist für jede konstante $n \times n$ -Matrix C mit $\det C \neq 0$ die Matrix $\Psi := \Phi \cdot C$ eine Fundamentalmatrix von (1.33) und jede Fundamentalmatrix von (1.33) lässt sich in dieser Gestalt schreiben.*

Beweis: (i) Sei $\Phi = (\varphi_1, \dots, \varphi_n)$ und $\Psi = (\psi_1, \dots, \psi_n)$ sowie $C = (c_{ik})$. Genau dann gilt $\Psi = \Phi \cdot C$, wenn $\psi_k = \sum_{\mu=1}^n c_{\mu k} \varphi_{\mu}$ für $k = 1, \dots, n$ ist. Mit φ_j als Lösung für $j = 1, \dots, n$ ist daher auch ψ_k Lösung von (1.33). Wegen $\det \Phi \neq 0$ und $\det C \neq 0$ folgt auch $\det \Psi(x) \neq 0$ für alle $x \in I$. Damit ist auch Ψ Fundamentalmatrix von (1.33).

(ii) Ist umgekehrt Ψ Fundamentalmatrix und sind $\varphi_1, \dots, \varphi_n$ Lösungen von (1.33), so ist jedes φ_k Linearkombination der ψ_j :

$$\varphi_k = \sum_{j=1}^n \tilde{c}_{jk} \cdot \psi_j, \quad \tilde{c}_{ik} \in \mathbb{R} \text{ (oder } \mathbb{C}\text{)}.$$

Dies ist gleichbedeutend mit

$$\Phi = \Psi \cdot \tilde{C}$$

mit $\tilde{C} = (c_{jk})$. Ist Φ Fundamentalmatrix, so sind Φ und Ψ nichtsinguläre Matrizen, und daher ist auch $\det \tilde{C} \neq 0$. Mit $C := \tilde{C}^{-1}$ gilt dann

$$\Psi = \Phi \cdot C.$$

□

Definition 1.29. *Die Determinante*

$$\det \Phi(x) = \det(\varphi_1(x), \dots, \varphi_n(x))$$

heißt **Wronski-Determinante** des Systems $\varphi_1, \dots, \varphi_n$ von Lösungen des Systems (1.33).

Bemerkung:

- a) Es gilt entweder $\det \Phi(x) \equiv 0$ auf I , oder $\det \Phi(x) \neq 0$ für alle $x \in I$.
- b) Die Wronski-Determinanten zweier Fundamentalsysteme $\Phi(x)$ und $\Psi(x)$ unterscheiden sich nur um einen konstanten Faktor $c \neq 0$:

$$\det \Phi(x) = c \cdot \det \Psi(x)$$

(dabei ist $c = \det C$, und C wie in Satz 1.28).

Bezeichnen wir mit $\text{Sp}(A)$ die **Spur** der Matrix $A = (a_{ij})$, d.h.

$$\text{Sp}(A) = \sum_{i=1}^n a_{ii},$$

so gilt der

Satz 1.30. *Die Wronski-Determinante einer Fundamentalmatrix $\Phi(x)$ von (1.33) genügt der linearen DGL*

$$\frac{d}{dx} \det \Phi(x) = (\text{Sp}A(x)) \cdot \det \Phi(x),$$

es gilt daher

$$\det \Phi(x) = c \cdot \exp \left(\int_{x_0}^x \text{Sp}A(t) dt \right)$$

mit einer Konstanten $c \neq 0$.

Beweis: Ist $\Phi(x) = (\varphi_{ik}(x))$ und

$$D = D(x) = \det \Phi(x) = \sum_{\pi \in S_n} \text{sgn}(\pi) \cdot \varphi_{1\pi(1)}(x) \cdots \varphi_{n\pi(n)}(x),$$

so berechnet sich die Ableitung von $D(x)$ folgendermaßen nach der Leibniz-Regel:

D' ist die Summe der Determinanten, die aus D dadurch entstehen, dass man die Elemente einer Zeile differenziert.

Nun ist

$$\varphi'_{ik}(x) = \sum_{j=1}^n a_{ij}(x)\varphi_{jk}(x),$$

und es folgt daher für beliebige $x \in I$

$$\begin{aligned} \frac{d}{dx}(\det \Phi(x)) &= \sum_{i=1}^n \det \begin{pmatrix} \varphi_{11}(x) & \dots & \varphi_{1n}(x) \\ \vdots & \dots & \vdots \\ \varphi'_{i1}(x) & \dots & \varphi'_{in}(x) \\ \vdots & \dots & \vdots \\ \varphi_{n1}(x) & \dots & \varphi_{nn}(x) \end{pmatrix} = \\ &= \sum_{i=1}^n \sum_{k=1}^n a_{ik}(x) \cdot \det \begin{pmatrix} \varphi_{11}(x) & \dots & \varphi_{1n}(x) \\ \vdots & \dots & \vdots \\ \varphi_{k1}(x) & \text{[i-te Zeile]} & \varphi_{kn}(x) \\ \vdots & \dots & \vdots \\ \varphi_{n1}(x) & \dots & \varphi_{nn}(x) \end{pmatrix} = \\ &= \sum_{i=1}^n a_{ii}(x) \cdot \det \Phi(x) = \text{Sp}A(x) \cdot \det \Phi(x). \end{aligned}$$

Bei dem vorletzten Schritt wurde benutzt, dass die Determinante verschwindet, wenn die Matrix zwei gleiche Zeilen enthält. Damit ist Satz (1.30) bewiesen. \square

Anwendung auf den Fall der homogenen linearen DGL n -ter Ordnung:

Gegeben sei die Gleichung

$$y^{(n)} + a_1(x) \cdot y^{(n-1)} + \dots + a_{n-1}(x)y' + a_n(x)y = 0$$

mit stetigen Abbildungen $a_i : I \rightarrow \mathbb{R}$ (bzw. \mathbb{C}).

Die Matrix $A(x)$ des zugeordneten linearen DGL-Systems erster Ordnung hat die Gestalt

$$A(x) = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \dots & 1 \\ -a_n & -a_{n-1} & \dots & -a_2 & -a_1 \end{pmatrix}.$$

Die m vielen Lösungsfunktionen $y_1(x), \dots, y_m(x)$ sind genau dann linear unabhängig, wenn die Vektoren

$$\begin{pmatrix} y_i(x_0) \\ y_i'(x_0) \\ \vdots \\ y_i^{(n-1)}(x_0) \end{pmatrix}$$

für $i = 1, \dots, m$ bei beliebigem festen $x_0 \in I$ linear unabhängig sind.

Die Wronski-Determinante des zugeordneten linearen DGL-Systems ist

$$\det Y(x) = \det \begin{pmatrix} y_1(x) & \dots & y_n(x) \\ \vdots & & \vdots \\ y_1^{n-1}(x) & \dots & y_n^{n-1}(x) \end{pmatrix}.$$

Die allgemeinen Ergebnisse zeigen: Entweder gilt $\det Y(x) \equiv 0$ auf ganz I , oder es ist $\det Y(x) \neq 0$ für alle $x \in I$.

Die Eigenschaft $\det Y(x) \neq 0$ mit einem gegebenen $x \in I$ gilt genau dann, wenn die Lösungen

$$y_1(x), \dots, y_n(x)$$

linear unabhängig sind. Wegen $\text{Sp}A(x) = -a_1(x)$ folgt mit Satz (1.30)

$$\det Y(x) = c \cdot \exp \left(- \int_{x_0}^x a_1(t) dt \right)$$

mit einer Konstanten c .

Nun zu den inhomogenen linearen Gleichungssystemen

$$y' = A(x)y + b(x). \tag{1.34}$$

Die Lösbarkeit von (1.34) ist dann gegeben, wenn man für das zugehörige homogene System ein Fundamentalsystem von Lösungen kennt:

Satz 1.31. *Ist $\Phi(x)$ eine Fundamentalmatrix des homogenen Systems $y' = A(x)y$, so wird die allgemeine Lösung von (1.34) gegeben durch*

$$y(x) = \Phi(x) \left(c + \int_{x_0}^x (\Phi(t))^{-1} \cdot b(t) dt \right).$$

Dabei ist c ein konstanter Vektor, $(\Phi(t))^{-1}$ ist die zu $\Phi(t)$ inverse Matrix.

Beweis: Der Beweis erfolgt ähnlich dem eindimensionalen Fall durch die Methode der Variation der Konstanten: Wir machen den Ansatz

$$y(x) = \Phi(x) \cdot z(x)$$

mit einem Spaltenvektor $z(x)$. Wegen

$$\Phi'(x) = A(x)\Phi(x)$$

und

$$\begin{aligned} y'(x) &= \Phi'(x) \cdot z(x) + \Phi(x) \cdot z'(x) = \\ &= A(x)\Phi(x)z(x) + \Phi(x) \cdot z'(x) = A(x)y(x) + \Phi(x)z'(x) \end{aligned}$$

gilt $y'(x) = A(x)y(x) + b(x)$ genau dann, wenn

$$\Phi(x)z'(x) = b(x)$$

ist, d.h.

$$z'(x) = (\Phi(x))^{-1} \cdot b(x),$$

und damit folgt

$$z(x) = c + \int_{x_0}^x (\Phi(t))^{-1} b(t) dt.$$

Damit ist Satz (12.12) gezeigt. □

Manchmal lässt sich die Gestalt von $A(x)$ vereinfachen, indem man das DGL-System mit einer nichtsingulären, auf I stetig differenzierbaren $n \times n$ -Matrix $S(x)$ transformiert:

$$y(x) := S(x) \cdot z(x).$$

Dann hat man

$$y'(x) = S'(x) \cdot z(x) + S(x) \cdot z'(x)$$

und andererseits

$$y'(x) = A(x)y + b(x),$$

also:

$$z'(x) = S^{-1}(x)(A(x)S(x) - S'(x)) \cdot z(x) + S^{-1}(x) \cdot b(x).$$

Die Methode der Transformation des DGL-Systems ist insbesondere bei DGL-Systemen mit konstanten Koeffizienten sehr nützlich. Hierzu mehr im nächsten Paragraphen.

1.7 Lineare Differentialgleichungssysteme mit konstanten Koeffizienten

Es geht um DGL-Systeme der Form

$$y'(x) = A \cdot y(x) + b(x) \quad (1.35)$$

mit einer **konstanten** $n \times n$ -Matrix $A = (a_{ik})$. Die Funktion $b : I \rightarrow \mathbb{R}$ (oder \mathbb{C}) sei stetig. Komplexwertige Funktionen bzw. komplexwertige Koeffizienten werden in diesem Paragraphen ausdrücklich zugelassen.

In dem Beweis von Satz (1.24) hatten wir die Norm $|A|$ einer quadratischen $n \times n$ -Matrix $A = (a_{ik})$ mit Koeffizienten in \mathbb{R} (bzw. \mathbb{C}) eingeführt:

$$|A| := n \cdot \sup_{i,k} |a_{ik}|.$$

Es gelten folgende Eigenschaften:

$$\begin{aligned} |c \cdot A| &= |c| \cdot |A| \quad (c \in \mathbb{C}), \\ |A + B| &\leq |A| + |B|, \\ |AB| &\leq |A| \cdot |B|, \end{aligned}$$

und insbesondere $|A^m| \leq |A|^m$ für $m \in \mathbb{N}$.

Für $m \in \mathbb{N}$ ist

$$s_m(A) := \sum_{\mu=0}^m \frac{1}{\mu!} A^\mu$$

wieder eine quadratische Matrix, und wegen

$$|s_m(A)| \leq \sum_{\mu=0}^m \frac{1}{\mu!} |A|^\mu$$

folgt sofort die Existenz von $\lim_{m \rightarrow \infty} s_m(A)$. Dabei ist der Grenzübergang komponentenweise zu verstehen.

Wir können daher definieren:

$$\exp A := \sum_{\mu=0}^{\infty} \frac{1}{\mu!} A^\mu. \quad (1.36)$$

Ist A reell, so auch $\exp A$, und ferner gilt: $\exp 0 = E = E_n$ ist die Einheitsmatrix.

Multipliziert man A mit der reellen oder komplexen Zahl x , so gilt:

$$\exp(Ax) = \sum_{\mu=0}^{\infty} \frac{1}{\mu!} (Ax)^{\mu} = \sum_{\mu=0}^{\infty} \frac{A^{\mu}}{\mu!} x^{\mu}$$

ist eine auf ganz \mathbb{C} konvergente Potenzreihe, insbesondere ist $\exp(Ax)$ beliebig oft differenzierbar, und es gilt

$$\begin{aligned} \exp(Ax)' &= \sum_{\mu=1}^{\infty} \frac{A^{\mu}}{(\mu-1)!} x^{\mu-1} = A \cdot \sum_{\mu=1}^{\infty} \frac{A^{\mu-1}}{(\mu-1)!} x^{\mu-1} = \\ &= A \cdot \exp(Ax) = \exp(Ax) \cdot A. \end{aligned} \quad (1.37)$$

Hieraus folgt mit Induktion

$$(\exp(Ax))^{(m)} = A^m \cdot \exp Ax = \exp(Ax) \cdot A^m.$$

Entwickelt man $\exp Ax$ um einen beliebigen Punkt $x_1 \in \mathbb{C}$ in eine Taylorreihe, so erhält man

$$\begin{aligned} \exp Ax &= \sum_{\mu=0}^{\infty} \frac{(\exp Ax_1) A^{\mu}}{\mu!} (x - x_1)^{\mu} = \\ &= \exp Ax_1 \cdot \exp A(x - x_1), \end{aligned}$$

d.h. mit $x_2 = x - x_1$ ergibt sich

$$\exp A(x_1 + x_2) = \exp Ax_1 \cdot \exp Ax_2.$$

Speziell für $x_1 = 1$, $x_2 = -1$ hat man

$$E = \exp 0 = \exp A \cdot \exp(-A),$$

d.h. die Matrix $\exp A$ ist nichtsingulär, und es ist

$$(\exp A)^{-1} = \exp(-A).$$

Zurück zum DGL-System (1.35) und dem zugehörigen homogenen System

$$y' = Ay. \quad (1.38)$$

Für $\Phi(x) := \exp Ax$ gilt nach dem oben Gezeigten:

$$\Phi'(x) = A \cdot \Phi(x),$$

und zerlegt man Φ in Spaltenvektoren $\Phi = (\varphi_1, \dots, \varphi_n)$, so hat man also

$$\Phi'_{\nu}(x) = A \cdot \varphi_{\nu} \quad \text{für } \nu = 1, \dots, n.$$

Es gilt daher der

Satz 1.32. Die Matrix $\Phi(x) = \exp(Ax)$ ist eine Fundamentalmatrix des DGL-Systems $y' = Ay$.

Problem: Wie kann man $\exp(Ax)$ explizit berechnen? Es stellt sich heraus, dass man die unendliche Reihe $\exp(Ax) = \sum_{\mu=0}^{\infty} \frac{A^\mu}{\mu!} x^\mu$ nicht wirklich zu bilden braucht, sondern mit einigen algebraischen Umformungen auskommt. Dazu bemerken wir zunächst:

Ist S eine nichtsinguläre $n \times n$ -Matrix, so gilt

$$(S^{-1}AS)^\mu = S^{-1}A^\mu \cdot S \quad \text{für } \mu = 0, 1, \dots,$$

also:

$$\begin{aligned} \exp(S^{-1}ASx) &= \sum_{\mu=0}^{\infty} \frac{(S^{-1}AS)^\mu}{\mu!} x^\mu = \\ &= S^{-1} \cdot \left(\sum_{\mu=0}^{\infty} \frac{A^\mu}{\mu!} x^\mu \right) \cdot S = S^{-1} \cdot \exp(Ax) \cdot S, \end{aligned}$$

d.h. $\exp Ax = S \cdot \exp(S^{-1}AS) \cdot S^{-1}$.

Man wähle nun S so, dass $\exp(S^{-1}ASx)$ sich möglichst einfach berechnen lässt.

1. Spezialfall: Wir nehmen an, man kann S so wählen, dass $A^* := S^{-1}AS$ eine Diagonalmatrix wird (komplexe Diagonalelemente seien zugelassen):

$$A^* = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & & \ddots & \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} =: [\lambda_1, \dots, \lambda_n].$$

Dann gilt

$$(A^*)^\mu = [\lambda_1^\mu, \dots, \lambda_n^\mu]$$

und daher

$$\begin{aligned} \exp(A^*x) &= \sum_{\mu=0}^{\infty} \frac{(A^*x)^\mu}{\mu!} = \sum_{\mu=0}^{\infty} \left[\frac{\lambda_1^\mu}{\mu!} x^\mu, \dots, \frac{\lambda_n^\mu}{\mu!} x^\mu \right] = \\ &= \left[\sum_{\mu=0}^{\infty} \frac{\lambda_1^\mu}{\mu!} x^\mu, \dots, \sum_{\mu=0}^{\infty} \frac{\lambda_n^\mu}{\mu!} x^\mu \right] = [e^{\lambda_1 x}, \dots, e^{\lambda_n x}]. \end{aligned}$$

Nun ist mit $S \cdot \exp A^*x \cdot S^{-1}$ auch $S \cdot \exp A^*x$ eine Fundamentalmatrix (vergl. Satz 1.28). Bezeichnet man die Spalten von S mit y_1, \dots, y_n , so sind

$$y_1 \exp(\lambda_1 x), \dots, y_n \exp(\lambda_n x)$$

die Spalten von $S \cdot \exp(A^*x)$.

Es bleiben die Zahlen λ_ν und die Vektoren y_ν zu bestimmen.

Feststellung 1.33. Die Zahlen λ_ν sind die Eigenwerte von A , d.h. die Nullstellen des charakteristischen Polynoms der Matrix A :

$$\det(A - Ex) = \chi_A(x) = \prod_{\nu=1}^n (\lambda_\nu - x),$$

wobei dieses Produkt ein Polynom n -ten Grades in x ist.

Feststellung 1.34. y_1, \dots, y_n sind Eigenvektoren von A zu den Eigenwerten $\lambda_1, \dots, \lambda_n$. Denn sind e_ν für $\nu = 1, \dots, n$ die Einheitsvektoren, so ist offenbar

$$A^* \cdot e_\nu = \lambda_\nu e_\nu,$$

d.h. $S^{-1}ASe_\nu = e_\nu \cdot \lambda_\nu$, was gleichbedeutend ist mit $ASe_\nu = S \cdot e_\nu \lambda_\nu$.

Die Spalten von S waren aber gerade die Vektoren y_ν , d.h. es ist $Se_\nu = y_\nu$. Daher muss gelten

$$A \cdot y_\nu = y_\nu \cdot \lambda_\nu \quad \text{für } \nu = 1, \dots, n.$$

Sind umgekehrt y_ν Eigenvektoren von A zum Eigenwert λ_ν und setzt man

$$S = (y_1, \dots, y_n),$$

so gilt

$$AS = S \cdot [\lambda_1, \dots, \lambda_n].$$

Wegen $\det(A - E\lambda_\nu) = 0$ gibt es zu jedem Eigenwert λ_ν von A von Null verschiedene Eigenvektoren.

Es bleibt die Frage, ob man die $y_\nu \neq 0$ so wählen kann, dass $S = (y_1, \dots, y_n)$ nicht singulär wird.

Das ist im Allgemeinen leider nicht möglich. Es gilt aber Folgendes: Sind die Eigenwerte $\lambda_1, \dots, \lambda_n$ von A paarweise verschieden, und ist y_ν für $\nu = 1, \dots, n$ ein nicht verschwindender Eigenvektor zu λ_ν , so sind y_1, \dots, y_n linear unabhängig und damit $S = (y_1, \dots, y_n)$ nichtsingulär. (Zum Beweis vergl. [8].) Man hat daher den

Satz 1.35. Die Eigenwerte $\lambda_1, \dots, \lambda_n$ der Matrix A seien paarweise verschieden. Ist y_ν ein nicht verschwindender Eigenvektor von A zum Eigenwert λ_ν für $\nu = 1, \dots, n$, so bilden

$$y_1 e^{\lambda_1 x}, \dots, y_n e^{\lambda_n x}$$

ein Fundamentalsystem von Lösungen des DGL-Systems $y' = Ay$.

Bemerkung: Das Fundamentalsystem ist im Allgemeinen nicht reell, selbst wenn A reell ist. Um ein reelles Fundamentalsystem zu erhalten, bildet man

$$\operatorname{Re}(y_1 e^{\lambda_1 x}), \operatorname{Im}(y_1 e^{\lambda_1 x}), \dots, \operatorname{Re}(y_n e^{\lambda_n x}), \operatorname{Im}(y_n e^{\lambda_n x}).$$

Dies sind $2n$ viele reelle Lösungen, unter denen man n über \mathbb{R} linear unabhängige Lösungen finden kann.

2. Fall = Allgemeiner Fall: Man kann im Allgemeinen keine Diagonalgestalt von A finden, aber es gilt:

Satz 1.36 (Jordansche Normalform). Jede $n \times n$ -Matrix A kann durch Transformation mit einer nichtsingulären Matrix S auf die sog. **Jordansche Normalform** gebracht werden. D.h. $S^{-1}AS$ besteht dann aus Kästchen der Form

$$\begin{pmatrix} \lambda_k & 1 & 0 & \dots & 0 \\ 0 & \lambda_k & 1 & \dots & 0 \\ 0 & 0 & \lambda_k & \dots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & 0 & 0 & \ddots & 1 \\ 0 & 0 & 0 & \dots & \lambda_k \end{pmatrix}$$

für $k = 1, \dots, s$. Dabei stehen die verschiedenen Eigenwerte λ_k von A entsprechend ihrer Vielfachheit in der Hauptdiagonalen der Matrix, und in jedem Kästchen tritt jeweils nur ein Eigenwert auf, möglicherweise in mehreren Kästchen derselbe.

Wenn man die gleichen Eigenwerte in Kästchen zusammenfasst, lässt sich folgende Form erreichen: $S^{-1}AS = A^*$ besteht aus Kästchen der Form

$$\begin{pmatrix} \lambda_k & * & * & \dots & * \\ 0 & \lambda_k & * & \dots & * \\ 0 & 0 & \lambda_k & \dots & * \\ \vdots & & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda_k \end{pmatrix}$$

mit jeweils der Größe n_k . Dabei deuten die Sterne an, dass in den betreffenden Bereichen des Schemas möglicherweise von Null verschiedene Elemente stehen. Genauer: außerhalb der Diagonalen stehen Nullen, ausgenommen die Diagonale, die die Elemente mit Index $(\nu, \nu + 1)$ enthält und die aus Einsen besteht.

Schreibt man abkürzend $A^* = [A_1, \dots, A_k]$, wobei A_ν wie oben ist, so hat man die Aufgabe, $\exp A^*x$ zu berechnen.

Wegen $(A^*)^\nu = [A_1^\nu, \dots, A_k^\nu]$ ist

$$\exp(A^*x) = [\exp A_1x, \dots, \exp A_kx].$$

Man braucht daher nur noch $\exp Bx$ für eine Matrix B der Form

$$B = \begin{pmatrix} \lambda & * & * & \cdots & * \\ 0 & \lambda & * & \cdots & * \\ 0 & 0 & \lambda & \cdots & * \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda \end{pmatrix}$$

zu berechnen.

Es ist $Bx = E\lambda x + (B - E\lambda)x$ (wobei E die Einheitsmatrix bezeichnet).

Die Matrizen $E\lambda x$ und $(B - E\lambda)x$ sind offenbar vertauschbar, und deswegen gilt

$$\exp Bx = \exp(E\lambda x) \cdot \exp((B - E\lambda)x).$$

$B - E\lambda$ ist eine Matrix, die in und unterhalb der Hauptdiagonalen nur Nullen enthält. Die n -te Potenz einer solchen Matrix ist aber die Nullmatrix, wenn n die Reihenzahl ist. Denn mit vollständiger Induktion sieht man:

Ist C eine Matrix, die in und unterhalb der Hauptdiagonalen nur Nullen enthält, so verschwinden in C^μ alle Elemente $c_{ik}^{(\mu)}$ mit $k < i + \mu$. Das ist für $\mu = 1$ richtig, und wenn dies für ein $\mu \geq 1$ gilt, so folgt

$$c_{ik}^{(\mu+1)} = \sum_{\nu=1}^m c_{i\nu} c_{\nu k}^{(\mu)} = \sum_{\nu=i+1}^{k-\nu} c_{i\nu} c_{\nu k}^{(\mu)} = 0 \quad \text{für } k < i + \mu + 1,$$

d.h. die Aussage gilt für $\mu + 1$.

Die Reihe für $\exp((B - E\lambda)x)$ bricht daher spätestens mit dem $(m - 1)$ -ten Glied ab:

$$\exp((B - E\lambda)x) = \sum_{\mu=0}^{m-1} \frac{(B - E\lambda)^\mu}{\mu!} x^\mu.$$

Die Koeffizienten sind Polynome höchstens vom Grad $m - 1$. Es ist

$$\exp(E_m \lambda x) = e_m \cdot e^{\lambda x},$$

und damit besteht $\exp(A^* x)$ aus Kästchen der Form

$$e^{\lambda_\kappa x} \cdot Q_\kappa(x).$$

Dabei ist Q_κ eine n_κ -reihige Matrix, deren Hauptdiagonale aus Einsen besteht und die unterhalb der Hauptdiagonale nur Nullen enthält. Die Elemente oberhalb der Hauptdiagonale sind Polynome $q_{\mu\nu}^\kappa(x)$ vom Grad $\leq \nu - \mu \leq n_\kappa - 1$.

Die Matrix $S \cdot \exp A^* x$ ist Fundamentalmatrix von

$$y' = Ay.$$

Bezeichnet man mit y_1, \dots, y_n die Spalten von S , so sind die ersten n_1 Spalten von $S \cdot \exp(A^* x)$

$$y_1 e^{\lambda_1 x}, \quad (y_2 + q_{12}^{(1)}(x)y_1) e^{\lambda_1 x}, \dots, (y_{n_1+1} + q_{n_1-1, n_1}^1(x) \cdot y_{n_1-1} + \dots + q_{1, n_1}^1(x)y_1) e^{\lambda_1 x}.$$

Man erhält so insgesamt den

Satz 1.37. *Die verschiedenen Eigenwerte der Matrix A seien $\lambda_1, \dots, \lambda_k$, ihre Vielfachheiten seien n_1, \dots, n_k . Dann gibt es ein Fundamentalsystem von $y' = Ay$, welches aus jeweils n_κ Lösungen der Gestalt $q(x) \cdot e^{\lambda_\kappa x}$ besteht, dabei ist $q(x)$ ein Vektor von Polynomen des Grades $\leq n_\kappa - 1$ für $\kappa = 1, \dots, k$.*

Beispiele: Zunächst betrachten wir das System

$$y' = Ay \quad \text{mit} \quad A = \begin{pmatrix} 0 & 1 \\ -b & -2a \end{pmatrix} \quad (1.39)$$

(dies entspricht der linearen DGL zweiter Ordnung: $y'' + 2ay' + by = 0$). Dann ist $\det(A - Ex) = x^2 + 2ax + b$. Die Eigenwerte von A sind

$$\lambda_1 = -a + \sqrt{a^2 - b}$$

und

$$\lambda_2 = -a - \sqrt{a^2 - b}.$$

Man rechnet nach: $\begin{pmatrix} 1 \\ \lambda_\nu \end{pmatrix}$ ist für $\nu = 1, 2$ Eigenvektor zum Eigenwert λ_ν . Falls $a^2 - b \neq 0$, also $\lambda_1 \neq \lambda_2$, so hat man die zwei linear unabhängigen Lösungen

$$\varphi_1(x) = \begin{pmatrix} 1 \\ \lambda_1 \end{pmatrix} e^{\lambda_1 x}, \quad \varphi_2(x) = \begin{pmatrix} 1 \\ \lambda_2 \end{pmatrix} e^{\lambda_2 x}.$$

Ist $a^2 - b = 0$, so ist der Lösungsraum der linearen Gleichung $(A + aE)y = 0$ eindimensional.

Eigenvektor ist $\begin{pmatrix} 1 \\ -a \end{pmatrix}$, und es gibt daher keine nichtsinguläre Matrix S , deren Spalten Eigenvektoren von A sind. D.h.: A kann nicht auf Diagonalgestalt transformiert werden. Die Jordansche Normalform von A ist

$$A^* = \begin{pmatrix} -a & 1 \\ 0 & -a \end{pmatrix}$$

Man rechnet leicht nach, dass mit $S = \begin{pmatrix} 1 & 0 \\ -a & 1 \end{pmatrix}$

$$S^{-1} = \begin{pmatrix} 1 & 0 \\ a & 1 \end{pmatrix}$$

gilt. Es ist

$$\begin{aligned} \exp(A^*x) &= e^{-ax} \cdot \exp \begin{pmatrix} 0 & x \\ 0 & 0 \end{pmatrix} = \\ &= e^{-ax} \left(E + \begin{pmatrix} 0 & x \\ 0 & 0 \end{pmatrix} \right) = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} e^{-ax} \end{aligned}$$

und damit

$$S \cdot \exp(A^*x) = \begin{pmatrix} 1 & x \\ -a & 1 - ax \end{pmatrix}$$

Fundamentalsystem ist

$$\varphi_1(x) = \begin{pmatrix} 1 \\ -a \end{pmatrix} e^{-ax}, \quad \varphi_2(x) = \begin{pmatrix} x \\ 1 - ax \end{pmatrix} e^{-ax}.$$

Sind a, b reell, so ist auch dieses Fundamentalsystem reell.

Im ersten Fall, also wenn $a^2 - b \neq 0$ gilt, müssen λ_1, λ_2 nicht mehr notwendig reell sein: Ist $a^2 - b < 0$, so sind $y_i \in \mathbb{C} \setminus \mathbb{R}$ zwei konjugiert komplexe Nullstellen, nämlich

$$\lambda_j = -a \pm \sqrt{a^2 - b} = -a \pm i\sqrt{b - a^2}$$

für $j = 1, 2$.

Setzt man $\omega := \sqrt{b - a^2}$, so hat man daher aus

$$\operatorname{Re} \varphi_1, \quad \operatorname{Im} \varphi_1, \quad \operatorname{Re} \varphi_2, \quad \operatorname{Im} \varphi_2$$

ein reelles Fundamentalsystem zu bestimmen:

Für die ersten Komponenten der Spaltenvektoren erhält man so als reelles Fundamentalsystem

$$e^{-ax} \cdot \cos(\omega x), \quad e^{-ax} \cdot \sin(\omega x).$$

2 Kurven- und Flächenintegrale, Vektoranalysis

Dieser Teil ist keine Ausarbeitung der Vorlesung, sondern nur ein kurzes Exzerpt, das die wesentlichen Sätze und Definitionen enthält. Die Beweise werden in diesem Exzerpt nicht ausformuliert.

2.1 Orientierte Flächen

Definition 2.1. Eine Orientierung eines Vektorraums V über \mathbb{R} wird gegeben durch Angabe einer geordneten Basis e_1, \dots, e_n von V . Ist f_1, \dots, f_n eine weitere Basis, so definiert diese dieselbe (bzw. entgegengesetzte) Orientierung, wenn für die durch

$$f_i = \sum_{k=1}^n a_{ik} e_k \quad i = 1, \dots, n$$

gegebene Transformation $\det(a_{ik}) > 0$ (bzw. < 0) gilt.

Die Basen von V werden auf diese Weise in zwei Klassen eingeteilt, und jede Klasse definiert eine Orientierung von V . Jeder n -dimensionale euklidische Vektorraum hat genau zwei Orientierungen. Ist etwa $V = \mathbb{R}^3$, so definiert die kanonische Basis (e_1, e_2, e_3) die positive Orientierung. Diese kann man sich mit den Fingern der rechten Hand oder der "Schraubenregel" veranschaulichen.

Sei weiter V ein n -dimensionaler orientierter Vektorraum und H der Halbraum, der von der Hyperebene W begrenzt wird. Dann gibt es eine Basis e_1, \dots, e_n von V mit folgenden Eigenschaften:

- (1) e_2, \dots, e_n bildet eine Basis von W ;
- (2) $e_1 \notin H$.

Die Basis e_1, \dots, e_n repräsentiert eine Orientierung von V , und e_2, \dots, e_n repräsentiert eine Orientierung von W . Man nennt diese die durch die in V gegebene Orientierung **induzierte** Orientierung von W . Diese ist unabhängig von der speziellen Wahl der Basis $\{e_1, \dots, e_n\}$.

Definition 2.2. Eine Punktmenge $F \subseteq \mathbb{R}^n$ heißt **p -dimensionales glattes Flächenstück**, wenn es eine offene Punktmenge $A \subseteq \mathbb{R}^p$ und eine Abbildung $\varphi : A \rightarrow F$ gibt mit:

- (1) φ ist topologische Abbildung: d.h. bijektiv, stetig, und φ^{-1} ist stetig.
- (2) φ ist stetig differenzierbar.
- (3) $\text{Rang} \left(\frac{\partial \varphi}{\partial t_j} \right) = p$ für alle $t \in A$.

Seien $\varphi_1, \dots, \varphi_n$ die Koordinatenfunktionen von φ bezüglich irgendeines Koordinatensystems des \mathbb{R}^n . Dann heißt (A, φ) **Parameterdarstellung** des Flächenbereiches.

Bemerkung.

- (i) Der Fall $p = n$ ist zugelassen, d.h. offene Teilmengen des \mathbb{R}^n sind auch n -dimensionale glatte Flächenstücke.
- (ii) Für $p = 1$ erhält man eine glatte Kurve mit gewissen Zusatzbedingungen: φ ist in diesem Fall injektiv, d.h. enthält keine Selbstdurchdringungen und wird auch nicht mehrfach durchlaufen.
- (iii) Am Rand kann ein Flächenstück „wild“ werden, wie dies z.B. bei dem glatten 2-dimensionalen Flächenstück

$$z = \cos \frac{r}{1-r}, \quad r = \sqrt{x^2 + y^2}, \quad x^2 + y^2 < 1$$

der Fall ist: z , in Abhängigkeit von r , wird „wild“ bei Annäherung $r \rightarrow 1$. Um solche Fälle auszuschließen, definiert man:

Definition 2.3. Ein p -dimensionales glattes Flächenstück F heißt **glatt bis zum Rand**, wenn F eine Parameterdarstellung (A, φ) besitzt, so dass A beschränkt ist und sich φ auf einer offenen Menge $B \subseteq \mathbb{R}^p$ mit $B \supseteq \bar{A}$ fortsetzen läßt, so dass $\varphi : B \rightarrow \varphi(B) \subseteq \mathbb{R}^n$ noch die Bedingungen (1), (2), (3) aus (2.2) erfüllt.

Beispiel 2.4. Gegeben sei die Kegelfläche $z = \sqrt{x^2 + y^2}$ für $0 < x^2 + y^2 < 1$, also ohne die Spitze. Diese Kegelfläche ist ein glattes Flächenstück, allerdings nicht glatt bis zum Rand, da an der Spitze keine Tangentialebene existiert.

Verschiedene Parameterdarstellungen hängen folgendermaßen zusammen:

Satz 2.5. Sind $(A, \varphi), (B, \psi)$ zwei Parameterdarstellungen desselben glatten Flächenstücks F , dann existiert eine topologische stetig differenzierbare Abbildung $\chi : A \rightarrow B$ mit $\det \left(\frac{\partial \chi_i}{\partial t_j} \right) \neq 0$ und $\varphi(t) = \psi(\chi(t))$ für $t \in A$. D.h. $\varphi = \psi \circ \chi$.

Wir führen nun die Orientierung eines Flächenstücks ein. Diese ist gegeben durch eine Parameterdarstellung, indem man \mathbb{R}^p durch die kanonische Basis e_1, \dots, e_p als orientiert betrachtet. (Wie gewohnt sei $e_i = (0, \dots, 0, 1, 0, \dots, 0)$ mit der 1 an der i -ten Stelle.) Zwei Parameterdarstellungen $(A, \varphi), (B, \psi)$ liefern dieselbe (oder entgegengesetzte) Orientierung von F , falls gilt

$$\det \left(\frac{\partial \chi_i}{\partial t_j} \right) > 0 \quad (< 0)$$

für alle $t \in A$. Dabei setzen wir $\chi := \psi^{-1} \circ \varphi$.

Bemerkung. Es kann vorkommen, dass die Funktionaldeterminante in bestimmten Teilbereichen von A positiv und in anderen negativ ist. Auf diese Weise zerfällt ein Flächenstück möglicherweise in mehrere, voneinander getrennte Teile.

Definition 2.6. Eine Punktmenge F des \mathbb{R}^n heißt **p -dimensionale glatte Fläche**, wenn es zu jedem Punkt $x \in F$ eine offene Kugel $K(x, r) \subseteq \mathbb{R}^n$ gibt, so dass $K \cap F$ ein p -dimensionales glattes Flächenstück ist.

(Glatte Flächen können durch glatte Flächenstücke F_i — i aus einer Indexmenge I — überdeckt werden, wobei I nicht endlich oder abzählbar zu sein braucht.) F heißt **orientierbar**, wenn es eine Überdeckung durch Flächenstücke F_i gibt mit **verträglichen** Parameterdarstellungen (A_i, φ_i) , d.h.

$$F_i \cap F_j \neq \emptyset \Rightarrow \det \left(\frac{\partial (\chi_{ij})_k}{\partial t_i} \right) > 0$$

mit

$$\chi_{ij} : \varphi_i^{-1}(F_i \cap F_j) \rightarrow \varphi_j^{-1}(F_i \cap F_j).$$

Zwei Orientierungen von F mittels $\{F_i\}_{i \in I}$ bzw. $\{F'_j\}_{j \in J}$ heißen **gleich**, wenn die durch F_i definierte Orientierung mit der durch F'_j definierten Orientierung verträglich ist im obigen Sinne.

Bemerkung. Es gibt Beispiele von Flächen, die keine Orientierung besitzen — etwa das Möbiusband.

Wir betrachten nun die induzierte Orientierung des Randes einer orientierten Fläche F . Dabei stellt sich die Frage, was „Randpunkte“ einer Fläche sind. Damit sind nicht Randpunkte von F im \mathbb{R}^n gemeint, da sonst im Fall $p < n$ jeder Punkt von F ein Randpunkt wäre.

Wir definieren

$$\bar{\partial}F := \bar{F} \setminus F$$

als Randpunkte der Fläche F .

Bemerkung. Nicht jede Fläche hat Randpunkte. Die Fläche

$$\{x = (x_1, \dots, x_n) : \sum_{i=1}^n x_i^2 = 1\}$$

hat beispielsweise keine Randpunkte.

Definition 2.7. Ein Punkt $x \in \bar{\partial}F$ heißt **guter Randpunkt**, wenn es eine Kugel $K(x, r)$ gibt, so dass $F \cap K$ ein bis zum Rand glattes Flächenstück ist mit einer im Sinne von (2.3) fortsetzbaren Parameterdarstellung (A, φ) mit: $\varphi(0) = x$, $A \subseteq H = \{t : t_1 < 0\}$, und es gibt ein $d > 0$ so dass $H \cap K(0, d) \subseteq A$.

Bemerkung. Sämtliche Randpunkte aus einer genügend kleinen Umgebung eines guten Randpunktes x sind wieder gute Randpunkte.

Bezeichnung: ∂F sei definiert als die Menge der guten Randpunkte von F .

Bemerkung. ∂F ist eine $(p - 1)$ -dimensionale glatte Fläche. Ist F orientierbar, so induziert diese Orientierung eine solche auf ∂F . Diese hängt nicht von der speziellen Wahl der Parameterdarstellung (A_i, φ_i) ab.

Ist speziell $A \subseteq \mathbb{R}^2$ ein zweidimensionales glattes Flächenstück, so ist ∂A eine glatte Kurve. Wählt man die Orientierung auf A , die durch $x_1 = x$, $x_2 = y$ gegeben ist, so ist die auf ∂A induzierte Orientierung gerade diejenige Durchlaufrichtung, bei der A „zur Linken“ liegt (mathematisch positive Richtung: gegen den Uhrzeigersinn).

Hilfssatz 2.8. *Zu einer p -dimensionalen glatten Fläche F gibt es eine (möglicherweise abbrechende) Folge von glatten beschränkten Flächenstücken F_i , die bis zum Rand glatt sind und folgende Eigenschaften haben:*

- (1) $F = \cup_{i=1}^{\infty} F_i$.
- (2) Ist K kompakt und $K \subseteq F$, dann existiert eine Zahl $i(K)$, so dass $F_i \cap K = \emptyset$ für alle $i \geq i(K)$.

Bemerkung.

- (i) Für jedes i ist $\overline{F_i}$ kompakt. Denn für jede Überdeckung von F im Sinne von (2) kann, bei festem i , $F_i \cap F_j \neq \emptyset$ nur für endlich viele F_j gelten.
- (ii) Aus (2.8) folgt: Jede kompakte glatte Fläche kann schon durch endlich viele bis zum Rand glatte Flächenstücke überdeckt werden.

2.2 Flächenintegrale

Wir haben bisher über einen Bereich A des \mathbb{R}^p integriert, wobei dieses Integral als Grenzwert einer geeigneten Folge von Näherungssummen definiert war. Nun soll über eine p -dimensionale Fläche im \mathbb{R}^n integriert werden. Dabei stellt sich das Problem, dass wir bisher noch keine Maße auf gekrümmten Flächen kennen, und es fragt sich, welche Art von Funktionen überhaupt integriert werden soll.

Hierzu führen wir folgende Überlegung durch: Der Parameterbereich A einer Parameterdarstellung (A, φ) von F wird in kleine Teilbereiche A_i zerlegt. In jedem A_i wählen wir einen Punkt $t^{(i)}$. Anstelle von $\varphi(A_i) \subseteq F$ betrachten wir das Bild von A_i unter der zum Punkt $t^{(i)}$ gehörenden approximierenden linearen Abbildung $\varphi'(t^{(i)})$. Dieses Bild ist ein kleiner Bereich aus der p -dimensionalen Tangentialebene von F im Punkt $\varphi(t^{(i)})$. Ist dann m_i der p -dimensionale Inhalt dieses Bereichs, so ist

$$\sum_i f(\varphi(t^{(i)}))m_i$$

eine Näherungssumme für das gesuchte Integral. Es fehlt allerdings noch eine Berücksichtigung der Orientierung, und wir haben gesehen, dass die Orientierung (z.B. bei Kurvenintegralen) eine wichtige Rolle spielt. Ein Beispiel

aus der Physik: Sei $v(x)$ das Geschwindigkeitsfeld einer Flüssigkeit in Abhängigkeit von $x \in \mathbb{R}^3$ und F ein Flächenstück; wir fragen nach der pro Zeiteinheit durch F hindurchströmenden Flüssigkeitsmenge. Diese hängt offenbar von der Richtung der Flüssigkeit durch F ab. Ist nun F orientiert und sind a_1, a_2 zwei Tangentialvektoren in einem Punkt $x \in F$, welche die gegebene Orientierung repräsentieren, so ist

$$\det(a_1, a_2, v(x)) > 0 \quad \text{oder} \quad < 0$$

je nachdem, ob $v(x)$ nach der einen oder der anderen Seite von F zeigt. Der Absolutbetrag

$$|\det(a_1, a_2, v(x))|$$

ist das Volumen des durch $a_1, a_2, v(x)$ aufgespannten Parallelotops und damit näherungsweise die durch das Flächenelement hindurchströmende Flüssigkeitsmenge. Durch die Orientierung wird festgelegt, welche Richtung als die positive angesehen wird. Die Determinante ändert bei einer Vertauschung von a_1 und a_2 , also einer Änderung der Orientierung, das Vorzeichen. Allgemeiner gilt: Falls $b_i = \sum_{j=1}^2 c_{ij} a_j$ für $i = 1, 2$ gilt, so ist

$$\det(b_1, b_2, v) = \det(a_1, a_2, v) \cdot \det(c_{ij}).$$

Daher betrachten wir nun Funktionen des folgenden Typs, wobei auch der Fall der nichtorientierten Fläche mit berücksichtigt wird:

Definition 2.9. Sei F eine p -dimensionale glatte Fläche im \mathbb{R}^n . Eine Funktion $\phi : F \times (\mathbb{R}^n)^p \rightarrow \mathbb{R}$ heiÙe **vom Typ (2.9)**, falls gilt: Mit $b_i = \sum_j c_{ij} a_j$, $i = 1, \dots, p$ ist

$$(a) \quad \phi(x; b_1, \dots, b_p) = \phi(x; a_1, \dots, a_p) \cdot |\det(c_{ij})|, \text{ bzw. falls } F \text{ orientiert ist}$$

$$(b) \quad \phi(x; b_1, \dots, b_p) = \phi(x; a_1, \dots, a_p) \cdot \det(c_{ij}).$$

Bemerkung.

(i) Ist ϕ vom Typ (2.9) und $g : F \rightarrow \mathbb{R}$ eine beliebige Funktion, so ist $g(x) \cdot \phi$ vom Typ (2.9).

(ii) Mit ϕ_1, ϕ_2 ist auch $\phi_1 \pm \phi_2$ vom Typ (2.9).

Nun definieren wir für Funktionen des Typs (2.9) die Integrierbarkeit und das Integral, und zwar zunächst über ein Flächenstück F .

Sei (A, φ) eine Parameterdarstellung von F und $x(t) := \varphi(t)$. Eine Funktion ϕ vom Typ (2.9) (b oder a, je nachdem ob F orientiert ist oder nicht) heißt **integrierbar über F** , wenn

$$\int_F \phi := \int_A \phi(x(t); \frac{\partial x}{\partial t_1}, \dots, \frac{\partial x}{\partial t_p}) dt_1, \dots, dt_p \quad (2.1)$$

existiert. In diesem Fall heißt $\int_F \phi$ das **Flächenintegral** von ϕ über F .

Bemerkung. Diese Definition ist unabhängig von der Wahl der Parameterdarstellung von F . Um dies zu zeigen, wendet man die p -dimensionale Kettenregel und die Transformationsformel an.

Sei nun F kein Flächenstück mehr, sondern eine ganze Fläche. Wie wäre in solch einem Fall $\int_F \phi$ zu definieren? Man tut dies wie folgt: man wählt eine Zerlegung $F = \cup_i F_i$ in paarweise disjunkte Flächenstücke F_i und setzt

$$\int_F \phi = \sum_i \int_{F_i} \phi.$$

D.h. wenn $\chi_i(x) = \chi_{F_i}(x)$ die charakteristische Funktion von F_i ist, soll gelten

$$\int_F \phi = \sum_i \int_{F_i} \chi_i \phi.$$

Allgemeiner definieren wir

Definition 2.10. *Eine Zerlegung der Einheit (Partition der Einheit) von F ist eine Folge χ_i ($i = 1, 2, \dots$) mit*

- (a) $0 \leq \chi_i(x) \leq 1$ für alle $x \in F$;
- (b) $\sum_{i=1}^{\infty} \chi_i(x) \equiv 1$ auf F ;
- (c) Für jedes i existiert ein bis zum Rand glattes Flächenstück F_i , so dass $\chi_i(x) = 0$ für alle $x \notin F_i$ gilt und $\{F_i\}$ eine Überdeckung von F im Sinne von (2.8) ist. Ist $(A, x(t))$ eine Parameterdarstellung von F_i , so sei $\chi_i(x(t))$ messbar.

Bemerkung.

- (i) Aus (c) folgt, dass es in jedem Punkt $x \in F$ nur endlich viele χ_i mit $\chi_i(x) \neq 0$ gibt.
- (ii) Jeder Überdeckung $\{F_i\}$ von F mit den Eigenschaften (2.8) kann man eine Partition der Einheit auf F zuordnen. Hierzu wählen wir eine zweite Überdeckung $F = \cup_i G_i$, mit $\overline{G_i} \subseteq F_i$, und setzen

$$\eta_i(x) := \chi_{G_i}(x).$$

Dann ist $\sum_i \eta_i(x)$ für jedes $x \in F$ eine endliche Summe und ≥ 1 . Daher leistet

$$\chi_i(x) := \frac{\eta_i}{\sum_i \eta_i(x)}$$

das Gewünschte.

Ist $\{\chi_i\}_{i \in \mathbb{N}}$ eine Partition der Einheit auf F mit zugehörigen Flächenstücken F_i , so definieren wir

$$\int_F \phi := \sum_{i=1}^{\infty} \int_{F_i} \chi_i \phi \tag{2.2}$$

falls $\sum_{i=1}^n \int_{F_i} |\phi \cdot \chi_i|$ konvergiert. Diese Definition ist unabhängig von der Wahl der χ_i .

Wir wollen nun den Flächeninhalt von p -dimensionalen glatten Flächen definieren. Für dieses Problem spielt die Frage der Orientierung keine Rolle. Wir gehen vor wie bei in den Überlegungen zu Beginn dieses Paragraphen. Wir wählen für

$$\phi(x; a_1, \dots, a_p)$$

den p -dimensionalen Inhalt des von a_1, \dots, a_p aufgespannten Parallelotops, d.h. wir verlangen

$$\phi(x; e_1, \dots, e_p) = 1$$

für beliebige Orthonormalsysteme e_1, \dots, e_p und

$$\phi(x; a_1, \dots, a_p) = |\det(c_{ij})|$$

für beliebige Systeme von Vektoren $a_i = \sum_{j=1}^p c_{ij} e_j$ ($i = 1, \dots, p$). Man kann nun zeigen, dass die letztgenannte Bedingung nicht von der speziellen Wahl

der Orthonormalbasis e_1, \dots, e_p abhängt. Wir bezeichnen dann die Funktion ϕ auch mit dF und definieren

$$\int_F dF$$

als den **Flächeninhalt von F** . Ist F ein Flächenstück mit Parameterdarstellung $(A, x(t))$, so gilt

$$\int_F dF = \int_A \sqrt{\det \left(\frac{\partial x}{\partial t_i} \cdot \frac{\partial x}{\partial t_k} \right)} dt_1, \dots, dt_p. \quad (2.3)$$

Spezialfälle:

(i) Für $p = 1$ entspricht dieser Flächeninhalt der Bogenlänge einer Kurve:

$$L(f; a, b) = \int_a^b |f'(t)| dt.$$

(ii) Sei F ein $(n-1)$ -dimensionales Flächenstück im \mathbb{R}^n mit $x_n = f(x_1, \dots, x_{n-1})$.

Wir setzen $(x_1, \dots, x_{n-1}) = (t_1, \dots, t_{n-1})$ und erhalten

$$\frac{\partial x}{\partial t_i} = \left(\delta_{1i}, \dots, \delta_{n-1,i}, \frac{\partial f}{\partial x_i} \right)$$

und damit

$$\frac{\partial x}{\partial t_i} \cdot \frac{\partial x}{\partial t_k} = \delta_{ik} + f_i f_k \quad \text{mit } f_i := \frac{\partial f}{\partial x_i}.$$

Die Determinante $\det \left(\frac{\partial x}{\partial t_i} \cdot \frac{\partial x}{\partial t_k} \right)$ ist damit gleich

$$\begin{vmatrix} 1 + f_1^2 & f_1 f_2 & \cdots & f_1 f_{n-1} \\ f_2 f_1 & 1 + f_2^2 & \cdots & f_2 f_{n-1} \\ \cdots & \cdots & \ddots & \vdots \\ f_{n-1} f_1 & f_{n-1} f_2 & \cdots & 1 + f_{n-1}^2 \end{vmatrix}.$$

Der i -te Zeilenvektor ist

$$(\delta_{1i}, \dots, \delta_{n-1,i}) + (f_i f_1, f_i f_2, \dots, f_i f_{n-1}),$$

und die Determinante wird dementsprechend in eine Summe von 2^{n-1} vielen Einzeldeterminanten zerlegt. Falls in einer dieser Einzeldeterminanten mehr als eine Zeile vom zweiten Typ vorkommt, sind zwei

Zeilen linear abhängig, so dass die Einzeldeterminante in diesem Fall Null wird. Damit bleibt

$$\det \left(\frac{\partial x}{\partial t_i} \cdot \frac{\partial x}{\partial t_k} \right) = 1 + \sum_{i=1}^{n-1} f_i^2.$$

So erhält man für den Flächeninhalt von F die Formel

$$\int_F dF = \int_A \sqrt{1 + \left(\frac{\partial f}{\partial x_1} \right)^2 + \left(\frac{\partial f}{\partial x_2} \right)^2 + \cdots + \left(\frac{\partial f}{\partial x_{n-1}} \right)^2} dx_1 \dots dx_{n-1}. \quad (2.4)$$

Beispielsweise ist dies für $n = 2$ die Bogenlänge der durch $y = f(x)$, $a \leq x \leq b$ definierten Kurve:

$$\int_a^b \sqrt{1 + (f'(t))^2} dt.$$

2.3 Die Graßmannalgebra eines Vektorraums

Sei T ein reeller Vektorraum, $\dim_{\mathbb{R}} T = n$. Man definiert

$$T^* = \text{Hom}_{\mathbb{R}}(T, \mathbb{R}),$$

d.h. T^* sei der n -dimensionale Vektorraum der Linearformen auf T .

Mit T^p bezeichnen wir das p -fache kartesische Produkt von T , also die Menge $\{(\xi_1, \dots, \xi_p) : \xi_i \in T\}$.

Definition 2.11. Eine p -**Linearform** (oder auch **Multilinearform**) ist eine Abbildung $\varphi : T^p \rightarrow \mathbb{R}$ mit folgenden Eigenschaften:

(i) Für $1 \leq \nu \leq p$ gilt

$$\varphi(\xi_1, \dots, \xi'_\nu + \xi''_\nu, \xi_p) = \varphi(\xi_1, \dots, \xi'_\nu, \dots, \xi_p) + \varphi(\xi_1, \dots, \xi''_\nu, \dots, \xi_p).$$

(ii) Für $1 \leq \nu \leq p$ und $a \in \mathbb{R}$ gilt

$$\varphi(\xi_1, \dots, a\xi_\nu, \dots, \xi_p) = a\varphi(\xi_1, \dots, \xi_\nu, \dots, \xi_p).$$

Bemerkung. Eine 0-Linearform ist eine reelle Zahl, und die Menge aller 1-Linearformen auf T ist T^* . Für jedes $p > 0$ bildet die Menge aller p -Linearformen einen reellen Vektorraum. Man bezeichnet p als die **Dimension** von φ oder nennt φ eine **Multilinearform der Dimension p** .

Definition 2.12. Eine p -Multilinearform φ heißt **alternierend** oder **antisymmetrisch**, bzw. **äußere p -Form**, wenn für jedes ν mit $1 \leq \nu < p$

$$\varphi(\xi_1, \dots, \xi_\nu, \xi_{\nu+1}, \dots, \xi_p) = -\varphi(\xi_1, \dots, \xi_{\nu+1}, \xi_\nu, \dots, \xi_p)$$

gilt.

Jede 1-Linearform ist per definitionem alternierend, und die Menge der äußeren Nullformen ($p = 0$) wird als \mathbb{R} definiert.

Satz 2.13. Für jedes p bilden die alternierenden p -Formen einen reellen Vektorraum, den wir mit E^p bezeichnen. Es ist $E^0 = \mathbb{R}$ und $E^1 = T^*$.

Bemerkung.

(i) Ist $\varphi \in E^p$ und $\nu \neq \mu$, so gilt

$$\varphi(\xi_1, \dots, \xi_\nu, \dots, \xi_\mu, \dots, \xi_p) = -\varphi(\xi_1, \dots, \xi_\mu, \dots, \xi_\nu, \dots, \xi_p).$$

(ii) Ist $\varphi \in E^p$ und $\xi_\nu = \xi_\mu$ für gewisse $\nu \neq \mu$, so gilt

$$\varphi(\xi_1, \dots, \xi_\nu, \dots, \xi_\mu, \dots, \xi_p) = 0.$$

Definition 2.14. Sei φ eine p -Multilinearform. Unter dem **alternierenden Anteil von φ** versteht man die durch

$$[\varphi](\xi_1, \dots, \xi_p) := \frac{1}{p!} \sum_{\pi \in \gamma_p} \text{sgn}(\pi) \varphi(\xi_{\pi(1)}, \dots, \xi_{\pi(p)})$$

erklärte p -Multilinearform $[\varphi]$. Für $p = 0$ bzw. $p = 1$ setzt man $[\varphi] := \varphi$. $\text{sgn}(\pi)$ ist das Signum der jeweiligen Permutation π .

Hilfssatz 2.15. Ist φ eine alternierende p -Multilinearform, so ist $[\varphi] = \varphi$.

Hilfssatz 2.16. Für jede p -Multilinearform φ ist $[\varphi]$ eine alternierende p -Multilinearform.

Folgerung 2.17. Es gilt stets $[[\varphi]] = [\varphi]$.

Hilfssatz 2.18. Seien φ, ψ zwei p -Multilinearformen. Dann gilt $[\varphi + \psi] = [\varphi] + [\psi]$, und für $a \in \mathbb{R}$ ist $[a\varphi] = a \cdot [\varphi]$.

Wir definieren nun das Produkt zweier Multilinearformen. Sei φ eine p -Multilinearform und ψ eine q -Multilinearform. Dann sei

$$(\varphi \cdot \psi)(\xi_1, \dots, \xi_p, \eta_1, \dots, \eta_q) = \varphi(\xi_1, \dots, \xi_p) \cdot \psi(\eta_1, \dots, \eta_q).$$

Die Abbildung

$$\varphi \cdot \psi : T^{p+q} \rightarrow \mathbb{R}$$

ist eine $p + q$ -Linearform, und es gelten folgende Regeln:

1. $(\varphi \cdot \psi) \cdot \chi = \varphi \cdot (\psi \cdot \chi)$.
2. $(\varphi + \psi) \cdot \chi = \varphi \cdot \chi + \psi \cdot \chi$.
3. $\varphi \cdot (\psi + \chi) = \varphi \cdot \psi + \varphi \cdot \chi$.
4. $a(\varphi \cdot \psi) = (a\varphi) \cdot \psi = \varphi \cdot (a\psi)$ für alle $a \in \mathbb{R}$.

Hilfssatz 2.19. Es ist $[\varphi \cdot \psi] = [[\varphi] \cdot [\psi]]$.

Definition 2.20. Für alternierende p -Formen φ und alternierende q -Formen ψ wird das **äußere Produkt** $\varphi \wedge \psi$ definiert durch

$$\varphi \wedge \psi = \frac{(p+q)!}{p! \cdot q!} [\varphi \cdot \psi].$$

$\varphi \wedge \psi$ ist eine alternierende $p + q$ -Form, und die Abbildung

$$\wedge : E^p \times E^q \rightarrow E^{p+q}$$

liefert eine Algebra $E = \sum_{p \geq 0} E^p$ über dem Vektorraum T . Man bezeichnet diese als die **Graßmann-Algebra** über T .

Satz 2.21.

(a) Für $\varphi, \psi \in E^p$ und $\chi \in E^q$ gilt

$$(\varphi + \psi) \wedge \chi = \varphi \wedge \chi + \psi \wedge \chi.$$

(b) Für $\varphi, \psi \in E^p$ und $\chi \in E^q$ ist

$$\chi \wedge (\varphi + \psi) = \chi \wedge \varphi + \chi \wedge \psi.$$

(c) Für $\varphi \in E^p, \psi \in E^q, \chi \in E^r$ gilt

$$\varphi \wedge (\psi \wedge \chi) = (\varphi \wedge \psi) \wedge \chi.$$

(d) Für $a \in \mathbb{R}, \varphi \in E^p, \psi \in E^q$ ist

$$a(\varphi \wedge \psi) = (a\varphi) \wedge \psi = \varphi \wedge (a\psi).$$

Bemerkung. Die äußere Multiplikation ist im allgemeinen nicht kommutativ, und für 1-Formen φ, ψ gilt

$$\varphi \wedge \psi = -\psi \wedge \varphi.$$

2.4 Tangentialvektoren, Derivationen und Pfaffsche Formen

Sei $x_0 \in \mathbb{R}^n$. Weiter sei D_{x_0} die Menge der in einer Umgebung $U(x_0)$ von x_0 definierten Funktionen f , die in x_0 differenzierbar sind.

Bemerkung. Ist $g \in D_{x_0}$ und f stetig in x_0 , und gilt $g(x_0) = 0$, so gehört $g \cdot f$ stets zu D_{x_0} .

Definition 2.22. Ein **Tangentialvektor** (oder eine **Derivation**) in x_0 ist eine lineare Abbildung $D : D_{x_0} \rightarrow \mathbb{R}$ mit folgenden Eigenschaften:

(i) $D(1) = 0$.

(ii) $D(gf) = 0$, falls f stetig in x_0 ist, $g \in D_{x_0}$ und $f(x_0) = g(x_0) = 0$ gilt.

Lemma 2.23. Sind f, g in D_{x_0} und ist D eine Derivation in x_0 , so gilt

$$D(fg) = f(x_0) \cdot D(g) + D(f) \cdot g(x_0).$$

Bemerkung. Die Menge der Derivationen in x_0 bildet einen Vektorraum.

Definition 2.24. Der Vektorraum der Derivationen in x_0 heißt **Tangentenraum** (des \mathbb{R}^n) in x_0 und wird mit T_{x_0} bezeichnet. Statt von Derivationen in x_0 sprechen wir auch von **Tangentialvektoren** in x_0 .

Satz 2.25. Die Derivationen $\frac{\partial}{\partial x_\nu}$ in x_0 für $\nu = 1, \dots, n$ bilden eine Basis des Vektorraums T_{x_0} .

Motivation für den Namen „Tangentialvektor“:

Wir nehmen an, $\varphi : I \rightarrow \mathbb{R}^n$ sei eine glatte Kurve, und es gelte $x_0 = \varphi(t_0)$. Dann ist die Tangente in x_0 gegeben durch

$$x = x_0 + (t - t_0)\varphi'(t_0).$$

Der Richtungsvektor dieser Tangente ist

$$\varphi'(t_0) = (\varphi'_1(t_0), \dots, \varphi'_n(t_0)).$$

Ist f in einer Umgebung von x_0 definiert und in x_0 differenzierbar, so ist $f \circ \varphi$ in t_0 differenzierbar (siehe Kettenregel) und es gilt

$$(f \circ \varphi)'(t_0) = \sum_{\nu=1}^n \varphi'_\nu(t_0) \cdot \frac{\partial f}{\partial x_\nu}(\varphi(t_0)).$$

Die Derivation

$$D = \sum_{\nu=1}^n \varphi'_\nu(t_0) \frac{\partial}{\partial x_\nu}$$

ist die Richtungsableitung von f in der Richtung φ .

Jeder parametrisierten Kurve φ entspricht so in natürlicher Weise ein Tangentialvektor in x_0 . Man kann auf diese Weise auch jeden Tangentialvektor

$$D = \sum_{\nu=1}^n a_\nu \frac{\partial}{\partial x_\nu}$$

in x_0 erhalten. Hierzu nimmt man die Strecke durch x_0 mit dem Richtungsvektor (a_1, \dots, a_n) .

Zur Transformation von Tangentialvektoren:

Sei $M \subseteq \mathbb{R}^n$ offen und $x_0 \in M$. Weiter sei $F : M \rightarrow \mathbb{R}^m$ eine Abbildung mit $F(x_0) = y_0$. Ist f in x_0 differenzierbar, so induziert F einen Vektorraumhomomorphismus

$$F_* : T_{x_0} \rightarrow T_{y_0}$$

wie folgt: Sei $f \in D_{y_0}$ (definiert auf einer Umgebung $U(y_0)$). Dann gibt es eine Umgebung $U(x_0) \subseteq M$ mit $F(U(x_0)) \subseteq U(y_0)$. Dann ist $f \circ F : U(x_0) \rightarrow \mathbb{R}$ wohldefiniert, und wir setzen

$$(F_*D)(f) := D(f \circ F)$$

für alle $f \in D_{y_0}$ und für $D \in T_{x_0}$.

Satz 2.26. Für $D \in T_{x_0}$ ist F_*D eine Derivation in y_0 , und F_* ist Vektorraumhomomorphismus von T_{x_0} nach T_{y_0} .

Satz 2.27. Sei $F : M \rightarrow \mathbb{R}^m$, $G : N \rightarrow \mathbb{R}^l$ mit $x_0 \in M^\circ$ und $y_0 \in N^\circ$. Seien F, G in x_0 bzw. y_0 differenzierbar. Dann gilt

$$(G \circ F)_* = G_* \circ F_*$$

Bemerkung. Wählt man Basen

$$\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n} \quad \text{von } T_{x_0}$$

und

$$\frac{\partial}{\partial y_1}, \dots, \frac{\partial}{\partial y_m} \quad \text{von } T_{y_0},$$

so wird F_* beschrieben durch die Funktionalmatrix

$$\begin{pmatrix} f_{1x_1}(x_0) & \cdots & f_{1x_n}(x_0) \\ \vdots & & \vdots \\ f_{mx_1}(x_0) & \cdots & f_{mx_n}(x_0) \end{pmatrix},$$

wobei F der Spaltenvektor $(f_1, \dots, f_m)^{tr}$ ist (siehe Kettenregel).

Differentiale und kovariante Tangentialvektoren

Sei $f : U(x_0) \rightarrow \mathbb{R}$ differenzierbar in x_0 . Wir wollen nun ein „Differential“ $df : T_{x_0} \rightarrow \mathbb{R}$ definieren.

Definition 2.28. $T_{x_0}^* = \text{Hom}_{\mathbb{R}}(T_{x_0}, \mathbb{R})$ heißt **kovarianter Tangentialraum**. Ein Element $\omega \in T_{x_0}^*$ heißt **Kovektor** in x_0 . Das **Differential** $df \in T_{x_0}^*$ einer differenzierbaren Funktion f in x_0 wird durch

$$df \left(\sum_{\nu=1}^n a_{\nu} \frac{\partial}{\partial x_{\nu}} \right) := \sum_{\nu=1}^n a_{\nu} \frac{\partial f}{\partial x_{\nu}}(x_0)$$

definiert, d.h. $df(D) := D(f)$.

Allgemeiner setzt man (mit $F : M \rightarrow \mathbb{R}^m$, $M = M(x_0)$ Umgebung von x_0 , $F(x_0) = y_0$):

$$F_* : T_{x_0} \rightarrow T_{y_0}$$

und

$$F^* : T_{y_0}^* \rightarrow T_{x_0}^*$$

sowie

$$(F^*\omega)(D) := \omega(F_*D),$$

d.h. $F^*(\omega) := \omega \circ F_*$.

Bemerkung.

- (i) Man kann zeigen: Für Funktionen $F : M \rightarrow N$, $G : N \rightarrow \mathbb{R}^l$ mit $F(x_0) = y_0$, $G(y_0) = z_0$ gilt

$$(G \circ F)^* = F^* \circ G^*.$$

- (ii) Mit den Bezeichnungen von (2.28) ist

$$dx_1, \dots, dx_n$$

die zu $\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n}$ duale Basis von $T_{x_0}^*$; d.h.

$$dx_i \left(\frac{\partial}{\partial x_k} \right) = \delta_{ik}.$$

Ist in jedem Punkt einer offenen Menge $M \subseteq \mathbb{R}^n$ ein Differential $\omega = \omega(y)$ gegeben, so kann man schreiben

$$\omega = \sum_{\nu=1}^n f_{\nu}(x) dx_{\nu}.$$

Solche Ausdrücke heißen **Pfaffsche Formen** oder **Differentialformen**.

2.5 Alternierende Differentialformen

Definition 2.29. Die alternierenden p -Linearformen auf dem Tangentialraum T_{x_0} heißen **p -dimensionale alternierende (oder auch äußere) Differentialformen** im Punkt x_0 (oder kurz **äußere p -Formen**). Eine **alternierende Differentialform** in x_0 ist ein Element der Grassmannalgebra $E = \sum_{p \geq 0} E^p$ über T_0 .

Insbesondere ist $E^1 = T_{x_0}^*$ der Vektorraum der Pfaffschen Formen in x_0 . E^1 wird aufgespannt von den Formen dx_ν ($\nu = 1, \dots, n$) mit

$$dx_\nu(\xi) = \xi(x_\nu).$$

Zur Bestimmung der Dimension der Räume E^p ein

Hilfssatz 2.30.

- (a) Für $p > n$ ist jede p -Differentialform φ Null.
- (b) Wenn eine äußere p -Form φ auf allen p -Tupeln der Gestalt

$$\left(\frac{\partial}{\partial x_{i_1}}, \dots, \frac{\partial}{\partial x_{i_p}} \right)_{1 \leq i_1 < \dots < i_p \leq n}$$

verschwindet, so ist $\varphi = 0$.

Hilfssatz 2.31. Für das äußere Produkt aus $dx_{i_1}, \dots, dx_{i_p}$ gilt:

$$(dx_{i_1} \wedge \dots \wedge dx_{i_p}) \left(\frac{\partial}{\partial x_{j_1}}, \dots, \frac{\partial}{\partial x_{j_p}} \right) = \begin{cases} 0 & \text{falls } \{j_1, \dots, j_p\} \neq \{i_1, \dots, i_p\} \\ \text{sgn} \begin{pmatrix} i_1 \dots i_p \\ j_1 \dots j_p \end{pmatrix} & \text{sonst} \end{cases}$$

wobei sgn das Signum der gegebenen Permutation bezeichnet, also $+1$, wenn j_1, \dots, j_p durch eine gerade Zahl von Vertauschungen aus i_1, \dots, i_p hervorgeht, und -1 sonst.

Satz 2.32. Für jedes $p \geq 0$ ist $\dim E^p = \binom{n}{p}$ und damit $\dim E = \sum_{p=0}^n \dim E^p = 2^n$. Die speziellen p -Formen

$$dx_{i_1} \wedge \dots \wedge dx_{i_p}, \quad 1 \leq i_1 < i_2 < \dots < i_p \leq n$$

bilden für $p \geq 1$ eine Basis von E^p ; d.h. jede p -Form φ läßt sich eindeutig in der Form

$$\varphi = \sum_{1 \leq i_1 < \dots < i_p \leq n} a_{i_1 \dots i_p} dx_{i_1} \wedge \dots \wedge dx_{i_p}$$

darstellen. Die Darstellung von φ in dieser Form heißt **Normalform** oder **Grundform** von φ .

Es ist stets $dx_i \wedge dx_l = -dx_l \wedge dx_i$ und insbesondere $dx_i \wedge dx_i = 0$. Ferner gilt

$$\begin{aligned} (dx_i \wedge dx_l) \left(\frac{\partial}{\partial x_j}, \frac{\partial}{\partial x_k} \right) &= dx_i dx_l \left(\frac{\partial}{\partial x_j}, \frac{\partial}{\partial x_k} \right) - dx_l dx_i \left(\frac{\partial}{\partial x_k}, \frac{\partial}{\partial x_j} \right) \\ &= \frac{\partial x_i}{\partial x_j} \frac{\partial x_l}{\partial x_k} - \frac{\partial x_l}{\partial x_k} \frac{\partial x_i}{\partial x_j}. \end{aligned}$$

Hiermit kann man nun zeigen:

Satz 2.33. Für alle $\varphi \in E^p$ und $\psi \in E^q$ gilt

$$\varphi \wedge \psi = (-1)^{pq} \psi \wedge \varphi.$$

Folgerung 2.34. Ist p ungerade und $\varphi \in E^p$, so gilt $\varphi \wedge \varphi = 0$.

Wir betrachten nun Produkte von n Pfaffschen Formen:

$$\begin{aligned} \varphi_1 \wedge \dots \wedge \varphi_n &= \sum a_{1\nu_1} dx_{\nu_1} \wedge \dots \wedge \sum a_{n\nu_n} dx_{\nu_n} = \\ &= \sum a_{1\nu_1} \dots a_{n\nu_n} \operatorname{sgn} \begin{pmatrix} 1 \dots n \\ \nu_1 \dots \nu_n \end{pmatrix} dx_1 \wedge \dots \wedge dx_n = \\ &= \det(a_{\mu\nu}) dx_1 \wedge \dots \wedge dx_n. \end{aligned}$$

Also ist $\varphi_1 \wedge \dots \wedge \varphi_n$ genau dann Null, wenn die φ_ν linear abhängig sind.

Sei nun $U \subseteq \mathbb{R}^n$ offen, $F : U \rightarrow \mathbb{R}^m$ differenzierbar und $x_0 \in U$ sowie $F(x_0) = y_0$. Dann induziert F eine Abbildung $F_* : T_{x_0} \rightarrow T_{y_0}$. Es gilt

$$F_*(D)(g) = D(g \circ F), \quad D \in T_{x_0}, \quad g \in D_{y_0}.$$

Ist nun φ eine p -Form in y_0 (nicht notwendig alternierend) und $p \geq 1$, so setzen wir

$$(\varphi \circ F)(\xi_1, \dots, \xi_p) := \varphi(F_*\xi_1, \dots, F_*\xi_p), \quad \xi_i \in T_{x_0},$$

und für den Fall, dass φ eine Nullform ist, $\varphi \circ F := \varphi$. In jedem Fall ist $\varphi \circ F$ wieder eine p -Form, und mit φ ist auch $\varphi \circ F$ alternierend.

Sei $V = V(y_0) \subseteq \mathbb{R}^m$ offen und $G : V \rightarrow \mathbb{R}^k$. Dann gilt für p -Formen φ in $y_0 = G(y_0)$

$$\varphi \circ (G \circ F) = (\varphi \circ G) \circ F,$$

wie aus $(G \circ F)_* = G_* \circ F_*$ folgt.

Allgemeiner sei $\varphi \in E_{y_0}$ aus der Graßmannalgebra über T_{y_0} . Ist $\varphi = \sum_{p \geq 0} \varphi_p$, so setzt man

$$\varphi \circ F = \sum_{p \geq 0} \varphi_p \circ F.$$

Satz 2.35. Die Abbildung $\varphi \mapsto \varphi \circ F$ ist ein Homomorphismus von E_{y_0} nach E_{x_0} , der für jedes p den Raum $E_{y_0}^p$ in $E_{x_0}^p$ abbildet.

Satz 2.36. Ist $\varphi = \sum a_{i_1 \dots i_p} dy_{i_1} \wedge \dots \wedge dy_{i_p} \in E_{y_0}^p$ und $F = (f_1, \dots, f_m)$, so gilt

$$(\varphi \circ F) = \sum a_{i_1 \dots i_p} df_{i_1} \wedge \dots \wedge df_{i_p}.$$

Sei nun $M \subseteq \mathbb{R}^n$ offen und T_x, E_x^p, E_x usw. wie oben.

Definition 2.37. Eine p -Differentialform (bzw. Differentialform) auf M ist eine Vorschrift φ , die jedem $x \in M$ genau ein Element $\varphi(x) \in E_x^p$ (bzw. $\in E_x$) zuordnet, d.h.

$$\varphi(x) = \sum a_{i_1 \dots i_p}(x) dx_{i_1} \wedge \dots \wedge dx_{i_p}.$$

φ heißt in $x_0 \in M$ **stetig**, bzw. **differenzierbar** oder **k -mal stetig differenzierbar**, falls sämtliche Koeffizientenfunktionen $a_{i_1 \dots i_p}(x)$ die entsprechende Eigenschaft haben.

Die unendlich oft differenzierbaren Formen auf einer offenen Menge bilden einen \mathbb{R} -Vektorraum, bezeichnet mit \mathcal{E}^p ($p \geq 0$) bzw. eine \mathbb{R} -Algebra $\mathcal{E} = \mathcal{E}^0 + \mathcal{E}^1 + \dots + \mathcal{E}^n$. Wir wollen auf \mathcal{E} einen Differentialoperator d („äußere Ableitung“) einführen.

Hierzu sei f in x_0 differenzierbar, $D \in T_{x_0}$. Durch $(df)D = D(f)$ ist dann eine Pfaffsche Form $df \in E_{x_0}^1$ erklärt (das „totale Differential“ von f). Es gilt dann die Leibnizsche Regel

$$d(fg) = f(x_0)dg + g(x_0)df$$

für $f, g \in D_{x_0}$. Es ist

$$df = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x_0)dx_i.$$

Ist f in jedem Punkt differenzierbar, so ist df eine wohldefinierte Differentialform auf M .

Satz 2.38. Sei $U \subseteq \mathbb{R}^n$ offen, $\mathcal{E} = \mathcal{E}^0 + \dots + \mathcal{E}^n$ die Graßmannalgebra der unendlich oft differenzierbaren Differentialformen auf U . Dann gibt es genau eine lineare Abbildung $d : \mathcal{E} \rightarrow \mathcal{E}$ mit folgenden Eigenschaften:

(1) Für $f \in \mathcal{E}^0$ ist df das totale Differential von f .

(2) Für $\varphi \in \mathcal{E}^p$, $\psi \in \mathcal{E}^q$ ist

$$d(\varphi \wedge \psi) = d\varphi \wedge \psi + (-1)^p \varphi \wedge d\psi.$$

(3) $ddF = 0$ für $f \in \mathcal{E}^0$.

Hierbei ist für $\varphi = \sum a_{i_1 \dots i_p} dx_{i_1} \wedge \dots \wedge dx_{i_p}$ stets $d\varphi = \sum da_{i_1 \dots i_p} dx_{i_1} \wedge \dots \wedge dx_{i_p}$.

An Stelle von (3) gilt allgemeiner:

Satz 2.39. Sei $U \subseteq \mathbb{R}^n$ offen und φ eine auf $U \subseteq \mathbb{R}^n$ definierte differenzierbare Differentialform, die in $x_0 \in U$ zweimal differenzierbar ist. Dann gilt $dd\varphi = 0$.

Definition 2.40. Eine (in x_0) differenzierbare Differentialform φ heißt (**in x_0**) **geschlossen**, wenn $d\varphi = 0$ (bzw. $(d\varphi)_{x_0} = 0$) gilt. Die Form φ heißt (**in x_0**) **exakt**, wenn es eine (in x_0) differenzierbare Differentialform ψ mit $d\psi = \varphi$ gibt (bzw. $(d\psi)_{x_0} = \varphi_{x_0}$).

Folgerung 2.41. *Ist φ auf der offenen Menge $U \subseteq \mathbb{R}^n$ einmal differenzierbar und exakt, so ist φ geschlossen.*

Für das Verhalten von Differentialformen und äußerer Ableitung bei differenzierbaren Abbildungen erhält man leicht den

Satz 2.42. *Seien $M \subseteq \mathbb{R}^n$ und $N \subseteq \mathbb{R}^m$ offen und sei φ in $y_0 \in N$ differenzierbar. Sei F in x_0 mit $y_0 = F(x_0)$ zweimal differenzierbar. Dann gilt*

$$d(\varphi \circ F) = d\varphi \circ F.$$

2.6 Der Stokessche Integralsatz

Der Integralsatz, den wir in diesem Paragraphen behandeln, stellt eine Beziehung zwischen einem Flächenintegral und einem Integral über den Rand ∂F des gegebenen Flächenstücks her: Unter gewissen Voraussetzungen gilt

$$\int_F d\omega = \int_{\partial F} \omega. \quad (2.5)$$

Dazu ist zunächst für $p+1$ -dimensionale Flächenstücke F und p -dimensionale Differentialformen ω

$$\int_F d\omega \text{ bzw. } \int_{\partial F} \omega$$

zu erklären.

Definition 2.43. *Sei $F \subseteq \mathbb{R}^n$ ein p -dimensionales Flächenstück mit Parameterdarstellung $(A, x(t))$. Für $\omega = f dg_1 \wedge \dots \wedge dg_p$ setzen wir*

$$\int_F f dg_1 \wedge \dots \wedge dg_p = \int_A f(x(t)) \det \left(\frac{\partial g_i(x(t))}{\partial t_j} \right) dt_1 \dots dt_p. \quad (2.6)$$

Ist speziell $F = A$, d.h. F eine offene Teilmenge des \mathbb{R}^p sowie $g_i = t_i = x_i$, so haben wir dann

$$\int_A f dt_1 \wedge \dots \wedge dt_p = \int_A f(t) dt_1 \dots dt_p. \quad (2.7)$$

Bemerkung: (2.6) garantiert gerade, dass die Unabhängigkeit von der speziellen Wahl der Parameterdarstellung vorliegt.

Zu (2.7): Man beachte, dass hier keine Unabhängigkeit von der Reihenfolge der dt_i vorliegt. Falls nämlich π eine ungerade Permutation ist, geht das Integral auf der linken Seite in sein Negatives über (denn ist $g_i = t_{\pi(i)}$, so $\frac{\partial g_i}{\partial t_j} = \text{sgn}(\pi)$). Daher kann

$$f dt_1 \wedge \dots \wedge dt_p$$

nur über ein orientiertes Flächenstück integriert werden: A ändert bei ungerader Permutation π seine Orientierung.

Zum Stokesschen Satz:

Stokes selbst hat den Satz nur für zweidimensionale Flächen im \mathbb{R}^3 betrachtet; wir wollen beliebige Dimensionen $p < n$ zulassen.

Voraussetzungen 2.44 (für (2.5)). *Sei F eine orientierte $(p+1)$ -dimensionale glatte Fläche im \mathbb{R}^n , ω eine p -Form und stetig differenzierbar in einer offenen Menge $B \subseteq \mathbb{R}^n$, welche $F \cup \partial F$ umfasst. Ferner sollen beide Integrale $\int_F d\omega$ und $\int_{\partial F} \omega$ existieren, und weiter ist dafür zu sorgen, dass „schlechte“ Randpunkte nicht gefährlich werden können. Dies kann auf zwei Weisen geschehen:*

- a) *Der Träger von ω , also die abgeschlossene Hülle der Punkte x mit $\omega(x) \neq 0$, kann als beschränkt vorausgesetzt werden.*
- b) *$\overline{\partial F} \setminus \partial F$ ist einzuschränken.*

So erhält man folgende erste Version des Stokesschen Satzes:

Satz 2.45. *Zusätzlich zu (2.44) setzen wir voraus, dass es eine kompakte Menge $K \subseteq \mathbb{R}^n$ gibt, so dass $K \cap (\overline{\partial F} \setminus \partial F) = \emptyset$ und $\omega(x) = 0$ für $x \notin K$ gilt. Ferner besitze F lokale Parameterdarstellungen, die zweimal stetig differenzierbar sind, auch in den guten Randpunkten von F . Dann gilt*

$$\int_F d\omega = \int_{\partial F} \omega.$$

Man zeigt Satz (2.45) zunächst für den Fall, dass F ein Quader im \mathbb{R}^{p+1} ist. Für beliebige Flächen wird benötigt:

Hilfssatz 2.46. *Sei $K \subseteq \mathbb{R}^n$ kompakt. Seien $x^{(i)} \in K$ endlich viele Punkte, und es gebe Zahlen r_i, r'_i mit $0 < r_i < r'_i$ so dass $K \subseteq \cup_i K(x^{(i)}, r_i)$ gilt. Dann existieren Funktionen $\chi_i(x)$, für die gilt:*

- (i) χ_i sind auf dem \mathbb{R}^n unendlich oft differenzierbar;
- (ii) $0 < \chi_i(x) < 1$ für alle x ;
- (iii) $\chi_i(x) = 0$ für $x \notin K(x^{(i)}, r_i)$;
- (iv) $\sum_i \chi_i(x) = 1$ für alle $x \in K$.

Mit anderen Worten: Im \mathbb{R}^n existieren passend zu endlichen Überdeckungen von K Zerlegungen der Einheit mit beliebig oft differenzierbaren Funktionen χ_i .

Mit diesem Hilfssatz wird in der Vorlesung die erste Version des Stokesschen Satzes (Formel 2.5) bewiesen. Wir schwächen nun die Voraussetzungen von (2.5) wie folgt ab: Der Träger K von ω darf von nun an auch schlechte Randpunkte enthalten, aber

$$(\bar{\partial}F \setminus \partial F) \cap K$$

darf nicht zu groß sein. Genauer gilt:

Satz 2.47. *Die Formel (2.5) gilt unter folgenden Voraussetzungen: Zusätzlich zu (2.44) sei ω auf ganz \bar{F} stetig und sei $K \subseteq \mathbb{R}^n$ eine kompakte Menge mit $\omega(x) = 0$ für $x \notin K$. Es gebe zu jedem $x \in \bar{F} \cap K$ eine Umgebung U , eine offene Menge $A \subseteq \mathbb{R}^{p+1}$, Teilmengen $A' \subseteq \partial A$ und $A'' \subseteq \bar{A} \setminus (A \cup A')$ und eine auf einer \bar{A} umfassenden offenen Menge definierte und zweimal stetig differenzierbare Abbildung h in den \mathbb{R}^n mit folgenden Eigenschaften:*

- (i) $h(A) = F \cap U$, $h(A') = \partial F \cap U$, und in jedem Punkt von $A \cup A'$ hat die Funktionalmatrix von h den Rang $p + 1$;
- (ii) $h(A'') \supseteq (\bar{\partial}F \setminus \partial F) \cap K \cap U$ und A'' ist Bild eines kompakten Teils des \mathbb{R}^{p-1} unter einer stetig differenzierbaren Abbildung g .

Anmerkung: Die Bedingung (i) entspricht den Differenzierbarkeitsvoraussetzungen von (2.45). (ii) ersetzt die Forderung $(\bar{\partial}F \setminus \partial F) \cap K = \emptyset$.

2.7 Das Poincarésche Lemma

Eine p -Form φ heißt **exakt**, wenn es eine $(p-1)$ -Form ω mit $\varphi = d\omega$ gibt. In diesem Fall ist φ geschlossen, d.h. $d\varphi = 0$ (siehe Kapitel 2.5, Satz (2.39)). Es stellt sich die Frage, in welchen Fällen umgekehrt jede exakte Form geschlossen ist. Antwort: Dies hängt vom Definitionsbereich von φ ab, lokal gilt die Umkehrung.

Definition 2.48. Ein offener Bereich $U \subseteq \mathbb{R}^n$ heißt **zusammenziehbar auf den Punkt** $x^{(0)} \in U$, wenn es eine zweimal stetig differenzierbare Abbildung h des topologischen Produktes $U \times I$ von U mit dem Einheitsintervall $I = [0, 1]$

$$U \times I = \{(x_1, \dots, x_n, t) \in \mathbb{R}^{n+1} : (x_1, \dots, x_n) \in U, t \in [0, 1]\}$$

auf U gibt, so dass

$$h(x, 1) = x \quad \text{und} \quad h(x, 0) = x^{(0)}$$

für alle $x \in U$ gilt. A heißt **bogenweise zusammenhängend**, wenn sich je zwei Punkte aus A mit einer glatten Kurve verbinden lassen.

Bemerkung. Offensichtlich ist eine nicht bogenweise zusammenhängende Menge auch nicht zusammenziehbar.

Definition 2.49. Unter einem **Gebiet** verstehen wir eine offene bogenweise zusammenhängende Punktmenge.

Bemerkung. Nicht jedes Gebiet ist zusammenziehbar; Kreisringe sind beispielsweise nicht zusammenziehbar.

Lemma 2.50 (Poincarésches Lemma). Sei U ein auf einen Punkt zusammenziehbares Gebiet des \mathbb{R}^n und ψ eine auf U stetig differenzierbare geschlossene p -Form ($0 < p \leq n$). Dann existiert eine auf U stetig differenzierbare $(p-1)$ -Form ω mit $\psi = d\omega$.

Bemerkung.

- (i) Ist U nicht zusammenziehbar, so gibt es nach diesem Lemma wenigstens zu jeder ganz in U gelegenen Kugel eine $(p-1)$ -Form ω , so dass in dieser Kugel $\psi = d\omega$ gilt. Jede geschlossene Form ist also lokal exakt.
- (ii) Existiert eine Lösung ω von $\psi = d\omega$, so ist diese nicht eindeutig bestimmt. So ist etwa für $p = 1$ mit ω auch $\omega + \text{const}$ stets eine Lösung. Entsprechendes gilt für $p > 1$: mit ω ist dann auch $\omega + d\chi$ Lösung, wenn χ eine zweimal stetig differenzierbare $(p-2)$ -Form ist. Umgekehrt sind dadurch auch sämtliche Lösungen gegeben.

2.8 Vektoranalysis im dreidimensionalen Raum

Wir wenden nun die Sätze der letzten zwei Paragraphen an, um klassische Ergebnisse für den Fall des \mathbb{R}^3 zu erhalten. Hierzu betrachten wir 0-, 1-, 2- und 3-dimensionale Differentialformen.

Sei zunächst f eine 0-Form, also eine Funktion. Dann gilt

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz,$$

wenn $(x, y, z) = \mathfrak{s}$ den Ortsvektor im \mathbb{R}^3 bezeichnet. Ist weiter $d\mathfrak{s} = (dx, dy, dz)$ das vektorielle Differential, so können wir df als formales Skalarprodukt

$$df = \text{grad } f \cdot d\mathfrak{s} \tag{2.8}$$

schreiben.

Sei nun φ eine 1-Form:

$$\varphi = a_1 dx + a_2 dy + a_3 dz = \mathbf{a} \cdot d\mathfrak{s}.$$

Dann gilt

$$d\varphi = \left(\frac{\partial a_3}{\partial y} - \frac{\partial a_2}{\partial z} \right) dy \wedge dz + \left(\frac{\partial a_1}{\partial z} - \frac{\partial a_3}{\partial x} \right) dz \wedge dx + \left(\frac{\partial a_2}{\partial x} - \frac{\partial a_1}{\partial y} \right) dx \wedge dy.$$

Der Vektor

$$\left(\frac{\partial a_3}{\partial y} - \frac{\partial a_2}{\partial z}, \frac{\partial a_1}{\partial z} - \frac{\partial a_3}{\partial x}, \frac{\partial a_2}{\partial x} - \frac{\partial a_1}{\partial y} \right) =: \text{rot } \mathbf{a}$$

heißt die **Rotation** des Vektorfeldes $\mathbf{a} = (a_1, a_2, a_3)$. Setzen wir nun $d\mathcal{F} = (dy \wedge dz, dz \wedge dx, dx \wedge dy)$, so ergibt sich

$$d\varphi = \text{rot } \mathbf{a} \cdot d\mathcal{F}. \quad (2.9)$$

Entsprechend gehen wir für 2-Formen $\varphi = a_1 dy \wedge dz + a_2 dz \wedge dx + a_3 dx \wedge dy = \mathbf{a} \cdot d\mathcal{F}$ vor. Es ist

$$d\varphi = \left(\frac{\partial a_1}{\partial x} + \frac{\partial a_2}{\partial y} + \frac{\partial a_3}{\partial z} \right) dx \wedge dy \wedge dz = \text{div } \mathbf{a} dV, \quad (2.10)$$

wobei $dV = dx \wedge dy \wedge dz$ als orientiertes Volumenelement aufgefasst wird. Der Skalar

$$\frac{\partial a_1}{\partial x} + \frac{\partial a_2}{\partial y} + \frac{\partial a_3}{\partial z} =: \text{div } \mathbf{a}$$

heißt die **Divergenz** von \mathbf{a} .

Die Differentialoperatoren grad (Gradient), div (Divergenz) und rot (Rotation) können alle mit demselben Differentialoperator

$$\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right),$$

dem sogenannten **Nabla-Operator**, geschrieben werden: Es ist

$$\text{grad } f = \nabla f, \quad \text{div } \mathbf{a} = \nabla \cdot \mathbf{a}, \quad \text{rot } \mathbf{a} = \nabla \times \mathbf{a},$$

wobei $\nabla \times \mathbf{a}$ das Vektorprodukt zweier Vektoren im \mathbb{R}^3 ist. Das Skalarprodukt des Nabla-Operators mit sich selbst ergibt den **Laplace-Operator** Δ :

$$\nabla \cdot \nabla f = \Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} = \text{div grad } f.$$

Die Regeln der Vektorrechnung lassen sich übertragen. So liefert die Rechenregel $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c}) \cdot \mathbf{b} - (\mathbf{a} \cdot \mathbf{b}) \cdot \mathbf{c}$ die Formel

$$\nabla \times (\nabla \times \mathbf{c}) = \nabla(\nabla \cdot \mathbf{c}) - (\nabla \cdot \nabla)\mathbf{c},$$

anders geschrieben

$$\text{rot rot } \mathbf{c} = \text{grad div } \mathbf{c} - \Delta \mathbf{c}. \quad (2.11)$$

Indem wir φ als C^2 -Differentialform voraussetzen, erhalten wir für Nullformen (Funktionen) f mittels $ddf = 0$:

$$\text{rot grad } f = 0 \quad (2.12)$$

und für 1-Formen $\varphi = \mathbf{a} \cdot d\mathbf{s}$

$$\operatorname{div} \operatorname{rot} \mathbf{a} = 0. \quad (2.13)$$

Weitere Aussagen erhält man durch Verwendung eines Spezialfalls des Poincaré'schen Lemmas, nämlich des Falls der 1-Form $\psi = \mathbf{a} \cdot d\mathbf{s}$. Genau dann ist $d\psi = 0$, wenn $\operatorname{rot} \mathbf{a} = 0$ gilt, und wir haben damit das Ergebnis

Folgerung 2.51. *Gilt für eine Vektorfunktion \mathbf{a} in einem zusammenziehbaren Gebiet G $\operatorname{rot} \mathbf{a} = 0$, so gibt es eine in G stetig differenzierbare Funktion f mit $\operatorname{grad} f = \mathbf{a}$.*

Dies besitzt eine wichtige physikalische Anwendung: Ist \mathbf{a} ein Kraftfeld, das in der Form $\mathbf{a} = \operatorname{grad} f$ darstellbar ist, so nennt man \mathbf{a} eine **Potentialkraft** und f ein **Potential** von \mathbf{a} . (2.51) besagt, dass jedes rotationsfreie (wirbelfreie) Kraftfeld ein Potential besitzt.

Für 2-Formen $\psi = \mathbf{a} \cdot d\mathcal{F}$ gilt entsprechend $d\psi = 0$ genau dann, wenn $\operatorname{div} \mathbf{a} = 0$ ist. So erhält man

Folgerung 2.52. *Gilt $\operatorname{div} \mathbf{a} = 0$ in einem zusammenziehbaren Gebiet G , so gibt es eine Vektorfunktion \mathbf{b} derart, dass $\mathbf{a} = \operatorname{rot} \mathbf{b}$ gilt.*

Wir wenden nun den allgemeinen Satz von Stokes (2.47) auf zweidimensionale Flächen an, danach auf dreidimensionale Flächen (d.h. offene Bereiche des Raumes). Dazu betrachten wir zunächst 1-Formen $\varphi = \mathbf{a}(\mathbf{s}) \cdot d\mathbf{s}$. Aus (2.47) und (2.9) folgt dann

Satz 2.53 (Klassischer Satz von Stokes). *F sei eine beschränkte orientierte zweidimensionale glatte Fläche, und $\bar{\partial}F \setminus \partial F$ bestehe aus endlich vielen Punkten ($\bar{\partial}F$ ist also eine stückweise glatte, geschlossene Kurve). Dann gilt für ein auf \bar{F} stetiges, auf $\partial F \cup F$ stetig differenzierbares Vektorfeld $\mathbf{a}(\mathbf{s})$*

$$\int_{\partial F} \mathbf{a}(\mathbf{s}) \cdot d\mathbf{s} = \int_F \operatorname{rot} \mathbf{a}(\mathbf{s}) \cdot d\mathcal{F}.$$

Bemerkung. $\bar{\partial}F$ braucht nicht zusammenhängend zu sein. Aber die Orientierung jedes glatten Kurvenstücks von $\bar{\partial}F$ muss gerade die durch die Fläche F induzierte Orientierung sein.

Bei einer dreidimensionalen Fläche ist eine 2-Form $\varphi = \mathbf{a} \cdot d\mathcal{F}$ zu betrachten, und es ergibt sich der

Satz 2.54 (Satz von Gauß (im Raum)). $G \subseteq \mathbb{R}^3$ sei offen und beschränkt, und es gelte die Voraussetzung (2.47) bzw. (2.45). Dann gilt für ein auf \overline{G} stetiges, auf $G \cup \partial G$ stetig differenzierbares Vektorfeld $\mathbf{a} = \mathbf{a}(\mathfrak{s})$

$$\int_{\partial G} \mathbf{a} \cdot d\mathcal{F} = \int_G \operatorname{div} \mathbf{a} \, dV.$$

(Vgl. 2.10.) In dem Spezialfall $\mathbf{a}(\mathfrak{s}) = g(\mathfrak{s}) \cdot \operatorname{grad} h(\mathfrak{s})$ ist $\operatorname{div} \mathbf{a} = \operatorname{grad} g \cdot \operatorname{grad} h + g \Delta h$. So ergibt sich die

Satz 2.55 (Erste Greensche Formel).

$$\int_{\partial G} g \operatorname{grad} h \cdot d\mathcal{F} = \int_G (\operatorname{grad} g \cdot \operatorname{grad} h + g \cdot \Delta h) \, dV.$$

Indem man nun zunächst g mit h vertauscht und dann beide Gleichungen subtrahiert, erhält man auch noch die

Satz 2.56 (Zweite Greensche Formel).

$$\int_{\partial G} (g \cdot \operatorname{grad} h - h \cdot \operatorname{grad} g) \cdot d\mathcal{F} = \int_G (g \Delta h - h \Delta g) \, dV.$$

3 Funktionenräume

3.1 Banachräume und Hilberträume

Definition 3.1. Ein Vektorraum V über \mathbb{R} bzw. \mathbb{C} heißt **normiert mit Norm** $\|\cdot\|$, wenn gilt:

- (1) $\|x\| \geq 0$ für alle $x \in V$, und aus $\|x\| = 0$ folgt $x = 0$;
- (2) $\|c \cdot x\| = |c| \cdot \|x\|$ für alle $x \in V$, $c \in \mathbb{R}$ bzw. \mathbb{C} ;
- (3) $\|x + y\| \leq \|x\| + \|y\|$ (die Dreiecksungleichung).

Als Beispiel normierter Vektorräume betrachten wir Funktionenräume: Sei $A \subseteq \mathbb{R}^n$ messbar mit Lebesgue-Maß $m(A) \neq 0$. Sei f eine über A integrierbare Funktion mit $\int_A |f(x)|dx < \infty$ und g mit denselben Eigenschaften. Wir setzen $\|f\|_1 := \int_A |f(x)|dx$ und wollen diese Abbildung auf einem geeigneten Vektorraum zu einer Norm machen. Zunächst gilt offenbar

$$\int_A |cf(x)|dx = |c| \cdot \int_A |f(x)|dx,$$

d.h. (2), und es gilt auch

$$\int_A |f(x) + g(x)|dx \leq \int_A |f(x)|dx + \int_A |g(x)|dx < \infty,$$

d.h. (3). Aber (1) gilt ohne weiteres nicht, denn wenn zwei verschiedene Funktionen f und g sich nur auf einer Nullmenge unterscheiden, ist $\int_A |f(x)|dx = \int_A |g(x)|dx$. Man löst dieses Problem, indem man folgende Äquivalenzrelation einführt:

$$f \sim g : \iff m\{x \in A : f(x) \neq g(x)\} = 0,$$

d.h. Äquivalenz von f und g besagt, dass f mit g fast überall in A übereinstimmt. Ist nun $\underline{f} := \{g : f \sim g\}$ die Äquivalenzklasse von f , so gilt der

Satz 3.2. Die Klassen über A integrierbarer Funktionen mit $\int_A |f(x)|dx < \infty$ bilden mit der Norm $\|\underline{f}\|_1 := \int_A |f(x)|dx < \infty$ einen normierten Vektorraum, der mit $L_1(A)$ bezeichnet wird.

Bemerkung. Es gilt $\dim(L_1(A)) = \infty$. Denn wird A in abzählbar viele disjunkte Teilmengen zerlegt, so sind deren charakteristische Funktionen linear unabhängig in $L_1(A)$.

Weitere Beispiele:

Sei $V = \mathbb{R}^n$, $(x_1, \dots, x_n) = x$. Die Abbildungen

$$\|x\|_1 := \sum_{i=1}^n |x_i|, \quad \|x\|_2 := \sqrt{\sum_{i=1}^n |x_i|^2}, \quad \|x\|_\infty := \max_{i=1, \dots, n} |x_i|$$

sind jeweils Normen.

Bemerkung.

- (a) Für jedes reelle $p \geq 1$ ist $\|x\|_p := (\sum_{i=1}^n |x_i|^p)^{1/p}$ eine Norm.
- (b) $\lim_{p \rightarrow \infty} \|x\|_p = \|x\|_\infty$, was die Bezeichnung $\|x\|_\infty$ rechtfertigt.

Im Fall a) ist die Dreiecksungleichung die sogenannte Minkowskische Ungleichung (siehe z.B. Analysis I, Forster, S. 114-116).

Definition 3.3. Ein normierter Vektorraum V heißt **vollständig**, wenn jede Cauchy-Folge

$$(x_m) \in V, \quad \|x_m - x_l\| \rightarrow 0 \quad \text{für } m, l \rightarrow \infty$$

in V einen Grenzwert besitzt, d.h. wenn ein $x \in V$ existiert mit $\|x_m - x\| \rightarrow 0$ für $m \rightarrow \infty$. Ein vollständiger normierter Vektorraum heißt **Banachraum**.

Bemerkung 3.4. $L_1(A)$ ist ein Banachraum.

Nach den normierten Vektorräumen betrachten wir nun Räume, die statt mit einer Norm sogar mit einem Skalarprodukt versehen sind, was eine stärkere Bedingung ist.

Definition 3.5. Ein Vektorraum V über \mathbb{R} bzw. \mathbb{C} heißt **Prä-Hilbertraum**, wenn es eine Abbildung

$$\varphi : \begin{cases} V \times V \rightarrow \mathbb{R} \text{ bzw. } \mathbb{C} \\ (x, y) \mapsto \langle x, y \rangle \end{cases}$$

gibt mit folgenden Eigenschaften:

(i) φ ist additiv in x und y .

(ii) $\langle cx, y \rangle = c\langle x, y \rangle$, $\langle x, cy \rangle = \bar{c}\langle x, y \rangle$.

(iii) $\langle x, y \rangle = \overline{\langle y, x \rangle}$.

(iv) $\langle x, x \rangle \geq 0$, und für $x \neq 0$ ist $\langle x, x \rangle > 0$.

Eine Abbildung φ mit diesen Eigenschaften heißt **Skalarprodukt** auf V .

Korollar 3.6. In diesem Fall ist V mit der Norm $\|x\| := \sqrt{\langle x, x \rangle}$ ein normierter Vektorraum.

Um dies zu zeigen, beweist man die Dreiecksungleichung und verwendet dazu die folgende **Cauchy-Schwarzsche Ungleichung**:

Satz 3.7. Sei V ein Vektorraum über \mathbb{C} mit Skalarprodukt $\langle x, y \rangle$. Dann gilt

$$|\langle x, y \rangle| \leq \sqrt{\langle x, x \rangle} \cdot \sqrt{\langle y, y \rangle}$$

für $x, y \in V$.

Zum Beweis betrachtet man ein Skalarprodukt der Form

$$\langle x - ty, x - ty \rangle = \langle x, x \rangle - t\langle y, x \rangle - \bar{t}\langle x, y \rangle + |t|^2\langle y, y \rangle \quad (3.1)$$

und setzt dabei speziell $t := \frac{\langle x, y \rangle}{\langle y, y \rangle}$.

Definition 3.8. Ein vollständiger Prä-Hilbertraum heißt **Hilbertraum**.

Beispiele für Hilberträume:

Für jedes $n > 0$ ist $V = \mathbb{K}^n$, mit $\mathbb{K} = \mathbb{R}$ oder \mathbb{C} , ein Hilbertraum mit dem Skalarprodukt

$$\langle x, y \rangle = \sum_{i=1}^n x_i \bar{y}_i.$$

Für messbare Mengen $A \subseteq \mathbb{R}^n$ mit Lebesgue-Maß $m(A) > 0$ ist

$$L_2(A) := \left\{ \underline{f} : \int_A |f(x)|^2 dx < \infty \right\}$$

ein Hilbertraum, wobei \underline{f} die oben eingeführte Äquivalenzklasse ist, und zwar mit dem Skalarprodukt

$$\langle \underline{f}, \underline{g} \rangle := \int_A f(x) \overline{g(x)} dx.$$

Zum Beweis der Vollständigkeit geht man ähnlich wie im Fall des $L_1(A)$ vor und verwendet dabei den

Hilfssatz 3.9. *Ist $\lim_{m,n \rightarrow \infty} \|f_m - f_n\|_2 = 0$, so existiert eine Funktion $f \in L_2(A)$ mit $\lim_{n \rightarrow \infty} \|f - f_n\|_2 = 0$.*

Wir führen nun folgende Funktionenräume ein:

$C(A) := \{f : f \text{ stetig und beschränkt auf } A\}$ und

$C_0(\mathbb{R}^n) := \{f \in C(\mathbb{R}^n) : \exists \text{ kompakte Menge } Tr_f \subseteq \mathbb{R}^n \text{ mit } f(x) = 0 \text{ für alle } x \notin Tr_f\}$,

d.h. $C_0(\mathbb{R}^n)$ ist der Vektorraum der stetigen Funktionen auf \mathbb{R}^n mit kompaktem Träger. (Man schreibt auch $C_c(\mathbb{R}^n)$, wobei der Index c die Kompaktheit des Trägers bezeichnet; der **Träger** einer Funktion ist, wie unter (2.44) angemerkt, die abgeschlossene Hülle der Punkte x mit $f(x) \neq 0$).

Es ist dann $C(A)$ ein normierter Vektorraum mit der Norm $\|f\| = \sup_{x \in A} |f(x)|$, sogar ein Banachraum. $C_0(\mathbb{R}^n)$ ist mit jeder der Normen

$$\|\cdot\|_1, \quad \|\cdot\|_2, \quad \|\cdot\|_\infty$$

ein normierter Vektorraum, allerdings mit $\|\cdot\|_1$ oder $\|\cdot\|_2$ kein Banachraum.

Hilfssatz 3.10. *Die Elementarfunktionen, also die Funktionen der Form $f = \sum_{i=1}^m e_i \chi_i$, wobei χ_i die charakteristische Funktion eines Quaders Q_i ist, liegen dicht in jedem der folgenden Räume:*

(a) $L_1(\mathbb{R}^n)$

(b) $L_2(\mathbb{R}^n)$

(c) $H := L_1(\mathbb{R}^n) \cap L_2(\mathbb{R}^n)$.

Sei nun weiter C_0^p der Vektorraum der p -mal stetig differenzierbaren Funktionen $f : \mathbb{R}^n \rightarrow \mathbb{R}$, die außerhalb einer beschränkten Menge verschwinden, und sei $C_0^0(\mathbb{R}^n) = C_0(\mathbb{R}^n)$. Es gilt dann der

Satz 3.11. Für $p = 0, 1, 2, \dots$ oder $p = \infty$ ist $C_0^p(\mathbb{R}^n)$ eine dichte Teilmenge von $L_1(\mathbb{R}^n)$ und $L_2(\mathbb{R}^n)$. Zu jedem $f \in L_1(\mathbb{R}^n) \cap L_2(\mathbb{R}^n)$ gibt es eine Folge (f_n) aus $C_0^p(\mathbb{R}^n)$, die sowohl in der L_1 -Norm als auch in der L_2 -Norm gegen f konvergiert.

Bemerkung. $C_0(\mathbb{R}^n)$ ist nicht vollständig bezüglich der Norm $\|\cdot\|_1$ bzw. $\|\cdot\|_2$.

3.2 Fourierreihen

Wir betrachten nun periodische Funktionen f mit der Periode 1: $f(x) = f(x+1)$ für alle x . Wir setzen voraus $f \in L_2([a, a+1])$. Zu einer solchen Funktion gibt es die Fourierkoeffizienten

$$c_n = \int_a^{a+1} f(x) e^{-2\pi i n x} dx.$$

Die Funktionen

$$e_n(x) := e^{2\pi i n x}$$

bilden wegen $\langle e_m, e_n \rangle = \int_a^{a+1} e^{2\pi i(m-n)x} dx = \delta_{mn}$ ein Orthonormalsystem in $L_2(a, a+1)$. Für die Fourierkoeffizienten gilt $c_n = \langle f, e_n \rangle$, und $c_n e_n$ ist die orthogonale Projektion des Vektors f auf den Vektor e_n .

Satz 3.12. Sei N eine feste natürliche Zahl. Die Norm

$$\left\| f - \sum_{n=-N}^N a_n e_n \right\|_2$$

wird minimal für $a_n = c_n = \langle f, e_n \rangle$. Ferner gilt die **Besselsche Ungleichung**

$$\|f\|^2 \geq \sum_{n=-\infty}^{\infty} |c_n|^2.$$

Korollar 3.13. Für $f \in L_2(a, a+1)$ sind folgende Aussagen äquivalent:

(i) Zu jedem $\varepsilon > 0$ gibt es ein $N \in \mathbb{N}$ und $a_n \in \mathbb{C}$ mit $\|f - \sum_{n=-N}^N a_n e_n\| \leq \varepsilon$.

(ii) $\lim_{N \rightarrow \infty} \|f - \sum_{n=-N}^N c_n e_n\| = 0$, d.h. es gilt die Konvergenzaussage $f = \sum_{n=-\infty}^{\infty} c_n e_n$ in der L_2 -Konvergenz.

(iii) Es gilt die Parsevalsche Gleichung

$$\sum_{n=-\infty}^{\infty} |c_n|^2 = \|f\|^2.$$

Man kann nun sogar das Folgende zeigen:

Satz 3.14. Für jedes $f \in L_2(a, a+1)$ gilt die Parsevalsche Gleichung. Insbesondere ist $\{e_n\}$ ein vollständiges Orthonormalsystem.

Nun betrachten wir den sog. **Hilbertschen Folgenraum**

$$l_2 := \{(x_n)_{n \in \mathbb{Z}} : x_n \in \mathbb{C} \text{ und } \sum_{n=-\infty}^{\infty} |x_n|^2 < \infty\}.$$

Satz 3.15. l_2 ist ein Hilbertraum mit Skalarprodukt $\langle x, y \rangle := \sum_{n=-\infty}^{\infty} x_n \overline{y_n}$ und Norm $\|x\| = \sqrt{\langle x, x \rangle}$.

Es gilt der

Hilfssatz 3.16. Für eine in $L_2 = L_2(a, a+1)$ konvergente Folge $f_n \rightarrow f$ gilt $\langle f_n, e_m \rangle \rightarrow \langle f, e_m \rangle$ (mit $e_m(x) = e^{2\pi i m x}$).

Satz 3.17. Die Abbildung $\varphi : L_2(a, a+1) \rightarrow l_2$ mit $\varphi(f) = (c_n)$ ist Isomorphismus von Hilberträumen.

3.3 Fourier-Integrale

Wir betrachten in diesem Paragraphen das Analogon zur Fourierentwicklung für nichtperiodische Funktionen. Wieder wird eine „Entwicklung“ von Funktionen f nach Gliedern $e^{2\pi i\lambda x}$ durchgeführt, wobei aber λ nicht auf ganze Zahlen beschränkt ist, sondern kontinuierlich läuft.

Genauer: Sei $f \in C_0^1(\mathbb{R})$, $f(x) = 0$ für $|x| > K$. Wir entwickeln $f(Tx)$ für $T > 2K$ im Intervall $[-\frac{1}{2}, \frac{1}{2}]$ in eine Fourierreihe:

$$f(Tx) = \sum_{n=-\infty}^{\infty} c_n e^{2\pi i n x}.$$

Wir wollen zeigen: Mit

$$g(y) = \int_{-\infty}^{\infty} f(x) e^{-2\pi i x y} dx$$

gilt

$$f(x) = \int_{-\infty}^{\infty} g(y) e^{2\pi i x y} dy.$$

Der Beweis wird in der Vorlesung durch Anwendung des Konvergenzsatzes (Anhang B.30) von Lebesgue geführt. Man konstruiert eine Funktion $h(y)$ mit $|g_{T,x}(y)| \leq h(y)$ und $\int_{-\infty}^{\infty} h(y) dy < \infty$ und führt für $y \neq 0$ zweimal partielle Integration durch.

$$h(y) := \begin{cases} \int_{-\infty}^{\infty} |f(x)| dx & \text{für } |y| < 1 \\ \frac{\int_{-\infty}^{\infty} |f''(x)| dx}{4\pi^2 \cdot (|y| - \frac{1}{\varepsilon})^2} & \text{für } |y| \geq 1 \end{cases}$$

leistet für $T \geq 1$ das Gewünschte. Insgesamt folgt:

Definition 3.18. Die für alle $x \in \mathbb{R}$ definierte Funktion f sei zweimal stetig differenzierbar und es gelte $f(x) = 0$ für $|x| > K$. Dann heißt

$$g(y) = \int_{-\infty}^{\infty} f(x) e^{-2\pi i x y} dx$$

die **Fourier-Transformierte** von f , und es gilt die **Fouriersche Umkehrformel**

$$f(x) = \int_{-\infty}^{\infty} g(y) e^{2\pi i x y} dy.$$

Man nennt das Integral $\int_{-\infty}^{\infty} g(y)e^{2\pi ixy} dx$ das **Fourier-Integral** von f . Bezeichnungen: $g(y) = Ff(y)$, $f(x) = \overline{F}g(x)$.

Ist $f \in L_2(\mathbb{R})$, so kann man mit der Parsevalschen Gleichung zeigen

$$\|f\|_2 = \|Ff\|_2,$$

d.h. die Fourier-Transformation erhält die L_2 -Norm. Die Voraussetzung über f kann aber noch abgeschwächt werden. Es gilt sogar der

Satz 3.19.

1. Es gibt lineare Abbildungen $F, \overline{F} : L_2(\mathbb{R}) \rightarrow L_2(\mathbb{R})$ mit

$$Ff(y) = \int_{-\infty}^{\infty} f(x)e^{-2\pi ixy} dx, \quad \overline{F}f(y) = \int_{-\infty}^{\infty} f(x)e^{2\pi ixy} dx$$

für alle $f \in C_0^2(\mathbb{R})$.

2. Es gilt $\|f\|_2 = \|Ff\|_2 = \|\overline{F}f\|_2$ für alle $f \in L_2(\mathbb{R})$.

3. Die Gleichungen aus 1. gelten für alle $f \in L_1(\mathbb{R}) \cap L_2(\mathbb{R})$.

4. Es gilt $F\overline{F}f = f = \overline{F}Ff$ für alle $f \in L_2(\mathbb{R})$.

5. Es gilt $\langle Ff, Fg \rangle = \langle f, g \rangle$ für alle $f, g \in L_2(\mathbb{R})$.

3.4 Distributionen

In diesem Paragraphen beschäftigt uns die sogenannte **Diracsche Deltafunktion** δ . Dirac hat diese Funktion (die keine Funktion im strengen Sinn ist) 1927 eingeführt; die einleitenden Vorbemerkungen stammen aus Dirac, „The Physical Interpretation of Quantum Dynamics“.

Zunächst die beiden Eigenschaften, die unsere Funktion laut Definition besitzen soll:

1. $\delta(x) = 0$ für alle $x \neq 0$;
2. $\int_{-a}^a \delta(x) dx = 1$ für alle $a > 0$.

Genau genommen gibt es keine solche Funktion, und Dirac schreibt: „Streng genommen kann man δ nur als Grenzwert einer gewissen Funktionenfolge ansehen.“

Eine solche Funktionenfolge (δ_n) kann folgende Eigenschaften haben: $\delta_n(x) = 0$ für alle x mit $|x| > \frac{1}{n}$ sowie $\delta_n(x) \geq 0$ und $\int_{-1}^1 \delta_n(x) dx = 1$ für alle n . Ist dann φ eine stetige Funktion, so gilt

$$\lim_{n \rightarrow \infty} \int_{-1}^1 \delta_n(x) \cdot \varphi(x) dx = \varphi(0) =: \delta(\varphi), \quad (3.2)$$

wie aus dem Mittelwertsatz der Integralrechnung folgt. Wir verwenden dabei die folgende Form des Mittelwertsatzes: Sind $g, f : [a, b] \rightarrow \mathbb{R}$ stetig und $g \geq 0$ sowie $\int_a^b g(x) dx \neq 0$, so existiert ein Wert $x^* \in [a, b]$ mit $\int_a^b f(x)g(x) dx = f(x^*) \cdot \int_a^b g(x) dx$.

Ist φ sogar stetig differenzierbar, so gilt

$$\lim_{n \rightarrow \infty} \int_{-1}^1 \delta'_n(x) \varphi(x) dx = \lim_{n \rightarrow \infty} [\delta_n(x) \varphi(x)]_1^{-1} - \lim_{n \rightarrow \infty} \int_{-1}^1 \delta_n(x) \cdot \varphi'(x) dx = -\varphi'(0).$$

Definition 3.20. Eine Abbildung $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ heißt **Testfunktion**, wenn gilt:

1. φ ist beliebig oft stetig differenzierbar.
2. Es gibt ein $K \in \mathbb{R}^+$ mit $\varphi(x) = 0$ für alle $|x| > K$.

D.h. es soll $\varphi \in C_0^\infty(\mathbb{R})$ gelten.

Bemerkung. Die Testfunktionen bilden einen reellen Vektorraum. Dieser wird mit \mathcal{D} bezeichnet.

Definition 3.21. Eine Funktionenfolge $(\varphi_n) \in \mathcal{D}$ heißt **konvergent gegen** $\varphi \in \mathcal{D}$, in Zeichen

$$\varphi_n \xrightarrow{\mathcal{D}} \varphi,$$

falls gilt:

- (i) Es gibt ein Kompaktum $K \subseteq \mathbb{R}$, so dass $\text{Tr}(\varphi_n) \subseteq K$ für alle $n \in \mathbb{N}$ und $\text{Tr}(\varphi) \subseteq K$ gilt.

(ii) Für jedes $m \geq 0$ konvergiert die Folge $\varphi_n^{(m)}$ für $n \rightarrow \infty$ gleichmäßig auf K gegen $\varphi^{(m)}$. D.h. für alle Ableitungen und für φ_n wird die gleichmäßige Konvergenz gefordert.

Definition 3.22. Eine **Distribution** T ist eine stetige lineare Abbildung $T : \mathcal{D} \rightarrow \mathbb{R}$, also ein Vektorraumhomomorphismus mit

$$\varphi_n \xrightarrow{\mathcal{D}} \varphi \Rightarrow T\varphi_n \rightarrow T\varphi \text{ für } n \rightarrow \infty.$$

Die Menge aller Distributionen wird mit $\mathcal{D}' = \mathcal{D}'(\mathbb{R})$ bezeichnet und ist selbst wieder ein reeller Vektorraum.

Bemerkung. Man definiert ebenso für den n -dimensionalen Fall

$$\mathcal{D}(\mathbb{R}^n) = C_0^\infty(\mathbb{R}^n)$$

als die Menge der beliebig oft differenzierbaren Funktionen $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ mit kompaktem Träger. Der entsprechende Konvergenzbegriff lautet dann wie folgt: $\varphi_n \xrightarrow{\mathcal{D}} \varphi$, d.h. φ_n konvergiert in \mathcal{D} gegen φ , falls gilt:

- (i) Es gibt ein Kompaktum $K \subseteq \mathbb{R}^n$, so dass $Tr(\varphi_n) \subseteq K$ für alle $n \in \mathbb{N}$ und $Tr(\varphi) \subseteq K$ gilt.
- (ii) Für jeden Multiindex $\alpha \in \mathbb{N}^n$, $k \in \mathbb{N}$ konvergiert die Folge der Ableitungen

$$D^\alpha \varphi_k = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}} \varphi_k$$

gleichmäßig auf K gegen $D^\alpha \varphi$.

Bei $\mathcal{D}'(\mathbb{R}^n)$ verlangt man wieder die Stetigkeit von $T : \mathcal{D} \rightarrow \mathbb{R}$.

Beispiel 3.23. Sei $f : \mathbb{R}^n \rightarrow \mathbb{R}$ stetig. Es genügt auch die schwächere Bedingung der lokalen Integrierbarkeit von f , d.h. die Existenz von $\int_K f(x)dx$ für ein beliebiges Kompaktum K .

Für $\varphi \in \mathcal{D}(\mathbb{R}^n)$ setze

$$T_f(\varphi) := \int_{\mathbb{R}^n} f(x)\varphi(x)dx.$$

Diese Abbildung T_f ist dann stets Distribution.

Bemerkung 3.24. $T_f = T_g$ gilt genau dann, wenn $f(x) = g(x)$ für fast alle x ist.

Beispiel 3.25. Für $a \in \mathbb{R}^n$ und $\varphi \in \mathcal{D}(\mathbb{R}^n)$ gibt es stets die **Diracsche Delta-Distribution**:

$$\delta_a(\varphi) := \varphi(a).$$

Bemerkung 3.26. Setzt man

$$(S + T)(\varphi) := S(\varphi) + T(\varphi) \quad (S, T \in \mathcal{D}')$$

sowie

$$(cT)(\varphi) := c \cdot T(\varphi) \quad (c \in \mathbb{R}, T \in \mathcal{D}'),$$

so bilden die Distributionen einen \mathbb{R} -Vektorraum.

Man kann auch einen Konvergenzbegriff für Distributionen einführen:

Definition 3.27. $T_\nu \in \mathcal{D}'(\mathbb{R}^n)$ und $T \in \mathcal{D}'(\mathbb{R}^n)$ seien Distributionen. Man sagt, die Folge T_ν **konvergiert in \mathcal{D}' gegen T** , in Zeichen

$$T_\nu \xrightarrow{\mathcal{D}'} T,$$

falls $T_\nu(\varphi) \rightarrow T(\varphi)$ für alle $\varphi \in \mathcal{D}(\mathbb{R}^n)$ gilt. D.h. die Konvergenz einer Distribution wird so mit Funktionen φ aus \mathcal{D} „getestet“.

Bemerkung. Man kann zeigen, dass die δ -Distribution Grenzwert einer Folge von Distributionen T_{f_n} ist. Vgl. Forster, Analysis III, §17, Beispiel (17.4).

Differentiation von Distributionen:

Zur Erinnerung: Ist $\nu = (\nu_1, \dots, \nu_n)$ ein Multiindex, so ist ein Differentialoperator der Ordnung k wie folgt gegeben:

$$L = \sum_{|\nu| \leq k} a_\nu \cdot \frac{\partial^{\nu_1 + \dots + \nu_n}}{\partial x^\nu}$$

mit Funktionen $a_\nu \in C^\infty(\mathbb{R}^n)$. Wir wollen nun eine Abbildung $L : \mathcal{D}'(\mathbb{R}^n) \rightarrow \mathcal{D}'(\mathbb{R}^n)$ erklären, so dass für $f \in C^k(\mathbb{R}^n)$ stets

$$LT_f = T_{Lf}$$

gilt. Dazu benötigen wir den Begriff des zu einem Differentialoperator L **adjungierten** Operators. Zunächst der Fall $n = 1$: Es gilt

$$\begin{aligned} T_{f'}(\varphi) &= \int_{-\infty}^{\infty} f'(x)\varphi(x)dx = [f(x) \cdot \varphi(x)]_{\infty}^{-\infty} - \int_{-\infty}^{\infty} f(x)\varphi'(x)dx = \\ &= - \int_{-\infty}^{\infty} f(x)\varphi'(x)dx = -T(\varphi'). \end{aligned}$$

Ist $\varphi \in \mathcal{D}$, so setzen wir daher

$$T'(\varphi) := -T(\varphi'). \quad (3.3)$$

Mit φ ist auch $\varphi' \in \mathcal{D}$. Für $i \rightarrow \infty$ strebt $T\varphi_i^{(m)}$ für jedes feste m gegen Null. Daher ist $T'(\varphi)$ wieder eine Distribution. Mit Induktion über m erhalten wir

$$T^{(m)}(\varphi) = (-1)^m \cdot T(\varphi^{(m)}),$$

d.h. jede Distribution ist unendlich oft differenzierbar. Allgemeiner: wir betrachten die Formel

$$\int_{-\infty}^{\infty} f'(x)\varphi(x)dx = - \int_{-\infty}^{\infty} f(x)\varphi'(x)dx$$

und suchen nach einer Verallgemeinerung im n -dimensionalen Raum.

Definition 3.28. Sei L ein Differentialoperator der Ordnung k in der offenen Menge $U \subseteq \mathbb{R}^n$. Ein Differentialoperator M der Ordnung k in U heißt **adjungiert zu L** , falls

$$\int_U (Mf)(x)g(x)dx = \int_U f(x) \cdot (Lg)(x)dx$$

für alle $f \in C^k(U)$, $g \in C_0^k(U)$ gilt.

Bemerkung. Es gilt die Existenz und die Eindeutigkeit. Der eindeutig bestimmte Differentialoperator M aus (3.28) wird auch mit L^* bezeichnet.

Satz 3.29. Zu jedem linearen Differentialoperator L in der offenen Menge $U \subseteq \mathbb{R}^n$ gibt es einen adjungierten. Es gelten die Rechenregeln:

$$(i) \quad (\lambda L)^* = \lambda L^*.$$

$$(ii) \quad (L_1 + L_2)^* = L_1^* + L_2^*.$$

$$(iii) \quad (L_1 \circ L_2)^* = L_2^* \circ L_1^*.$$

Aufgrund dieser Rechenregeln besteht eine formale Analogie der Bildung eines adjungierten Operators zur Transposition von Matrizen.

Definition 3.30. Für $T \in \mathcal{D}'(\mathbb{R}^n)$ und lineare Differentialoperatoren L der Ordnung k definieren wir

$$(LT)(\varphi) := T(L^* \varphi).$$

Bemerkung. Wegen der Linearität von LT ist diese Definition sinnvoll. Ferner ist LT stetig und wieder Distribution.

Beispiel 3.31. Wir betrachten die **Heavyside-Funktion**

$$H(x) = \begin{cases} 0 & \text{für } x < 0 \\ 1 & \text{für } x \geq 0. \end{cases}$$

Aus ihr erhält man die Heavyside-Distribution

$$Y(\varphi) = \int_0^\infty \varphi(x) dx = T_H(\varphi).$$

Die Ableitung dieser Distribution ist

$$Y'(\varphi) = -Y(\varphi') = - \int_0^\infty \varphi'(x) dx = \varphi(0) = \delta(\varphi),$$

also die Delta-Distribution. Für die höheren Ableitungen gilt $\delta^{(n)}(\varphi) = (-1)^n \varphi^{(n)}(0)$.

Beispiel 3.32. Sei

$$g(x) = \begin{cases} 0 & \text{für } x \leq 0 \\ \frac{1}{\sqrt{x}} & \text{für } x > 0. \end{cases}$$

Zu dieser Funktion existiert $\mathcal{Z} = T_g$. Da jedoch $g'(x) = -\frac{1}{2x^{3/2}}$ für $x > 0$ nicht integrierbar ist, kann man nicht ohne weiteres $\mathcal{Z}' = T_{g'}$ erwarten. Es gilt

$$\begin{aligned} \mathcal{Z}'(\varphi) &= -\mathcal{Z}(\varphi') = - \int_0^\infty \frac{\varphi'(x)}{\sqrt{x}} dx = \\ &= - \lim_{\varepsilon \rightarrow 0} \int_\varepsilon^\infty \frac{\varphi'(x)}{\sqrt{x}} dx = \lim_{\varepsilon \rightarrow 0} \left[\frac{\varphi(\varepsilon)}{\sqrt{\varepsilon}} - \frac{1}{2} \int_\varepsilon^\infty \frac{\varphi(x)}{x^{3/2}} dx \right] = \\ &= \lim_{\varepsilon \rightarrow 0} \left[\frac{\varphi(0)}{\sqrt{\varepsilon}} - \frac{1}{2} \int_\varepsilon^\infty \frac{\varphi(x)}{x^{3/2}} dx \right]. \end{aligned}$$

Dies ist der Spezialfall der von Hadamard ("Le problème de Cauchy") eingeführten "partie finite".

Wir behandeln nun die Umkehrung der Differentiation von Distributionen, d.h. die Integration von Distributionen. Zu einer gegebenen Distribution T wird eine Distribution S mit $S' = T$ gesucht. In diesem Fall heißt S eine **Stamm-Distribution** zu T .

Bemerkung. Falls mit φ auch $x \mapsto \int_{-\infty}^x \varphi(t)dt$ wieder Testfunktion ist, erhält man durch

$$S(\varphi(x)) = -T \left(\int_{-\infty}^x \varphi(t)dt \right)$$

eine Stamm-Distribution wegen $S'(\varphi) = T \left(\int_{-\infty}^x \varphi'(t)dt \right) = T(\varphi(x))$.

Dies ist aber nicht immer der Fall:

Hilfssatz 3.33. *Eine Testfunktion ψ lässt sich genau dann als Ableitung einer Testfunktion φ darstellen, wenn $\int_{-\infty}^{\infty} \psi(x)dx = 0$ gilt.*

Bemerkung. Die Testfunktion φ , die der Gleichung $\varphi' = \psi$ genügt, ist durch ψ eindeutig bestimmt.

Satz 3.34. *T sei Distribution, φ_0 Testfunktion mit $\int_{-\infty}^{\infty} \varphi_0(x)dx = T_1\varphi_0 = 1$. Die durch*

$$S(\varphi) = -T \left(\int_{-\infty}^x \varphi(t) - \varphi_0(t) \cdot T_1\varphi dt \right)$$

definierte lineare Abbildung ist eine Stamm-Distribution von T . Zwei Stamm-Distributionen zu T unterscheiden sich höchstens um einen Summanden cT_1 mit einer Konstanten c .

Wir behandeln als nächstes die Multiplikation einer Funktion $\gamma \in C^\infty$ mit einer Distribution. Für eine solche Multiplikation soll die Forderung

$$\gamma \cdot T_f = T_{\gamma \cdot f}$$

stets erfüllt sein. Es folgt

$$\gamma \cdot \int_{-\infty}^{\infty} f(x)\varphi(x)dx = \int_{-\infty}^{\infty} \gamma(x)f(x)\varphi(x)dx = T_f(\gamma\varphi).$$

Definition 3.35. $\gamma T(\varphi) = T(\gamma\varphi)$ für beliebige Distributionen T und $\gamma \in C^\infty$.

Beispiel: Ist T die Delta-Distribution δ , so gilt

$$\gamma\delta(\varphi) = \delta(\gamma\varphi) = \gamma(0) \cdot \varphi(0) = \gamma(0) \cdot \delta(\varphi).$$

Es gilt stets die Regel $(\gamma T)'(\varphi) = \gamma T'(\varphi) + \gamma' \cdot T(\varphi)$ und damit das folgende Analogon zur Produktregel

$$(\gamma T)' = \gamma T' + \gamma' T.$$

Schwieriger ist es, die Multiplikation von zwei Distributionen zu erklären — etwa ein Produkt der Delta-Distribution mit sich selbst. Bevor wir darauf und auf die Faltungsoperation zu sprechen kommen, hier noch die Umkehrung der Multiplikation: Zu einer Distribution S und einer Funktion γ aus C^∞ wird eine Distribution T mit $\gamma T = S$ gesucht. Es erhebt sich die Frage, wie man in den Nullstellen von γ verfährt.

Beispiel: Sei $\gamma(x) = e^x - 1 = x \frac{e^x - 1}{x}$. Die Funktion $x \mapsto \frac{e^x - 1}{x}$ ist stetig fortsetzbar bei $x = 0$, und es ist $\lim_{x \rightarrow 0} \frac{e^x - 1}{x} = 1$. Es folgt $T^* = \frac{x}{e^x - 1} \cdot S$, und damit ist das Problem auf den Fall $\gamma = x$ reduziert.

Zweites Beispiel: Sei $\gamma(x) = \prod_{i=1}^n (a_i - x)$. Dann behandelt man die Terme $a_i - x$ einzeln und wendet das Verfahren für $\gamma = x$ schrittweise an. Bei der Frage nach einer Umkehrung der Multiplikation ist also stets der Fall $\gamma = x$ zu behandeln.

Hilfssatz 3.36. *Eine Testfunktion ψ lässt sich als $\psi(x) = x \cdot \varphi(x)$ mit einer weiteren Testfunktion φ genau dann darstellen, wenn $\psi(0) = 0$ gilt.*

Bemerkung. φ ist durch ψ eindeutig bestimmt, und ist $\{\psi_i\}$ eine Folge von Testfunktionen, die in \mathcal{D} gegen 0 konvergiert, so gilt dasselbe für die Folge $\{\varphi_i\}$.

Satz 3.37. *Sei S eine Distribution im \mathbb{R}^1 . Dann existiert eine Distribution T mit $xT = S$.*

Zum Beweis betrachtet man die eindeutige Darstellung von φ , wobei φ_0 eine fest gewählte Testfunktion mit $\varphi_0(0) = 1$ ist, in der Form $\varphi = \varphi(0) \cdot \varphi_0 + x\chi$. Man setzt $T(\varphi) := c_0\delta(\varphi) + S\chi$.

Im Spezialfall $S = 0$ ist $T(\varphi) = c_0 \cdot \delta(\varphi)$. Alle Vielfachen der δ -Distribution sind damit Lösungen.

Ein weiteres Beispiel: Zu lösen ist $xT = T_1$ und damit

$$T_1(\varphi) = \int_{-\infty}^{\infty} 1 \cdot \varphi(x) dx = \int_{-\infty}^{\infty} \varphi(x) dx.$$

Der Ansatz $T_{1/x}$ ist falsch, da $\frac{1}{x}$ bei $x = 0$ nicht integrierbar ist. Also beginnt man mit

$$\varphi = \varphi(0) \cdot \varphi_0 + \psi.$$

Aus dem Hilfssatz folgt $\psi = \varphi - \varphi_0 \cdot \varphi(0) = x \cdot \chi$ und damit $T(\varphi) = c_0 \delta(\varphi) + T_1(\chi)$. Wähle $c = 0$; für $x \neq 0$ ist

$$\chi(x) = \frac{\varphi(x) - \varphi(0)\varphi_0(x)}{x}.$$

Daher existiert

$$\lim_{\varepsilon \rightarrow 0} \int_{-\infty}^{\infty} \frac{\varphi(x) - \varphi(0)\varphi_0(x)}{x} dx = \int_{-\infty}^{\infty} \frac{\varphi(x) - \varphi(0)\varphi_0(x)}{x} dx = T_1(\chi) = T(\varphi).$$

Wählt man $\varphi_0(x)$ symmetrisch zur y -Achse, so ist

$$\int_{-\infty}^{-\varepsilon} \frac{\varphi_0(x)}{x} dx = - \int_{\varepsilon}^{\infty} \frac{\varphi_0(x)}{x} dx$$

und damit

$$T(\varphi) = \lim_{\varepsilon \rightarrow 0} \left(\int_{-\infty}^{-\varepsilon} \frac{\varphi(x)}{x} dx + \int_{\varepsilon}^{\infty} \frac{\varphi(x)}{x} dx \right).$$

Dieser Grenzwert heißt auch **Cauchyscher Hauptwert** von $\int_{-\infty}^{\infty} \frac{\varphi(x)}{x} dx$. Man beachte, dass jeder einzelne Term dieser Summe für $\varepsilon \rightarrow 0$ divergiert, und zwar in der Größenordnung $\log \varepsilon$ bzw. $-\log \varepsilon$.

Fourier-Transformation von Distributionen:

Ist φ eine Testfunktion, so ist die Fourier-Transformierte $F\varphi$ keine Testfunktion (außer im Fall $\varphi \equiv 0$). Aus diesem Grund konstruieren wir eine Erweiterung unseres bisherigen Raumes von Testfunktionen:

Definition 3.38. *Es sei*

$$S = \{ \varphi \in C^\infty(\mathbb{R}) : |x^p \cdot \varphi^{(q)}(x)| \leq k_{pq} \text{ für alle } x \in \mathbb{R} \},$$

oder allgemeiner im \mathbb{R}^n mit der Bedingung $|x^k \cdot D^q \cdot \varphi(x)| \leq c = c_{kq}$ für alle $x \in \mathbb{R}$, wobei $k = (k_1, \dots, k_n)$ und $q = (q_1, \dots, q_n)$ sowie $x^k = x_1^{k_1} \dots x_n^{k_n}$ ist sowie $D^q = \frac{\partial^{(q)}}{\partial x_1^{q_1} \dots \partial x_n^{q_n}}$.

In Worten: S sei der Raum der C^∞ -Funktionen, die mit ihren Ableitungen für $\|x\| \rightarrow \infty$ schneller gegen Null gehen als jede Potenz von $\|x\|^{-1} = \frac{1}{\|x\|}$. Offenbar gilt $S \supseteq \mathcal{D}$.

Für Funktionenfolgen $\varphi_k \in S$ definieren wir die Konvergenz gegen Null, in Zeichen $\varphi_k \xrightarrow{s} 0$, durch folgende Bedingungen:

1. Zu jedem Paar (p, q) existiert eine Konstante $C_{p,q}$, so dass $|x^p \cdot q_k^{(q)}| \leq C_{p,q}$ für alle $x \in \mathbb{R}$ und alle $k \in \mathbb{N}$ gilt.
2. Für jedes p geht $\varphi_k^{(p)}$ gleichmäßig gegen Null für $k \rightarrow \infty$.

(Alternativ: Abzählbares System von Normen $\|\varphi\|_p := \sup_{|K|, |q| \leq p} \|x^K D^q \varphi(x)\|$.)

Hilfssatz 3.39. Die Fouriertransformation F definiert eine lineare, stetige Abbildung von S in sich selbst (Beweis in der Vorlesung nur für $n = 1$). Insbesondere gehört mit φ auch $F\varphi$ wieder zu S .

Wir können nun Distributionen auf S betrachten („temperierte Distributionen“). Zunächst ist die Fouriertransformation einer Distribution zu definieren.

Forderung: Mit $T = T_f$ soll

$$T_{Ff}(\varphi) = \int_{-\infty}^{\infty} Ff(y)\varphi(y)dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x)e^{-2\pi ixy}dx\varphi(y)dy$$

sein.

Der Satz von Fubini liefert

$$\begin{aligned} T_{Ff}(\varphi) &= \int_{\mathbb{R}^2} f(x)e^{-2\pi ixy}\varphi(y)dx dy = \\ &= \int_{-\infty}^{\infty} f(x) \int_{-\infty}^{\infty} \varphi(y)e^{-2\pi ixy}dy dx = T_f(F\varphi). \end{aligned}$$

Wir definieren daher

Definition 3.40. $(FT)(\varphi) := T(F\varphi)$.

Damit ist FT als Komposition eines linearen, stetigen Operators und eines linearen, stetigen Funktionals wieder Distribution.

Rechenregeln:

1. $F(T)'(\varphi) = T'(F\varphi) = -T((F\varphi)').$

$$2. (FT)'(\varphi) = -(FT)(\varphi') = -T(F\varphi').$$

In der mehrdimensionalen Situation ist $S = S(\mathbb{R}^n)$ und $\left(\frac{\partial}{\partial y_i} T\right)(\varphi) = -T\left(\frac{\partial}{\partial y_i} \varphi\right)$.
Damit ist

$$\frac{\partial}{\partial y_i} FT = F(-2\pi i y_i T).$$

Sei nun eine Distribution S gegeben und T gesucht derart, dass $\Delta T = S$ gilt; dabei ist Δ der Laplace-Operator $\Delta = \frac{\partial^2}{\partial x_1^2} + \dots + \frac{\partial^2}{\partial x_n^2}$.

Der Ansatz $T = FR$ führt auf

$$\begin{aligned} \Delta FR &= \frac{\partial^2}{\partial x_1^2}(FR) + \dots + \frac{\partial^2}{\partial x_n^2}(FR) = \\ &= F(-4\pi^2 x_1^2 R - \dots - 4\pi^2 x_n^2 R) = \\ &= -4\pi^2(Fx_1^2 R + \dots + Fx_n^2 R) = -4\pi^2 F(x_1^2 + \dots + x_n^2)R. \end{aligned}$$

Da neben F auch $\bar{F} = F^{-1}$ auf Distributionen anwendbar ist, erhalten wir folgende Lösung:

$$(x_1^2 + \dots + x_n^2)R = -\frac{1}{4\pi^2} \cdot \bar{F}S.$$

Denn $\Delta T = S$ mit $T = FR$ ist gleichbedeutend mit $S = \Delta FR = -4\pi^2 F(x_1^2 + \dots + x_n^2)R$, und es gilt $F^{-1}\Delta FR = -4\pi^2 F^{-1}F(x_1^2 + \dots + x_n^2)R$ und daher $(x_1^2 + \dots + x_n^2)R = -\frac{1}{4\pi^2}(F^{-1}\Delta F)R$.

Also: Eine partielle Differentialgleichung im Raum der Distributionen kann auf die Frage der Auflösbarkeit von Gleichungen des Typs $\gamma \cdot T = S$ zurückgeführt werden.

Im Anhang sind zwei Zugänge zur n -dimensionalen Integrationstheorie dargestellt. Der erste folgt weitgehend einem Skript von E. Freitag [6], der zweite, der die klassische Maßtheorie benutzt, lehnt sich an das Skript von M. Kneser an [12].

A Integrationstheorie

A.1 Der Approximationssatz von Stone-Weierstraß

Der Satz von Stone-Weierstraß beschäftigt sich mit der gleichmäßigen Approximation einer stetigen Funktion $f : [a, b] \rightarrow \mathbb{R}$ durch eine vorgegebene Klasse von Funktionen h . Wir treffen die Einschränkung, dass es sich in jedem Fall um Klassen **stetiger** Funktionen handelt.

Satz A.1 (Approximationssatz von Stone-Weierstraß, erste Variante). *Es sei $X \subseteq \mathbb{R}^n$ eine kompakte Teilmenge, die mindestens zwei Punkte enthält. Weiterhin sei W eine Menge von stetigen reellwertigen Funktionen auf X mit den Eigenschaften*

$$1) f, g \in W \implies f + g \in W \text{ und } cf \in W \text{ für } c \in \mathbb{R}.$$

$$2) f \in W \implies |f| \in W.$$

3) (Punkt-trennungseigenschaft):

Seien a, b zwei verschiedene Punkte von X . Dann existiert ein $f \in W$ mit $f(a) = 0$ aber $f(b) \neq 0$.

Unter diesen Voraussetzungen gilt: Ist $f \in C(X)$ eine beliebige stetige Funktion auf X und $\varepsilon > 0$, so existiert eine Funktion

$$h \in W \text{ mit } \|f - h\| \leq \varepsilon$$

(wobei $\|\cdot\|$ die Supremumsnorm bezeichnet). Zusatz: Insbesondere existiert dann eine Folge (h_n) von Funktionen aus W , die gleichmäßig gegen f konvergiert. (Man setze $\varepsilon := \frac{1}{n}$ und bezeichne die zugehörige Funktion mit h_n .)

Beweis: Zunächst einige Bezeichnungen:

Sei $f : X \rightarrow \mathbb{R}$ eine Funktion. Dann sei

$$f^+ : X \rightarrow \mathbb{R}, \quad f^+(x) := \max(f(x), 0).$$

Offenbar gilt:

$$f^+ = \frac{1}{2}(|f| + f).$$

Insbesondere gilt

$$f \in W \implies f^+ \in W \quad (\text{wegen 1) und 2)}).$$

Sei $g : X \rightarrow \mathbb{R}$ eine weitere Funktion. Man definiert

$$f \vee g : X \rightarrow \mathbb{R}, \quad (f \vee g)(x) := \max(f(x), g(x)),$$

$$f \wedge g : X \rightarrow \mathbb{R}, \quad (f \wedge g)(x) := \min(f(x), g(x)).$$

Offenbar gilt

$$f \vee g = f + (g - f)^+,$$

$$f \wedge g = f - (g - f)^+.$$

Hieraus kann man sofort folgern:

$$f, g \in W \implies f \vee g \in W, \quad f \wedge g \in W.$$

Das gleiche gilt dann auch für endlich viele Funktionen f_1, \dots, f_n .

Nun soll (A.1) in mehreren Schritten bewiesen werden. Wir formulieren erst die Schritte.

1. Schritt. Seien $a, b \in X$ verschiedene Punkte und A, B zwei reelle Zahlen. Dann existiert $f \in W$ mit

$$f(a) = A, \quad f(b) = B.$$

2. Schritt. Sei $f : X \rightarrow \mathbb{R}$ stetig, $a, b \in X$ und $\varepsilon > 0$. Dann gibt es eine offene Umgebung V_b von b und eine Funktion $h_{a,b} \in W$ mit

$$h_{a,b}(a) = f(a) - \frac{\varepsilon}{2} \quad \text{und} \quad h_{a,b}(x) < f(x) \quad \text{für alle } x \in V_b \cap X.$$

3. Schritt. Sei $f : X \rightarrow \mathbb{R}$ stetig, $a \in X$ und $\varepsilon > 0$. Dann gibt es eine offene Umgebung U_a von a und eine Funktion $h_a \in W$ mit

$$f(x) > h_a(x) \quad \text{für alle } x \in U_a$$

und

$$h_a(x) > f(x) - \varepsilon \text{ für alle } x \in U_a \cap X.$$

4. Schritt: Beweis des Theorems.

Beweis des 1. Schrittes. Nach 3) existieren Funktionen $g, h \in W$ mit

$$g(a) \neq 0, \quad g(b) = 0 \text{ und } h(a) = 0, \quad h(b) \neq 0.$$

Man setze

$$f := \frac{A}{g(a)}g + \frac{B}{h(b)}h.$$

Insbesondere existiert zu jedem $a \in X$ eine Funktion $f \in W$ mit vorgegebenem Funktionswert in a .

Beweis des 2. Schrittes. Nach dem 1. Schritt existiert eine Funktion $h_{a,b} \in W$ mit

$$h_{a,b}(a) = f(a) - \frac{\varepsilon}{2} \quad \text{und} \quad h_{a,b}(b) = f(b) - \frac{\varepsilon}{2}.$$

Die Menge

$$V_b := \mathbb{R}^n \setminus \{x \in X : h_{a,b}(x) \geq f(x)\}$$

ist eine offene Umgebung von b (denn b ist offenbar in V_b enthalten, und $\{x \in X : h_{a,b}(x) \geq f(x)\}$ ist eine abgeschlossene Menge, weil sie das Urbild von $[0, \infty)$ unter der stetigen Abbildung $f - h_{a,b}$ ist).

Beweis des 3. Schrittes. Sei $f \in C(X, \mathbb{R})$, $a \in X$ und $\varepsilon > 0$. Zu jedem $b \in X$ existieren $h_{a,b} \in W$ und eine offene Umgebung V_b mit den Eigenschaften, die im zweiten Schritt formuliert wurden. Es gilt

$$X \subseteq \bigcup_{b \in X} V_b.$$

X ist kompakt, kann also durch endlich viele der V_b überdeckt werden:

$$X \subseteq V_{b_1} \cup \dots \cup V_{b_n} \quad (b_1, \dots, b_n \text{ geeignet}).$$

Wir definieren nun

$$h_a := h_{a,b_1} \wedge \dots \wedge h_{a,b_n}.$$

Dies ist wieder eine Funktion aus W . Nach Konstruktion von $h_{a,b}$ gilt ferner

$$h_a(x) \leq h_{a,b_j}(x) < f(x) \text{ für } x \in X \cap V_{b_j}.$$

Da die V_{b_j} ganz X überdecken, gilt

$$h_a(x) < f(x) \text{ für alle } x \in X.$$

Das ist eine der Behauptungen des dritten Schrittes. Es bleibt die Umgebung U_a zu konstruieren. Wegen

$$h_a(a) = f(a) - \frac{\varepsilon}{2} > f(a) - \varepsilon$$

ist

$$U_a := \mathbb{R}^n \setminus \{x \in X : h_a(x) \leq f(x) - \varepsilon\}$$

wieder eine offene Umgebung von a .

4. Schritt: Beweis des Theorems. Sei $f \in C(X)$. Zu jedem $a \in X$ wurde eine Funktion $h_a \in W$ konstruiert mit

$$f(x) > h_a(x) \text{ f\"ur alle } x \in X,$$

aber

$$h_a(x) > f(x) - \varepsilon \text{ f\"ur alle } x \in X \cap U_a,$$

wobei U_a eine offene Umgebung von a ist. Es gilt

$$X \subseteq \bigcup_{a \in X} U_a$$

und somit aufgrund der Kompaktheit von X bereits

$$X \subseteq U_{a_1} \cup \dots \cup U_{a_n}.$$

Man definiere nun

$$h := h_{a_1} \vee \dots \vee h_{a_n}.$$

Offenbar liegt h in W und es gilt

$$f(x) > h(x) > f(x) - \varepsilon \text{ f\"ur alle } x \in X,$$

insbesondere also

$$|f(x) - h(x)| < \varepsilon \text{ f\"ur alle } x,$$

was genau bedeutet $\|f - h\| < \varepsilon$. □

Beispiel A.2. Wir nehmen an, X sei ein eindimensionales kompaktes Intervall:

$$X = [a, b], \quad a < b$$

und W die Menge aller stetigen Funktionen $f : [a, b] \rightarrow \mathbb{R}$, für die es Stützstellen

$$a = a_0 < a_1 < \dots < a_n = b$$

gibt, so dass f für $0 \leq \nu < n$ linear auf $[a_\nu, a_{\nu+1}]$ ist.

Die Menge W erfüllt die in (A.1) geforderten Voraussetzungen. Daher kann man jede stetige Funktion durch stückweise lineare Funktionen gleichmäßig approximieren.

Satz A.3 (Approximationssatz von Stone-Weierstraß, zweite Variante). *Sei X eine nichtleere kompakte Teilmenge des \mathbb{R}^n und $W \subseteq C(X, \mathbb{R})$ eine Menge von stetigen reellwertigen Funktionen auf X mit den Eigenschaften*

1) $f, g \in W \Rightarrow f + g \in W$ und $cf \in W$ für $c \in \mathbb{R}$.

2) $f, g \in W \Rightarrow f \cdot g \in W$ und die konstanten Funktionen liegen in W .

3) W besitzt die Punkttrennungseigenschaft aus (A.1).

Unter diesen Voraussetzungen gilt: Ist $f \in C(X)$ eine beliebige stetige Funktion auf X und $\varepsilon > 0$, so existiert eine Funktion

$$h \in W \text{ mit } \|f - h\| \leq \varepsilon.$$

Der wesentliche Unterschied zu (A.1) besteht darin, dass man die Bedingung

$$f \in W \Rightarrow |f| \in W$$

durch die Bedingung

$$f, g \in W \Rightarrow f \cdot g \in W$$

ersetzt hat.

Beweis von A.3: Sei \overline{W} der Abschluss von W in $C(X)$, d.h. \overline{W} besteht aus allen stetigen Funktionen auf X , die sich gleichmäßig durch Funktionen aus W approximieren lassen. Also:

$$f \in \overline{W} \iff \text{zu jedem } \varepsilon > 0 \text{ existiert } h \in W \text{ mit } \|f - h\| < \varepsilon.$$

Zunächst machen wir uns klar, dass \overline{W} wieder die Eigenschaften 1)-3) von A.3 erfüllt. Es ist zu zeigen:

$$f, g \in \overline{W} \Rightarrow f + g, f \cdot g, cf \in \overline{W}.$$

- (a) Wir zeigen $f + g \in \overline{W}$. Sei $\varepsilon > 0$ vorgegeben und f_1, g_1 so gewählt, dass

$$\|f - f_1\| < \varepsilon \text{ und } \|g - g_1\| < \varepsilon$$

gilt. Es folgt

$$\|f + g - (f_1 + g_1)\| < \|f - f_1\| + \|g - g_1\| < 2\varepsilon.$$

- (b) Es soll $f \cdot g \in \overline{W}$ gezeigt werden. Sei wieder $\varepsilon > 0$ vorgegeben. Wir setzen

$$\delta := \min\left(1, \frac{\varepsilon}{1 + \|f\| + \|g\|}\right)$$

und bestimmen $f_1, g_1 \in W$ so, dass

$$\|f - f_1\| < \delta \text{ und } \|g - g_1\| < \delta$$

gilt. Es folgt

$$\begin{aligned} \|fg - f_1g_1\| &= \|-(f - f_1)(g - g_1) + f(g - g_1) + g(f - f_1)\| \\ &\leq \|f - f_1\| \cdot \|g - g_1\| + \|f\| \cdot \|g - g_1\| + \|g\| \cdot \|f - f_1\| \\ &\leq \delta^2 + \delta\|f\| + \delta\|g\| = \delta(\delta + \|f\| + \|g\|) \\ &\leq \delta(1 + \|f\| + \|g\|) \leq \varepsilon. \end{aligned}$$

Hierbei wurde neben der Dreiecksungleichung $\|f + g\| \leq \|f\| + \|g\|$ auch noch die Ungleichung $\|f \cdot g\| \leq \|f\| \cdot \|g\|$ benutzt, welche leicht zu verifizieren ist.

Damit besitzt also auch \overline{W} die Eigenschaften 1)-3) aus (A.3). Die Punkte-trennungseigenschaft ist dabei trivial, denn es gilt ja $W \subseteq \overline{W}$.

Jetzt zeigen wir sogar

$$f \in \overline{W} \implies |f| \in \overline{W}.$$

Wenn wir dies gezeigt haben, so sind wir fertig, denn nach der ersten Variante (A.1) kann dann jede stetige Funktion $f \in C(X)$ gleichmäßig durch Funktionen aus \overline{W} approximiert werden, d.h. ist $\varepsilon > 0$, so existiert $g \in \overline{W}$ mit

$$\|f - g\| < \varepsilon.$$

Andererseits existiert nach Definition von \overline{W} eine Funktion

$$h \in W \text{ mit } \|f - h\| = \|(f - g) + (g - h)\| < 2\varepsilon.$$

Wir haben also gezeigt, dass $\overline{\overline{W}} = \overline{W}$ gilt, was offenbar in beliebigen metrischen Räumen richtig ist. \square

Halten wir noch einmal fest:

Die zweite Variante A.3 des Approximationssatzes von Stone-Weierstraß ist zurückgeführt auf die erste, wenn man zeigen kann, dass für $f \in W$ die Funktion $|f|$ gleichmäßig durch Funktionen aus W approximiert werden kann. Wir werden sogar zeigen:

Die Funktion $|f|$ lässt sich gleichmäßig durch endliche Linearkombinationen von $1, f, f^2, \dots$ approximieren.

Der Schlüssel zum Beweis dieser Behauptung ist der

Hilfssatz A.4. *Auf dem Intervall $[-1, 1]$ existiert eine Folge von Polynomen*

$$p_n : [-1, 1] \rightarrow \mathbb{R},$$

die gleichmäßig gegen die Funktion $|x|$ konvergiert.

Wir werden diese Folge (p_n) explizit angeben. Der Gedankengang ist der folgende: Man kann

$$|x| = \sqrt{1 - (1 - x^2)} \text{ für } -1 \leq x \leq 1$$

schreiben. Setzt man

$$y = 1 - x^2,$$

so variiert y zwischen 0 und 1.

Es genügt also, eine Folge von Polynomen $q_n(y)$ zu konstruieren, die gleichmäßig gegen die Funktion

$$\sqrt{1 - y}$$

konvergiert. Der zugrunde gelegte Definitionsbereich ist hierbei

$$D := \{y : 0 \leq y \leq 1\}.$$

(Anschließend setze man dann $p_n(x) := q_n(1 - x^2)$. Dies sind dann Polynome in x).

Zur Approximation von $\sqrt{1 - y}$ verwendet man die **Taylorreihe**. (Diese haben wir in der [2], Abschnitt 5.2, im Zusammenhang mit der Taylorschen Formel behandelt. Dort haben wir uns allerdings um die Randpunkte des Konvergenzintervalls nicht gekümmert, weshalb wir das Ganze hier noch einmal aufrollen wollen.) Man ermittelt für die Taylorreihe sofort

$$\sum_{\nu=0}^{\infty} \binom{\frac{1}{2}}{\nu} (-1)^\nu y^\nu,$$

wobei der verallgemeinerte Binomialkoeffizient $\binom{\alpha}{\nu}$ definiert ist durch

$$\binom{\alpha}{\nu} = \frac{\alpha(\alpha-1)\dots(\alpha-\nu+1)}{\nu!}$$

Der Konvergenzradius dieser Reihe ist offenbar 1. Es ist aber dennoch nicht von vornherein klar, dass $\sqrt{1-y}$ durch die Reihe dargestellt wird. Jedenfalls wird durch die Reihe eine Funktion

$$f : (-1, 1) \rightarrow \mathbb{R}$$

dargestellt, die, wie man leicht nachrechnet, der Differentialgleichung

$$(1-y)f'(y) = -\alpha f(y)$$

(mit $\alpha = \frac{1}{2}$) genügt. (Gliederweise Differentiation ist im Innern des Konvergenzintervalls erlaubt!)

Hieraus folgert man leicht, dass

$$\left(\frac{f(y)}{(1-y)^\alpha} \right)' = 0$$

gilt. Also ist

$$f(y) = c \cdot (1-y)^\alpha.$$

Durch Spezialisierung ($y = 0$) ermittelt man $c = 1$.

Wir haben also gezeigt: Die Folge der Polynome

$$q_n(y) := \sum_{\nu=0}^n \binom{\frac{1}{2}}{\nu} (-1)^\nu y^\nu$$

konvergiert in $(-1, 1)$ gegen $\sqrt{1-y}$ und zwar ist die Konvergenz gleichmäßig in jedem Teilintervall $[-1 + \varepsilon, 1 - \varepsilon]$ (mit $1 > \varepsilon > 0$, ε beliebig). Mehr ist aus der allgemeinen Theorie der Potenzreihen nicht zu erhalten.

Wir wollten jedoch die Funktion $\sqrt{1-y}$, welche ja in $y = 1$ noch stetig ist, sogar gleichmäßig in $[0, 1]$ approximieren. (Der Punkt $y = 1$ entspricht gerade der Knickstelle $x = 0$ der Funktion $|x|$ und ist damit der Angelpunkt!)

Wir müssen also noch zeigen:

- 1) Die Folge $q_n(y)$ konvergiert auch für $y = 1$ und zwar gegen $\sqrt{1-1} = 0$.

2) Die Konvergenz in $[0, 1]$ ist gleichmäßig.

Beweis: Eine Abzählung der Vorzeichen im Binomialkoeffizienten $\binom{\alpha}{\nu}$ für $\nu \geq 1$ ergibt, dass

$$(-1)^\nu \binom{\frac{1}{2}}{\nu} = - \left| \binom{\frac{1}{2}}{\nu} \right|$$

gilt. Hieraus folgt

$$q_n(x) = 1 - \sum_{\nu=1}^n \left| \binom{\frac{1}{2}}{\nu} \right| y^\nu.$$

Damit ist klar, dass für $0 \leq y \leq 1$ gilt:

$$q_1(y) \geq q_2(y) \geq \dots$$

Außerdem ist

$$\sqrt{1-y} = \lim_{n \rightarrow \infty} q_n(y) \geq 0 \text{ für } 0 \leq y < 1.$$

Hieraus folgt $q_n(y) \geq 0$ für alle n und $0 \leq y < 1$.

Die Funktionen q_n sind stetig in $y = 1$, also folgt sogar

$$q_n(y) \geq 0 \text{ für alle } n \text{ und } 0 \leq y \leq 1.$$

Die Folge $(q_n(1))$ ist also monoton fallend und nach unten (durch 0) beschränkt. Sie konvergiert daher. Das bedeutet nichts anderes, als dass die Reihe

$$\sum_{\nu=0}^{\infty} \left| \binom{\frac{1}{2}}{\nu} \right|$$

konvergiert. Für alle $y \in [0, 1]$ gilt

$$\left| \binom{\frac{1}{2}}{\nu} (-1)^\nu y^\nu \right| \leq \left| \binom{\frac{1}{2}}{\nu} \right|.$$

Nach dem Majorantenkriterium konvergiert also die Folge $(q_n(y))$ daher schon gleichmäßig für $0 \leq y \leq 1$. Die Grenzfunktion

$$\lim_{n \rightarrow \infty} q_n(y)$$

ist daher stetig in $[0, 1]$.¹ Sie stimmt mit $\sqrt{1-y}$ in $[0, 1)$ überein. Da auch letztere Funktion in $y = 1$ stetig ist, müssen die beiden Funktionen auch dort übereinstimmen. Damit ist (A.4) bewiesen. \square

¹Die Stetigkeit folgt auch aus dem Abelschen Grenzwertsatz.

Nun ist der Approximationssatz in der zweiten Variante leicht zu beweisen. Sei etwa $f \in W$. Wir approximieren $|f|$. Sei zunächst

$$g := \frac{f}{\|f\| + 1}.$$

Dann gilt offenbar

$$|g(x)| \leq 1 \text{ für alle } x \in X.$$

Nach (A.4) existiert ein Polynom

$$p_n(x) = a_0 + a_1x + \dots + a_nx^n$$

mit der Eigenschaft

$$|p_n(x) - |x|| < \varepsilon \text{ für alle } x \in [-1, 1].$$

Es folgt insbesondere

$$|p_n(g(x)) - |g(x)|| < \varepsilon \text{ für alle } x \in X.$$

Die Funktion

$$p_n(g(x)) = a_0 + a_1g(x) + \dots + a_n(g(x))^n$$

ist eine Funktion aus W , damit ist gezeigt

$$|g| \in \overline{W} \implies |f| = (\|f\| + 1)|g| \in \overline{W}.$$

Folgerung A.5 (Weierstraßscher Approximationssatz). *Sei*

$$f : [a, b] \rightarrow \mathbb{R} \quad (a < b)$$

stetig. Es existiert eine Folge von Polynomen, die gleichmäßig gegen f konvergiert.

Beweis: Setze $X := [a, b]$ und $W := \{\text{Polynome auf } [a, b]\}$. Dann sind die Voraussetzungen von (A.3) erfüllt. \square

Es ist interessant festzustellen, dass obiger Beweis des Satzes von Stone-Weierstraß (A.3) sich zusammensetzt aus allgemeinen Betrachtungen und einem extremen Spezialfall ($f(x) = |x|$), in dem einmal eine solche Approximation explizit konstruiert werden musste.

Im Vorgriff auf das folgende Kapitel bemerken wir hier folgende Anwendung des Weierstraßschen Approximationssatzes: Es sei f eine stetige Funktion auf dem Rechteck

$$[a, b] \times [c, d] \quad (a < b, \quad c < d).$$

Dann gilt

$$\int_c^d \left(\int_a^b f(x, y) dx \right) dy = \int_a^b \left(\int_c^d f(x, y) dy \right) dx.$$

Beweis: Für Polynome

$$f(x, y) = \sum_{0 \leq i, k \leq n} a_{ik} x^i y^k$$

kann man dies leicht nachrechnen. Aufgrund des Approximationssatzes lässt sich aber jede stetige Funktion gleichmäßig durch Polynome approximieren. \square

Im nächsten Teil der Vorlesung folgen wir im Wesentlichen dem Skript [6] von E. Freitag.

A.2 Das Integral für stetige Funktionen mit kompaktem Träger

Gegeben sei ein abgeschlossener Quader Q im \mathbb{R}^n , welcher durch die n -Tupel

$$a = (a_1, \dots, a_n), \quad b = (b_1, \dots, b_n), \quad a_\nu \leq b_\nu \text{ für } \nu = 1, \dots, n$$

definiert sei:

$$Q = \{x \in \mathbb{R}^n : a_\nu \leq x_\nu \leq b_\nu\}.$$

Es sei eine stetige Funktion

$$f : Q \rightarrow \mathbb{R}$$

gegeben. Wir versuchen, das (mehrfache) Integral von f

$$\int_Q f(x) dx_1 \dots dx_n$$

zu definieren. Dies soll induktiv geschehen. Zunächst kann man bei festem x_2, \dots, x_n das Integral

$$\int_{a_1}^{b_1} f(x) dx_1$$

definieren, da f als Funktion von x_1 stetig und somit integrierbar ist.

Lässt man jetzt x_2, \dots, x_n wieder variieren, so kann man dieses Integral als Funktion von x_2, \dots, x_n auffassen. Definitionsbereich ist der Quader im \mathbb{R}^{n-1} , der durch

$$a_\nu \leq x_\nu \leq b_\nu \quad \text{für } 2 \leq \nu \leq n$$

definiert ist.

Jetzt will man über x_2 integrieren, d.h.

$$\int_{a_2}^{b_2} \left[\int_{a_1}^{b_1} f(x) dx_1 \right] dx_2$$

betrachten, usw. und schließlich soll dann das mehrfache Integral von f über Q durch die Formel

$$\int_Q f(x) dx_1 \dots dx_n = \int_{a_n}^{b_n} \dots \int_{a_1}^{b_1} f(x) dx_1 \dots dx_n$$

definiert werden. Allerdings muss man wissen, dass alle diese Integrationen durchführbar sind.

Hilfssatz A.6. *Die Funktion*

$$f : Q \rightarrow \mathbb{R}, \quad Q = \{x \in \mathbb{R}^n : a_\nu \leq x_\nu \leq b_\nu\} \quad (a_\nu \leq b_\nu)$$

sei stetig. Dann ist auch die Funktion

$$F : Q' \rightarrow \mathbb{R}, \quad Q' = \{(x_2, \dots, x_n) \in \mathbb{R}^{n-1} : a_\nu \leq x_\nu \leq b_\nu\}$$

$$F(x_2, \dots, x_n) = \int_{a_1}^{b_1} f(x_1, x_2, \dots, x_n) dx_1$$

stetig.

Beweis: Wesentliches Hilfsmittel zum Beweis ist der Satz von der gleichmäßigen Stetigkeit (2.14). Versuchen wir zunächst einmal, die Stetigkeit direkt zu beweisen, etwa im Punkt (x_2^0, \dots, x_n^0) . Man muss zeigen

$$|F(x_2, \dots, x_n) - F(x_2^0, \dots, x_n^0)| < \varepsilon \quad \text{für } |x_\nu - x_\nu^0| < \delta(\varepsilon) \quad (2 \leq \nu \leq n).$$

Dabei sei $\varepsilon > 0$ beliebig vorgegeben, also

$$\left| \int_{a_1}^{b_1} [f(x_1, x_2, \dots, x_n) - f(x_1, x_2^0, \dots, x_n^0)] dx_1 \right| < \varepsilon.$$

Nun gilt wegen der Stetigkeit von f tatsächlich

$$|f(x_1, x_2, \dots, x_n) - f(x_1, x_2^0, \dots, x_n^0)| < \varepsilon \quad \text{für } |x_\nu - x_\nu^0| < \delta(\varepsilon)$$

für $2 \leq \nu \leq n$, und man kann abschätzen:

$$|F(x_2, \dots, x_n) - F(x_2^0, \dots, x_n^0)| < \int_{a_1}^{b_1} \varepsilon dx_1 = (b_1 - a_1)\varepsilon$$

für $|x_\nu - x_\nu^0| < \delta(\varepsilon)$ ($2 \leq \nu \leq n$).

Das Auftreten der Konstanten $b_1 - a_1$ störte natürlich in keiner Weise, aber der Beweis funktioniert nur dann, wenn $\delta(\varepsilon)$ unabhängig von x_1 gewählt werden kann.

Dies folgt aber aus dem Satz von der gleichmäßigen Stetigkeit. □

Damit ist folgende Definition möglich:

Definition A.7. *Das Integral einer stetigen Funktion*

$$f : Q \rightarrow \mathbb{R}, \quad Q = \{x : a_\nu \leq x_\nu \leq b_\nu\} \quad (a_\nu \leq b_\nu)$$

auf einem kompakten Quader Q wird induktiv definiert durch

$$\int_Q f(x) dx_1, \dots, dx_n = \int_{Q'} \left[\int_{a_1}^{b_1} f(x_1, x_2, \dots, x_n) dx_1 \right] dx_2 \dots dx_n.$$

($Q' = \{(x_2, \dots, x_n) \in \mathbb{R}^{n-1}; \quad a_\nu \leq x_\nu \leq b_\nu\}$ für $2 \leq \nu \leq n$)

$$\int_{[a,b]} f(x) dx = \int_a^b f(x) dx \quad (a < b) \text{ im Falle } n = 1.$$

Schreibweise:

$$\int_Q f(x) dx = \int_{a_n}^{b_n} \dots \int_{a_1}^{b_1} f(x) dx_1 \dots dx_n.$$

Der Aufbau der Integrationstheorie ist vergleichsweise kompliziert. Für pragmatisch denkende Anwender ist vielleicht der Hinweis nützlich, dass das Integral für stetige Funktionen mit kompaktem Träger einer Veränderlicher im Sinne des Regelintegrals bereits ausreicht, um Volumenberechnungen durchzuführen. Man gehe folgendermaßen vor: Sei $A \subseteq \mathbb{R}^n$ eine einigermaßen anständige, sagen wir kompakte Teilmenge, welche man sich als massiven n -dimensionalen Körper vorstelle. Durchschneidet man diesen Körper mit der „Ebene“ $x_n = t$, so erhält man einen $(n-1)$ -dimensionalen Körper $A(t)$. Man will die Volumentheorie induktiv aufbauen, nimmt also an, dass man weiß, wie man das $(n-1)$ -dimensionale Volumen von $A(t)$ berechnen kann. Die Funktion $h(t) = \text{vol}(A(t))$ verschwindet außerhalb eines geeigneten Intervalls $[-C, C]$. In einigermaßen anständiger Situation darf man erwarten, dass sie eine Regelfunktion ist, so dass man

$$\text{vol}(A) := \int_{-C}^C h(t) dt$$

definieren kann. Aber dieser naive Weg gibt keine direkte Einsicht, warum beispielsweise das Volumen gegenüber Drehungen des Körpers invariant bleibt. Dennoch: Obige Vorgehensweise reicht aus für alle praktischen Belange der Volumenbestimmung und der eine oder andere Hörer mag mit dieser Erkenntnis zufrieden sein und die Mühe des Aufbaus einer leistungsstarken Integrationstheorie als überflüssig erachten. Dies ist legitim, wenn er sich in der Mathematik von der Analysis wegorientiert und beispielsweise eine algebraische Linie bevorzugt. Tieferer Einstieg in Gebiete wie Analysis, Wahrscheinlichkeitstheorie etc. erfordert jedoch eine leistungsstarke Integrationstheorie mit guten Grenzwertsätzen (Vertauschbarkeit von Integration mit Limesbildungen wie beispielsweise Differentiation).

Für die Integrationstheorie ist fundamental, dass man die Reihenfolge der Integrationen vertauschen darf, dass also beispielsweise im Falle $n = 2$

$$\int_{a_2}^{b_2} \int_{a_1}^{b_1} f(x_1, x_2) dx_1 dx_2 = \int_{a_1}^{b_1} \int_{a_2}^{b_2} f(x_1, x_2) dx_2 dx_1$$

gilt. Dies kann man an vielen Beispielen nachprüfen, muss aber streng bewiesen werden. Wir führen dies nochmals durch:

Gegeben sei eine Umordnung (Permutation)

$$\sigma = \{\sigma_1, \sigma_2, \dots, \sigma_n\} \quad \text{der Zahlen } \{1, \dots, n\},$$

d.h. jede Zahl 1 bis n kommt unter den σ_ν genau einmal vor.

Satz A.8. Gegeben sei eine stetige Funktion $f : Q \rightarrow \mathbb{R}$, Q ein kompakter Quader im \mathbb{R}^n und σ eine Permutation der Variablen. Dann gilt

$$\int_{a_n}^{b_n} \dots \int_{a_1}^{b_1} f(x) dx_1 \dots dx_n = \int_{a_{\sigma_n}}^{b_{\sigma_n}} \dots \int_{a_{\sigma_1}}^{b_{\sigma_1}} f(x) dx_{\sigma_1} \dots dx_{\sigma_n}.$$

Es kommt also auf die Reihenfolge der Integrationen nicht an.

Beweis: Die Behauptung ist trivial für Funktionen von dem speziellen Typ

$$f(x) = f_1(x_1) \cdot f_2(x_2) \dots f_n(x_n),$$

wobei

$$f_\nu : [a_\nu, b_\nu] \rightarrow \mathbb{R}$$

stetige Funktionen einer Variablen sind, denn dann zerfällt das Integral

$$\int_Q f(x) dx = \left(\int_{a_1}^{b_1} f_1(x_1) dx_1 \right) \dots \left(\int_{a_n}^{b_n} f_n(x_n) dx_n \right)$$

in ein Produkt von Integralen.

Wir bezeichnen mit A die Menge aller stetigen Funktionen $f : Q \rightarrow \mathbb{R}$, die sich als endliche Summe von Funktionen obigen speziellen Typs schreiben lassen, also

$$f(x) = \sum_{\nu=1}^m f_{1\nu}(x_1) \dots f_{n\nu}(x_n).$$

Es ist klar, dass die Behauptung auch für alle Funktionen dieser Klasse gilt. Satz A.8 wird dann mit Hilfe des *Approximationssatzes von Stone-Weierstraß* bewiesen (siehe A.1).

Zunächst behaupten wir:

Jede stetige Funktion $f : Q \rightarrow \mathbb{R}$ ist gleichmäßiger Grenzwert einer Folge von Funktionen aus A . Dazu weisen wir die Voraussetzungen des Approximationssatzes nach. Der einzige nicht völlig triviale Punkt ist die Punkttrennung. Seien also

$$x^0 = (x_1^0, \dots, x_n^0), \quad y^0 = (y_1^0, \dots, y_n^0)$$

zwei verschiedene Punkte aus Q , also etwa

$$x_\nu^0 \neq y_\nu^0.$$

Die Funktion f mit

$$f(x) = x_\nu - x_\nu^0$$

trennt die beiden Punkte ($f(x^0) = 0$, $f(y^0) \neq 0$) und liegt in A . Da man jede stetige Funktion $f : Q \rightarrow \mathbb{R}$ gleichmäßig approximieren kann durch Funktionen, für die Satz (A.8) schon bewiesen ist, muss man nur noch wissen, dass das mehrfache Integral bezüglich gleichmäßiger Konvergenz stabil ist. \square

Hilfssatz A.9. Sei $Q \subseteq \mathbb{R}^n$ ein kompakter Quader und

$$f : Q \rightarrow \mathbb{R}$$

eine stetige Funktion. Es gilt

$$\left| \int_Q f(x) dx_1 \dots dx_n \right| \leq \|f\| (b_1 - a_1) \dots (b_n - a_n),$$

wobei $\|f\|$ die Maximumsnorm bezeichnet.

Beweis: Aus den Rechenregeln für eine Veränderliche folgt unmittelbar

$$\int_Q f(x) dx_1 \dots dx_n \leq \int_Q g(x) dx_1 \dots dx_n,$$

falls

$$f(x) \leq g(x) \quad \text{für alle } x \in Q$$

gilt. Insbesondere gilt

$$\begin{aligned} \left| \int_Q f(x) dx_1 \dots dx_n \right| &\leq \int_Q |f(x)| dx_1 \dots dx_n \leq \\ &\int_Q \|f\| dx_1 \dots dx_n = \|f\| (b_1 - a_1) \dots (b_n - a_n). \end{aligned}$$

\square

Folgerung A.10. Die Folge von stetigen Funktionen

$$f_k : Q \rightarrow \mathbb{R}, \quad k = 1, 2, 3, \dots,$$

konvergiere gleichmäßig gegen f . Dann gilt

$$\int_Q f(x) dx_1 \dots dx_n = \lim_{k \rightarrow \infty} \int_Q f_k(x) dx_1 \dots dx_n.$$

Beweis:

$$\left| \int_Q f(x) dx_1 \dots dx_n - \int_Q f_k(x) dx_1 \dots dx_n \right| \leq \|f - f_k\| (b_1 - a_1) \dots (b_n - a_n) \rightarrow 0 \quad \text{für } k \rightarrow \infty.$$

□

Man sollte sich gut vor Augen halten:

Die Stabilität des Integrals für stetige Funktionen auf kompakten Quadern ist eine Trivialität, welche auf der banalen Abschätzung A.9 beruht.

Stetige Funktionen mit kompaktem Träger

Unter dem **Träger** einer Funktion

$$f : X \rightarrow \mathbb{R}, \quad X \subseteq \mathbb{R}^n$$

versteht man den Abschluss der Menge aller Punkte, in denen f nicht verschwindet:

$$\text{Träger } f = \overline{\{x \in X; f(x) \neq 0\}}.$$

Ein Punkt $x \in X$ gehört genau dann zum Träger von f , wenn $f(x) \neq 0$ ist, oder wenn es eine Folge von Punkten $x_n \in X$ gibt mit

$$x_n \longrightarrow x \quad \text{für } n \rightarrow \infty \quad \text{und } f(x_n) \neq 0 \text{ für alle } n.$$

Bezeichnung:

Klasse der stetigen Funktionen mit kompaktem Träger auf X :

$$C_c(X) := \{f : X \rightarrow \mathbb{R} : f \text{ stetig, Träger}(f) \text{ ist } \mathbf{kompakt}\}$$

Offenbar gilt

$$f, g \in C_c(X) \Rightarrow f + g, f \cdot g \text{ und } cf \in C_c(X) \quad (c \in \mathbb{R}).$$

Die konstanten Funktionen liegen nur dann in $C_c(X)$, wenn X selbst kompakt ist.

Bemerkung. Eine stetige Funktion

$$f : \mathbb{R}^n \rightarrow \mathbb{R}$$

gehört dann und nur dann zur Klasse $C_C(\mathbb{R}^n)$, wenn eine Zahl $R > 0$ existiert, so dass

$$f(x) = 0 \quad \text{für } \|x\| > R$$

gilt.

Beweis:

1. Wenn der Träger von f kompakt ist, so existiert eine Zahl $R > 0$, so dass

$$\text{Träger}(f) \subseteq K_R(0).$$

(Jedes Kompaktum ist beschränkt.)

2. Wenn ein solches R existiert, so ist der Träger von f beschränkt, außerdem ist er nach Konstruktion abgeschlossen und daher nach dem Überdeckungssatz von Heine-Borel kompakt.

Im Fall $n > 1$ erhält man stetige Funktionen mit kompaktem Träger durch

$$f(x) = f_1(x_1) \cdots f_n(x_n),$$

wobei die $f_i(x_i)$ solche in einer Veränderlichen sind. □

Das Integral für stetige Funktionen mit kompaktem Träger

Es sei

$$f : \mathbb{R}^n \rightarrow \mathbb{R}$$

eine stetige Funktion mit kompaktem Träger. Wir wählen $R > 0$, so dass gilt:

$$f(x) \neq 0 \implies |x_\nu| \leq R \quad \text{für } \nu = 1, \dots, n.$$

Dann definieren wir

$$\int_{\mathbb{R}^n} f(x) dx_1 \cdots dx_n := \int_Q f(x) dx_1 \cdots dx_n,$$

wobei Q den Quader

$$-R \leq x_\nu \leq R \quad \text{für } \nu = 1, \dots, n$$

bezeichne. Es ist klar, dass diese Definition von der Wahl von R nicht abhängt. R muss nur so groß sein, dass der Träger von f in Q liegt.

Das so definierte Integral für stetige Funktionen mit kompaktem Träger ist der Baustein des Lebesgue'schen Integrals, das nun mit Hilfe des Daniell-Lebesgue-Prozesses gewonnen werden soll.

A.3 Die Ausdehnung des Integrals auf halbstetige Funktionen

In A.2 haben wir die Klasse $C_c(\mathbb{R}^n)$ der stetigen Funktionen mit kompaktem Träger eingeführt und darauf ein Integral definiert:

$$f \mapsto \int_{\mathbb{R}^n} f(x) dx_1 \dots dx_n.$$

Wir stellen die Eigenschaften dieses Integrals, die wir im Folgenden benutzen, kurz zusammen.

Dabei benutzen wir die abkürzende Schreibweise

$$I(f) := \int_{\mathbb{R}^n} f(x) dx := \int_{\mathbb{R}^n} f(x) dx_1 \dots dx_n.$$

1) *Das Integral ist ein lineares Funktional, d.h.*

$$I(f + g) = I(f) + I(g), \quad I(cf) = cI(f).$$

2) *Das Integral ist „positiv“, d.h.*

$$I(f) \geq 0, \quad \text{falls } f(x) \geq 0 \text{ für alle } x \in \mathbb{R}^n.$$

Im Folgenden besteht das Problem, das Integral auf eine möglichst große Klasse von Funktionen auszudehnen.

Zunächst haben wir, das Integral für Funktionen einer Veränderlichen schon benutzend, das Integral für stetige Funktionen mit kompaktem Träger gewonnen. Allgemeinere Funktionen f versuchen wir jetzt, durch stetige Funktionen zu approximieren. *Dabei müssen wir allerdings das gelobte Land der gleichmäßigen Konvergenz verlassen, denn sonst kommen wir aus dem Bereich der stetigen Funktionen nicht heraus.*

Der Typ der Konvergenz, den wir betrachten wollen, ist die *monotone* Konvergenz.

A.4 Der Daniell-Lebesgue-Prozess, Teil 1: Das Integral für halbstetige Funktionen

Mit Hilfe der Integrationstheorie mehrerer Variablen will man u.a. mehrdimensionale Volumina von Bereichen im \mathbb{R}^n berechnen.

Sei also

$$f : D \rightarrow \mathbb{R}, \quad D \subseteq \mathbb{R}^n$$

eine Funktion, die der Einfachheit halber nirgends negativ sei,

$$f(x) \geq 0 \quad \text{für alle } x \in D \quad (\text{Schreibweise } f \geq 0).$$

Gesucht ist ein Maß für das Volumen des Bereiches im $(n+1)$ -dimensionalen Raum

$$\{(x, t) \in \mathbb{R}^{n+1}; \quad x \in D, \quad 0 \leq t \leq f(x)\}$$

Dieses Volumen soll gerade

$$\int_D f(x) dx_1 \dots dx_n$$

sein.

Man könnte natürlich wie im Falle $n = 1$ versuchen, das Integral durch Approximationen mit Hilfe von „Treppen“ zu definieren. Das ist möglich, aber schwierig. Im Fall $n = 1$ sind die Definitionsbereiche D in der Regel einfache Intervalle. Diese sind sehr einfach aufzuteilen in „kleine“ Intervalle, auf denen man dann die Treppen aufbaut.

Im Fall $n > 1$ kann schon der Definitionsbereich D relativ kompliziert sein. Man müsste ihn durch eine Quaderaufteilung pflastern (approximativ) und darauf die „Treppen“ aufbauen. Dieser Aufbau ist möglich. Stattdessen haben wir einen anderen Weg eingeschlagen. In sehr naheliegender Weise konnten wir direkt das Integral für stetige Funktionen definieren. Dazu machten wir einen Rückgriff auf das Regelintegral in einer Veränderlichen, das ja sehr leicht (relativ zum Fall $n \geq 1$) zu gewinnen war. Jetzt werden wir das Integral für allgemeinere Funktionentypen f dadurch gewinnen, dass wir f durch stetige Funktionen mit kompaktem Träger zu approximieren suchen.

Dabei müssen wir notgedrungen auf die gleichmäßige Konvergenz verzichten, sonst kommen wir nicht aus dem Bereich der stetigen Funktionen heraus.

Warum ist das Integral für allgemeinere Funktionstypen interessant?

Will man das n -dimensionale Volumen eines Bereiches $D \subseteq \mathbb{R}^n$ definieren und berechnen, so kann man folgendermaßen vorgehen:

Man betrachte die Funktion

$$f(x) = 1 \quad \text{für } x \in D.$$

Die Anschauung zeigt dann, dass die Definition

$$\text{Volumen}(D) = \int_D 1 dx_1 \dots dx_n$$

sinnvoll ist.

(Die Länge des Intervalls $[a, b]$ kann als Integral über die Funktion 1 interpretiert werden, also eindimensionales Volumen von $[a, b]$):

$$\int_a^b dx = b - a,$$

oder: Die Fläche der Kreisscheibe

$$E = \{(x, y) \in \mathbb{R} \times \mathbb{R} : x^2 + y^2 \leq 1\}$$

kann interpretiert werden als Volumen des dreidimensionalen Zylinders

$$Z = \{(x, y, t) \in \mathbb{R}^3 : x^2 + y^2 \leq 1, 0 \leq t \leq 1\},$$

das wäre das Integral

$$\text{Volumen von } Z = \int_{\substack{x^2+y^2 \leq 1 \\ 0 \leq t \leq 1}} dx dy dt.$$

Im folgenden werden wir nur Integrale über den \mathbb{R}^n berechnen, wir nehmen also an, dass die Funktion f auf dem ganzen \mathbb{R}^n definiert ist und streben von vorneherein das uneigentliche Integral

$$\int_{\mathbb{R}^n} f(x) dx_1 \dots dx_n$$

an.

Dies ist keine Einschränkung der Allgemeinheit: wenn eine Funktion

$$f : D \rightarrow \mathbb{R}, \quad D \subseteq \mathbb{R}^n$$

gegeben ist, so betrachten wir einfach $\tilde{f} : \mathbb{R}^n \rightarrow \mathbb{R}$,

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

und definieren

$$\int_D f(x) dx_1 \dots dx_n := \int_{\mathbb{R}^n} \tilde{f}(x) dx_1 \dots dx_n,$$

vorausgesetzt, dass die rechte Seite schon definiert ist. Für die Volumenmessung $\text{Volumen}(D)$ bedeutet dies folgendes: Man betrachte die sogenannte **charakteristische Funktion (Indikatorfunktion)** von D ,

$$\chi_D : \mathbb{R}^n \rightarrow \mathbb{R}, \quad \chi_D(x) = \begin{cases} 1 & \text{für } x \in D, \\ 0 & \text{für } x \notin D, \end{cases}$$

und definiert

$$\text{Volumen}(D) = \int_{\mathbb{R}^n} \chi_D(x) dx_1 \dots dx_n.$$

Damit ist gezeigt, dass das Konzept des Volumens eines Bereiches $D \subseteq \mathbb{R}^n$ sich unter den Begriffsapparat des Integrals

$$\int_{\mathbb{R}^n} f(x) dx_1 \dots dx_n, \quad f : \mathbb{R}^n \rightarrow \mathbb{R}$$

unterordnet, aber man beachte, dass die Funktion χ_D nicht stetig ist (die Randpunkte von D sind Unstetigkeitspunkte). Das Integral für stetige Funktionen mit kompaktem Träger reicht also keineswegs zur Berechnung von Volumina aus.

Wichtigstes Hilfsmittel für den ersten Schritt im Daniell-Lebesgue-Prozess ist

Satz A.11 (Dini). *Es sei X ein kompakter metrischer Raum und*

$$f_n : X \rightarrow \mathbb{R}, \quad n = 1, 2, 3, \dots$$

eine Folge von stetigen Funktionen, die monoton gegen Null fällt, d.h.

- a) $f_1(x) \geq f_2(x) \geq \dots$,
- b) $\lim_{n \rightarrow \infty} f_n(x) = 0$ für jedes $x \in X$.

Die Folge (f_n) konvergiert dann gleichmäßig gegen 0.

Beweis: Sei $\varepsilon > 0$ vorgegeben. Zu jedem $x \in X$ existiert eine natürliche Zahl $N(\varepsilon, x)$ mit der Eigenschaft

$$|f_n(x)| < \varepsilon \quad \text{für } n \geq N(\varepsilon, x).$$

Zu jedem x_0 existiert dann eine offene Umgebung $U(x_0)$ mit der Eigenschaft

$$|f_{N(\varepsilon, x_0)}(x)| < \varepsilon \quad \text{für } x \in U(x_0).$$

Endlich viele dieser Umgebungen überdecken X (Kompaktheit). Wir definieren

$$N = N(\varepsilon) = \max\{N(\varepsilon, x_0); x_0 \in \text{obiger endlicher Menge}\}.$$

Es gilt dann

$$|f_N(x)| < \varepsilon \quad \text{für alle } x.$$

Wegen der Monotonie der Folge $F_n(x)$ gilt sogar

$$|f_n(x)| < \varepsilon \quad \text{für alle } x \text{ und } n \geq N.$$

Das heißt gerade, dass f_n gleichmäßig gegen Null konvergiert. \square

Der Satz von Dini gibt uns die Möglichkeit, das Integral für stetige Funktionen mit kompaktem Träger auf eine große Klasse von Funktionen auszudehnen und zwar auf solche Funktionen f , die sich *monoton* durch eine Folge von Funktionen aus $C_c(\mathbb{R}^n)$ *approximieren* lassen.

Bei der technischen Durchführung der Integrationstheorie hat es sich als zweckmäßig erwiesen, auch Funktionen zuzulassen, die die Werte $\pm\infty$ annehmen dürfen, d.h. wir erweitern die reelle Zahlengerade \mathbb{R} durch Hinzufügen von weiteren Elementen, für die wir ∞ und $-\infty$ schreiben:

$$\overline{\mathbb{R}} := \mathbb{R} \cup \{\infty\} \cup \{-\infty\} \quad (\text{erweiterte Zahlengerade})$$

und vereinbaren die folgenden Rechenregeln:

- a) $\infty > x$ für alle $x \in \mathbb{R} \cup \{-\infty\}$.
- b) $-\infty < x$ für alle $x \in \mathbb{R} \cup \{\infty\}$.

Diese Erweiterung hat folgenden Vorteil: Jede Teilmenge $M \subseteq \overline{\mathbb{R}}$, sofern sie nicht leer ist, besitzt nun eine obere (natürlich auch untere) Grenze, die wir mit $\text{Sup}M$ bezeichnen wollen.²

²Wir haben dieselbe Konvention bereits im Zusammenhang mit der Berechnung des Konvergenzradius einer Potenzreihe verwendet: siehe Analysis I.

Das Supremum $\text{Sup } M$ einer nichtleeren Teilmenge $M \subseteq \overline{\mathbb{R}}$ ist die *kleinste* obere Schranke von M . Im selben Sinne verstehen wir $\text{Inf } M$.

Ist insbesondere $M \subseteq \mathbb{R}$ eine in \mathbb{R} nach oben beschränkte nicht leere Teilmenge, so ist $\text{Sup } M = \sup M$ das gewöhnliche Supremum. Ist hingegen M durch kein Element aus \mathbb{R} nach oben beschränkt, so ist $\text{Sup } M = \infty$.

Ist a_1, a_2, a_3, \dots eine Folge von Elementen aus \mathbb{R} , so verstehen wir unter dem Supremum dieser Folge einfach das Supremum der Menge der Folgenglieder:

$$\text{Sup}(a_n) = \text{Sup}\{a_1, a_2, \dots\}$$

Die erweiterte Zahlengerade hat allerdings den Nachteil, dass es unmöglich ist, die algebraischen Rechenregeln $(+, \cdot)$ auf \mathbb{R} so zu erweitern, dass die üblichen Gesetze erfüllt sind (Kommutativ-, Assoziativ- und Distributivgesetz). Zu keinen Widersprüchen führt die Konvention

$$\begin{aligned} \infty + x &= \infty && \text{für alle } x \in \mathbb{R} \cup \{\infty\}, \\ -\infty + x &= -\infty && \text{für alle } x \in \mathbb{R} \cup \{-\infty\}, \\ x \cdot \infty &= \infty && \text{für alle } x \in \mathbb{R}, x > 0, \\ x \cdot (-\infty) &= -\infty && \text{für alle } x \in \mathbb{R}, x > 0, \\ \infty \cdot \infty &= \infty, \\ \infty \cdot (-\infty) &= -\infty, \\ (-\infty) \cdot (-\infty) &= \infty, \end{aligned}$$

wovon sich der Leser (am besten dort wo es verwendet wird) überzeugen mag. **Nicht definiert werden jedoch:** $+\infty - \infty$ und $-\infty + \infty$.

Wir betrachten nun Funktionen, die man monoton durch stetige Funktionen mit kompaktem Träger approximieren kann.

Definition A.12. *Eine Funktion*

$$f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$$

gehört der Klasse B^+ an (Bairesche Klasse), wenn es eine Folge von stetigen Funktionen mit kompaktem Träger

$$f_\nu : \mathbb{R}^n \rightarrow \mathbb{R}$$

mit den Eigenschaften

$$a) f_1(x) \leq f_2(x) \leq f_3(x) \leq \dots$$

b) $f(x) = \text{Sup}_\nu f_\nu(x)$ für alle $x \in \mathbb{R}^n$

gibt.

Im Folgenden schreiben wir einfach $f_n \uparrow f$, wenn die Eigenschaften a) und b) erfüllt sind.

(Die Bezeichnung $f_n \downarrow f$ versteht sich von selbst.)

Es ist klar, dass die Funktionen aus B^+ den Wert ∞ annehmen können, aber auf keinen Fall $-\infty$. Man kann Funktionen aus der Klasse B^+ addieren, ohne diese Klasse zu verlassen, aber wenn f in B^+ enthalten ist, braucht $-f$ noch lange nicht in B^+ enthalten zu sein. Die Menge B^+ ist also kein Vektorraum. Immerhin gilt noch

$$f \in B^+, C \geq 0 \implies Cf \in B^+.$$

Bemerkung A.13. *Es seien $f \in B^+$ und $(f_\nu), (g_\nu)$ zwei Folgen von Funktionen aus C_c (stetige Funktionen mit kompaktem Träger) mit der Eigenschaft*

$$f_\nu \uparrow f, \quad g_\nu \uparrow f.$$

Dann gilt

$$\text{Sup}_\nu \left(\int_{\mathbb{R}^n} f_\nu(x) dx \right) = \text{Sup}_\nu \left(\int_{\mathbb{R}^n} g_\nu(x) dx \right).$$

Diese Bemerkung - wir werden sie gleich beweisen - gibt dann Anlass zu

Definition A.14. *Es sei $f \in B^+$. Das Integral von f wird durch die Formel*

$$\int_{\mathbb{R}^n} f(x) dx = \text{Sup}_\nu \left(\int_{\mathbb{R}^n} f_\nu(x) dx \right)$$

definiert, wobei $f_\nu \in C_c(\mathbb{R}^n)$ irgend eine Folge von Funktionen (stetig mit kompaktem Träger) ist, die f monoton wachsend approximiert, d.h. $f_\nu \uparrow f$.

Die Bemerkung A.13 besagt gerade, dass diese Definition unabhängig von der Wahl der Folge (f_ν) ist. Außerdem beachte man, dass eine Funktion $f \in C_c$ auch in B^+ liegt, man kann sie durch die konstante Folge $f, f, f \dots$ monoton wachsend approximieren.

Also: Es gilt $C_c \subseteq B^+$ und das durch A.14 definierte Integral stimmt auf C_c mit dem früher definierten Integral (A.2) überein.

Wir erinnern an die Bezeichnungen für die Funktionen

$$\begin{aligned} & f \vee g \text{ und } f \wedge g, \text{ die durch} \\ (f \vee g)(x) &= \max(f(x), g(x)), \\ (f \wedge g)(x) &= \min(f(x), g(x)), \end{aligned}$$

definiert sind.

Es gilt: $f, g \in C_c \implies f \vee g \text{ und } f \wedge g \in C_c$.

Wir zeigen zunächst folgendes:

Sei

$$f_\nu \uparrow f, \quad f_\nu \text{ stetig mit kompaktem Träger,}$$

und sei g irgendeine stetige Funktion mit kompaktem Träger mit der Eigenschaft

$$f \geq g.$$

Dann ist

$$\text{Sup} \int_{\mathbb{R}^n} f_\nu(x) dx \geq \int_{\mathbb{R}^n} g(x) dx.$$

Beweis von A.13: Die Folge $g - (f_\nu \wedge g)$ fällt offenbar monoton gegen Null. Nach dem *Satz von Dini* konvergiert sie daher gleichmäßig gegen Null, und da das Integral stabil gegenüber gleichmäßiger Konvergenz ist (A.10), gilt

$$\lim_{\nu \rightarrow \infty} \int_{\mathbb{R}^n} (f_\nu \wedge g) dx = \int_{\mathbb{R}^n} g(x) dx.$$

Nun ist

$$f_\nu \wedge g \leq f_\nu \quad \text{für alle } \nu,$$

es folgt daher

$$\int_{\mathbb{R}^n} g(x) dx \leq \text{Sup}_\nu \int_{\mathbb{R}^n} f_\nu(x) dx.$$

Es sei nun

$$g_\nu \uparrow f, \quad g_\nu \in C_c.$$

Nach dem, was oben bewiesen wurde, gilt

$$\int_{\mathbb{R}^n} g_\mu(x) dx \leq \text{Sup}_\nu \int_{\mathbb{R}^n} f_\nu(x) dx \quad \text{für jedes } \mu.$$

Insbesondere gilt

$$\operatorname{Sup}_\nu \int_{\mathbb{R}^n} g_\nu(x) dx \leq \operatorname{Sup}_\nu \int_{\mathbb{R}^n} f_\nu(x) dx.$$

Die umgekehrte Ungleichung gilt genauso, da man die Rollen von f_k und g_k vertauschen kann. \square

Dieser Beweis zeigt in Wirklichkeit noch etwas mehr, nämlich

Hilfssatz A.15. *Seien $f \leq g$ zwei Funktionen aus B^+ , dann gilt*

$$\int_{\mathbb{R}^n} f(x) dx \leq \int_{\mathbb{R}^n} g(x) dx.$$

Rechenregeln für die Bairesche Klasse

Wir nennen manchmal die Funktionen aus B^+ auch (unter-) halbstetig. Das hat seine Berechtigung in folgender Eigenschaft der Funktionen aus B^+ , die für stetige Funktionen wohlbekannt ist:

Bemerkung A.16. *Seien $f \in B^+$, C eine reelle Zahl und $a \in \mathbb{R}^n$ ein Punkt mit $f(a) > C$. Dann gilt*

$$f(x) > C \quad \text{in einer vollen Umgebung von } a.$$

Beweis: Sei

$$f_\nu \uparrow f, f_\nu \in C_c.$$

Nach Definition des Supremums existiert ein Index l mit

$$f(a) \geq f_l(a) > C.$$

Für die stetige Funktion f_l stimmt aber die Behauptung und damit erst recht für f . Allgemein nennt man Funktionen mit der in Bemerkung A.16 genannten Eigenschaft **unterhalbstetig** (oder **halbstetig nach unten**). Man kann umgekehrt zeigen, dass jede unterhalbstetige Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ mit der Eigenschaft $f(x) \geq 0$ außerhalb eines Kompaktums zu B^+ gehört. \square

Als nächstes zeigen wir, dass das Integral für Funktionen aus B^+ stabil gegenüber monotoner Approximation ist.

Hilfssatz A.17. Sei

$$f_1 \leq f_2 \leq f_3 \leq \dots$$

eine monotone Folge aus B^+ . Dann gilt

a) $f = \text{Sup} f_\nu \in B^+$,

b) $\int_{\mathbb{R}^n} f(x) dx = \text{Sup} \int_{\mathbb{R}^n} f_\nu(x) dx.$

Beweis: Sei

$$f_k \uparrow f, \quad f_k \in B^+ \quad (\text{Bairesche Klasse}).$$

Es existiert also

$$f_{1k} \leq f_{2k} \leq f_{3k} \dots \quad \text{mit } f_k = \text{Sup}_i f_{ik}.$$

Wir bilden die Funktionenfolge

$$g_r = \bigvee_{i+k \leq r} f_{ik} \quad (\text{also } g_r(x) = \max_{i+k \leq r} (f_{ik}(x))).$$

Offenbar gilt $g_r \in C_c$.

Außerdem gelten die Ungleichungen

$$g_1 \leq g_2 \leq g_3 \leq \dots \leq f.$$

Hieraus folgt

$$g = \text{Sup } g_k \leq f.$$

Beachtet man außerdem

$$f_{ik} \leq g_{i+k} \leq g \quad \text{für alle } i, k,$$

so folgt

$$f_k \leq g \quad \text{für alle } k = 1, 2, \dots$$

und daher

$$f = \text{Sup} f_k \leq g,$$

insbesondere also $f = g$.

Damit ist gezeigt:

$$g_k \uparrow f, \quad \text{also } f \in B^+.$$

Außerdem gilt

$$\int_{\mathbb{R}^n} f(x) dx = \text{Sup}_k \int_{\mathbb{R}^n} g_k(x) dx.$$

Aus den Ungleichungen

$$g_k \leq f_k \quad \text{für alle } k$$

folgt (A.15)

$$\int_{\mathbb{R}^n} g(x) dx \leq \int_{\mathbb{R}^n} f_k(x) dx \leq \int_{\mathbb{R}^n} f(x) dx$$

und hiermit

$$\text{Sup} \int_{\mathbb{R}^n} f_k(x) dx = \int_{\mathbb{R}^n} f(x) dx. \quad \square$$

□

Hilfssatz A.18. Seien $f, g \in B^+$ und a, b nicht negative Zahlen, $a \geq 0$, $b \geq 0$. Dann gilt $af + bg \in B^+$ und

$$\int_{\mathbb{R}^n} (af(x) + bg(x)) dx = a \int_{\mathbb{R}^n} f(x) dx + b \int_{\mathbb{R}^n} g(x) dx.$$

Beweis: Seien

$$f_k \uparrow f, \quad g_k \uparrow g, \quad f_k, g_k \in C_c.$$

Offenbar gilt

$$af_k + bg_k \uparrow af + bg \quad (\text{beachte: } a > 0, b > 0!)$$

und die Behauptung ist evident. □

Hier sieht man auch den wesentlichen Nachteil des Integrals für B^+ : Leider gilt im Allgemeinen nicht

$$f \in B^+ \implies -f \in B^+.$$

A.5 Der Daniell-Lebesgue-Prozess, Teil 2: Das äußere Integral

Wir ordnen nun **jeder** Funktion $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ ein gewisses äußeres Integral zu, das wir mit

$$\int^- f(x) dx_1 \dots dx_n = \int^- f(x) dx$$

bezeichnet werden. Zur Motivation nehmen wir einige Eigenschaften vorweg.

- 1) Wenn f in der Klasse B^+ liegt, dann stimmt das äußere Integral mit dem früher definierten Integral überein.
- 2) Das äußere Integral ist im Allgemeinen nicht additiv, es gilt z.B. im Allgemeinen nicht

$$\int^- f(x)dx = - \int^- (-f(x))dx.$$

Man nennt manchmal auch

$$\int_- f(x)dx := - \int^- (-f(x))dx$$

das *innere Integral* von $f(x)$.

Eine Funktion $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ heißt (im Lebesgue'schen Sinn) integrierbar, wenn äußeres und inneres Integral von f übereinstimmen und wenn diese einen endlichen Wert ($\neq \infty, -\infty$) annehmen.

Die Klasse $\mathcal{L}^1(\mathbb{R}^n)$ der Lebesgue-integrierbaren Funktionen hat dann alle die Eigenschaften, die man von einem „vernünftigen“ Integral erwartet. Diese werden dann in A.5 formuliert und bewiesen.

Definition A.19. Das *äußere Integral (oder Oberintegral)* einer Funktion

$$f : \mathbb{R}^n \rightarrow \mathbb{R}$$

wird durch die Formel

$$\int^- f(x)dx = \text{Inf}_g \left\{ \int_{\mathbb{R}^n} g(x)dx \right\}$$

definiert, wobei $g : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ alle Funktionen der Klasse B^+ mit der Eigenschaft

$$g(x) \geq f(x) \quad \text{für alle } x \in \mathbb{R}^n$$

durchläuft.

Fast unmittelbar aus der Definition kann man einige Eigenschaften des äußeren Integrals ableiten, die für das Folgende wichtig sind.

Zunächst wollen wir noch klarstellen, dass das Oberintegral überhaupt wohldefiniert ist. Da wir die Werte ∞ und $-\infty$ zulassen, muss dazu nur gezeigt werden, dass die Menge der Funktionen

$$g \in B^+ \text{ mit } g \geq f \quad (\text{d.h. } g(x) \geq f(x) \quad \text{für alle } x \in \mathbb{R}^n)$$

nicht leer ist. Dazu beachten wir einfach

Hilfssatz A.20. *Die Funktion*

$$f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}, \quad f(x) = \infty \quad \text{für alle } x \in \mathbb{R}^n$$

gehört der Klasse B^+ an.

Wir deuten den Beweis nur im Falle $n = 1$ an. Es ist dann sehr einfach, ihn auf den Fall $n > 1$ zu verallgemeinern. Die Funktion f kann approximiert werden durch die Folge von „Dreiecken“

$$f_k(x) = \begin{cases} -|x| + k & \text{für } |x| \leq k, \\ 0 & \text{für } |x| \geq k. \end{cases}$$

Damit ist also das äußere Integral wohldefiniert. Das äußere Integral ist erfreulicherweise ordnungstreu.

Hilfssatz A.21. *Es seien zwei Funktionen*

$$f, h : \mathbb{R}^n \rightarrow \overline{\mathbb{R}} \quad \text{mit } f \leq h$$

gegeben. Dann gilt für das äußere Integral

$$\int^- f(x) dx \leq \int^- h(x) dx.$$

Beweis: Ist $g \in B^+$ eine Funktion der Baireschen Klasse mit der Eigenschaft $g \geq h$, so gilt erst recht $g \geq f$. Bei der Definition des Oberintegrals h werden also weniger Funktionen zur Konkurrenz zugelassen als bei f . Dieses wird daher höchstens größer.

Für Funktionen aus der Baireschen Klasse B^+ bringt das äußere Integral nichts Neues. \square

Hilfssatz A.22. *Für jede Funktion $f \in B^+$ gilt*

$$\int_{\mathbb{R}^n} f(x) dx = \int^- f(x) dx = - \int^- (-f(x)) dx.$$

Beweis: 1. Teil: Die Gleichung

$$\int_{\mathbb{R}^n} f(x) dx = \int^- f(x) dx$$

ist klar, denn dann ist sogar

$$\int_{\mathbb{R}^n} f(x) dx = \min_{g \in B^+, g \geq f} \left\{ \int_{\mathbb{R}^n} g(x) dx \right\},$$

da unter den Funktionen $g \in B^+, g \geq f$ die Funktion f selbst vorkommt.

2. Teil: Die Ungleichung

$$-\int^- (-f(x)) dx \geq \int_{\mathbb{R}^n} f(x) dx :$$

(Beachte: wenn $f \in B^+$ enthalten ist, so braucht dies nicht für $-f$ zuzutreffen.)

Sei (f_ν) eine Folge von stetigen Funktionen mit kompaktem Träger, die f monoton wachsend approximiert

$$f_\nu \in C_c, \quad f_\nu \uparrow f.$$

Dann gilt

$$-f_\nu \geq -f,$$

also

$$\int_{\mathbb{R}^n} (-f_\nu)(x) dx \geq \int^- (-f)(x) dx$$

oder

$$-\int^- (-f(x)) dx \geq \int_{\mathbb{R}^n} f_\nu(x) dx.$$

Da dies für alle ν gilt, folgt die behauptete Ungleichung.

3. Teil: Die Ungleichung

$$\int^- (-f(x)) dx \geq - \int_{\mathbb{R}^n} f(x) dx.$$

Nach Definition des Oberintegrals als ein Infimum bedeutet dies nichts anderes als

$$\int_{\mathbb{R}^n} g(x) dx \geq - \int_{\mathbb{R}^n} f(x) dx,$$

wobei g alle Funktionen

$$g \in B^+ \quad \text{mit} \quad -f \leq g$$

durchläuft.

Dies wiederum ist äquivalent zu

$$\int_{\mathbb{R}^n} g(x)dx + \int_{\mathbb{R}^n} f(x)dx = \int_{\mathbb{R}^n} (g(x) + f(x))dx \geq 0,$$

was aber wegen $g + f \geq 0$ trivial ist. □

Die Gleichung

$$\int^- f(x)dx = - \int^- (-f(x))dx$$

ist für beliebige Funktionen f nicht richtig, wie komplizierte Gegenbeispiele zeigen, auf die wir hier nicht eingehen wollen.

Bezeichnung A.23 (Inneres Integral oder Unterintegral).

$$\int_- f(x)dx := - \int^- (-f(x))dx.$$

Allgemein gilt noch die folgende Ungleichung

Hilfssatz A.24. Für jede Funktion $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ gilt

$$\int_- f(x)dx \leq \int^- f(x)dx.$$

Beweis: Diese Ungleichung ist äquivalent zu

$$\int^- f(x)dx + \int^- (-f(x))dx \geq 0,$$

wenn man die beiden Fälle

$$\int^- f(x)dx = \pm\infty \quad \text{und gleichzeitig} \quad \int^- (-f(x))dx = \mp\infty$$

ausschließt. In diesen Ausnahmefällen gilt sogar $\int^- = \int_-$. Nach Definition des Oberintegrals als ein Infimum ist diese Behauptung äquivalent zu

$$\int_{\mathbb{R}^n} g(x)dx + \int_{\mathbb{R}^n} h(x)dx \geq 0$$

für alle halbstetigen Funktionen

$$g, h \in B^+, \quad \text{mit } g \geq f, h \geq -f.$$

Hieraus folgt $g + h \geq 0$ und somit ist alles klar, denn für halbstetige Funktionen aus B^+ gilt ja

$$\int_{\mathbb{R}^n} g(x)dx + \int_{\mathbb{R}^n} h(x)dx = \int_{\mathbb{R}^n} [g(x) + h(x)]dx.$$

□

Definition A.25. Eine Funktion $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ heißt integrierbar³, wenn

$$\int^- f(x)dx = \int_- f(x)dx$$

gilt und wenn dieser Wert endlich ist (also $\neq \infty, -\infty$).

Bezeichnung: Wenn eine Funktion f integrierbar ist, so setzen wir

$$\int_{\mathbb{R}^n} f(x)dx = \int^- f(x)dx = \int_- f(x)dx.$$

Diese Bezeichnung ist gerechtfertigt (Hilfssatz A.21).

Eine offensichtliche, den Begriff des Supremums umgehende Umformulierung des Begriffs der Integrierbarkeit ist:

Bemerkung A.26. Eine Funktion $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ ist genau dann integrierbar, wenn es zu jedem $\varepsilon > 0$ zwei Funktionen g, h der Baireschen Klasse mit folgenden beiden Eigenschaften gibt:

a) $g(x) \geq f(x), \quad h(x) \geq -f(x),$

b) $\int_{\mathbb{R}^n} (g + h)dx \leq \varepsilon.$

Aus dieser Umformulierung leiten wir eine wichtige Konsequenz ab.

³im Sinne von Lebesgue

Satz A.27. Sei f eine integrierbare Funktion. Dann sind auch die Funktionen f^+ und f^- integrierbar.

Wir erinnern an die Bezeichnungen

$$f^+(x) = \max(f(x), 0), \quad f^-(x) = -\min(f(x), 0).$$

Beweis: Wenn f eine Funktion der Baireschen Klasse ist, so trifft dies auch für die beiden Funktionen f^+ und $-f^-$ zu (weil der entsprechende Sachverhalt für stetige Funktionen mit kompaktem Träger gilt und weil die Bildungen ordnungstreu sind). Mit den Bezeichnungen von A.26 gilt $g^+(x) \geq f^+(x)$ und $-h^-(x) \geq -(-f)^-(x) = -f^+(x)$. Außerdem gilt $g^+(x) - h^-(x) \leq g(x) + h(x)$. Diese Ungleichung ist klar, wenn $g(x) \geq 0$, denn dann ist $g(x) = g^+(x)$ und außerdem gilt stets $-h^-(x) \leq h(x)$. Im Falle $g(x) < 0$ gilt $f(x) < 0$ und $g(x)$ und $h^-(x)$ sind beide 0. Andererseits ist $g(x) + h(x) \geq 0$ für alle x (auch in den Unendlichkeitsstellen von f). \square

A.6 Die integrierbaren Funktionen

In diesem und im nächsten Paragraphen wird sich zeigen, dass das Lebesgue-Integral alle Eigenschaften hat, die man von einem „vernünftigen“ Integral erwartet.

Im Folgenden werden wir häufig nur noch Funktionen $f : \mathbb{R}^n \rightarrow \mathbb{R}$ betrachten, die also nur endliche Werte annehmen. Die Werte $+\infty$ und $-\infty$ haben nur noch für die technische Durchführung der Theorie Bedeutung. Außerdem werden wir noch zeigen (vgl. A.7), dass eine *integrierbare* Funktion $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ in den Unendlichkeitsstellen (das sind Stellen $x \in \mathbb{R}^n$ mit $f(x) = +\infty$ oder $-\infty$) beliebig abgeändert werden kann, ohne dass der Wert des Integrals verändert wird. Die Unendlichkeitsstellen einer integrierbaren Funktion haben kein positives Volumen.

Bezeichnung. Menge der integrierbaren Funktionen ohne Unendlichkeitsstellen:

$$\mathcal{L}^1 = \mathcal{L}^1(\mathbb{R}^n) = \{f : \mathbb{R}^n \rightarrow \mathbb{R} : f \text{ integrierbar}\}.$$

Satz A.28. Die integrierbaren Funktionen aus \mathcal{L}^1 bilden einen Vektorraum und das Integral ist ein lineares Funktional, d.h. also:

$$f, g \in \mathcal{L}^1 \implies f + g \in \mathcal{L}^1 \text{ und } cf \in \mathcal{L}^1 \text{ f\u00fcr } c \in \mathbb{R},$$

au\u00dferdem

$$\int_{\mathbb{R}^n} (f(x) + g(x)) dx = \int_{\mathbb{R}^n} f(x) dx + \int_{\mathbb{R}^n} g(x) dx$$

und

$$\int_{\mathbb{R}^n} cf(x) dx = c \int_{\mathbb{R}^n} f(x) dx.$$

Der Beweis der Additivit\u00e4t beruht auf einer Ungleichung f\u00fcr das \u00e4u\u00dfere Integral:

Hilfssatz A.29. Es seien

$$f, g, h : \mathbb{R}^n \longrightarrow \overline{\mathbb{R}}$$

drei Funktionen mit der Eigenschaft

$$f(x) + g(x) = h(x),$$

falls die Summe $f(x) + g(x)$ wohldefiniert ist. (Wenn also $f(x) = \infty, g(x) = -\infty$ oder $f(x) = -\infty, g(x) = \infty$ gilt, so wird nichts gefordert, denn dann ist die Summe $f(x) + g(x)$ nicht definiert. Es ist dann gleichg\u00fcltig, was $h(x)$ f\u00fcr einen Wert annimmt.)

Dann gilt

$$\int^- f(x) dx + \int^- g(x) dx \geq \int^- h(x) dx,$$

falls die Summe auf der linken Seite definiert ist. F\u00fcr das Unterintegral gilt eine entsprechende Ungleichung in der anderen Richtung.

Im Spezialfall $g = -f, h = 0$ ist dies nichts anderes als Hilfssatz A.22. Der Beweis von A.29 erfolgt in Analogie zu dem von A.22.

Beweis: Man muss zeigen, dass f\u00fcr alle halbstetigen Funktionen

$$f^*, g^* \in B^+ \quad \text{mit } f^* \geq f, g^* \geq g$$

gilt

$$\int_{\mathbb{R}^n} f^*(x)dx + \int_{\mathbb{R}^n} g^*(x)dx \geq \int_{\mathbb{R}^n} h(x)dx.$$

Dies ist aber trivial, denn es gilt $f^* + g^* \geq h$ (auch in den Stellen, in denen $f(x) + g(x)$ nicht definiert ist!)

Der Beweis von A.28 ist eine triviale Folgerung aus A.29. Es gilt sogar mehr als wir in A.28 formuliert haben. Die in A.28 formulierte Additivität gilt auch für integrierbare Funktionen f, g mit eventuellen Unendlichkeitsstellen. Definiert man die Funktion h durch die Bedingungen $h(x) = f(x) + g(x)$ in allen Stellen x , in denen die Summe definiert ist und definiert man $h(x)$ an allen anderen Stellen beliebig - etwa $= 0$, so folgt aus A.29, dass auch h integrierbar ist und dass das Integral von h gleich der Summe der Integrale von f und g ist. In der Regel kommen wir mit der glatteren Variante A.28 aus. In Beweisen werden wir jedoch gelegentlich auf diese durch A.29 gegebene Verschärfung zurückgreifen.

Noch einfacher sieht man, dass

$$f \in \mathcal{L}^1 \Rightarrow cf \in \mathcal{L}^1 \text{ für } c \in \mathbb{R}$$

und

$$\int_{\mathbb{R}^n} cf(x)dx = c \int_{\mathbb{R}^n} f(x)dx \text{ für positive } c > 0.$$

Dies folgt unmittelbar aus der Definition des äußeren Integrals, wenn man beachtet, dass die entsprechende Gleichung für halbstetige Funktionen aus B^+ gilt. Der Übergang zu negativen c bei integrierbaren Funktionen f erfolgt über die Rechenregel

$$\int_{\mathbb{R}^n} (-f(x))dx = - \int_{\mathbb{R}^n} f(x)dx.$$

□

Satz A.30. *Ist $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ eine integrierbare Funktion, so ist auch ihr Betrag $|f|$ integrierbar. Ist $g : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ eine weitere integrierbare Funktion, so sind auch die Maxima und Minima $f \vee g$ und $f \wedge g$ integrierbar.*

Zum Beweis der ersten Aussage verwende man die Formel $f(x) = f^+(x) - f^-(x)$. Wir wissen, dass die Funktionen f^\pm integrierbar sind (A.27). Die Behauptung folgt aus der Linearität des Integrals, wobei man die verschärfte Fassung A.20 benötigt. Die beiden restlichen Aussagen beweist man analog.

Die Stärke des Lebesgue-Integrals liegt in seiner Stabilität gegenüber Grenzprozessen.

Satz A.31 (Beppo Levi). *Gegeben sei eine monoton wachsende Folge*

$$f_1 \leq f_2 \leq f_3 \leq \dots$$

von integrierbaren Funktionen aus $\mathcal{L}^1(\mathbb{R}^n)$. Die Folge der Integrale

$$\int_{\mathbb{R}^n} f_\nu(x) dx$$

sei beschränkt. Die durch

$$f(x) = \sup_{\nu \in \mathbb{N}} f_\nu(x)$$

definierte Funktion $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ ist integrierbar und es gilt

$$\int_{\mathbb{R}^n} f(x) dx = \lim_{\nu \rightarrow \infty} \int_{\mathbb{R}^n} f_\nu(x) dx.$$

Beweis: Wir betrachten die Funktionen

$$h_\nu = f_\nu - f_{\nu-1} \geq 0 \quad (f_0 = 0).$$

Es gilt

$$\sum_{\nu=1}^k h_\nu = f_k.$$

Da jede monotone und beschränkte Folge konvergiert, existieren die Grenzwerte

$$R = \sum_{\nu=1}^{\infty} \int_{\mathbb{R}^n} h_\nu(x) dx = \lim_{k \rightarrow \infty} \int_{\mathbb{R}^n} f_k(x) dx.$$

1) Es gilt

$$R \leq \int_{-} f(x) dx.$$

Beweis: Man beachte die Ungleichung (A.30)

$$\int_{\mathbb{R}^n} f_k(x) dx \leq \int_{-} f(x) dx \quad \text{für } k = 1, 2, \dots$$

2) Es gilt

$$\int^- f(x) dx \leq R$$

Der Beweis folgt offenbar aus der Ungleichung

$$\int^- f(x) dx \leq \sum_{\nu=1}^{\infty} \int^- h_{\nu}(x) dx,$$

welche es nun zu beweisen gilt.

Nach Definition des äußeren Integrals können wir halbstetige Funktionen

$$\bar{h}_{\nu} \in B^+ \quad \text{mit} \quad \bar{h}_{\nu} \geq h_{\nu}$$

und

$$\int_{\mathbb{R}^n} \bar{h}_{\nu} dx \leq \int_{\mathbb{R}^n} h_{\nu}(x) dx + \frac{\varepsilon}{2^{\nu}}$$

finden ($\varepsilon > 0$ beliebig vorgegeben).

Wir setzen

$$\bar{f} = \sup_k \sum_{\nu=1}^k \bar{h}_{\nu} \geq f.$$

Da das Integral für halbstetige Funktionen stabil gegenüber monotoner Konvergenz ist, folgt

$$\begin{aligned} \int^- f(x) dx &\leq \int_{\mathbb{R}^n} \bar{f}(x) dx = \\ &\sum_{\nu=1}^{\infty} \int_{\mathbb{R}^n} \bar{h}_{\nu}(x) dx \leq \sum_{\nu=1}^{\infty} \left\{ \int_{\mathbb{R}^n} h_{\nu}(x) dx + \frac{\varepsilon}{2^{\nu}} \right\}. \end{aligned}$$

Beachtet man

$$\sum_{\nu=1}^{\infty} \frac{\varepsilon}{2^{\nu}} = \varepsilon \quad (\text{geometrische Reihe}),$$

so folgt

$$\int^- f(x) dx \leq \sum_{\nu=1}^{\infty} \int_{\mathbb{R}^n} h_{\nu}(x) dx + \varepsilon.$$

Da dies für alle $\varepsilon > 0$ gilt, ist die Behauptung bewiesen. □

Satz A.32 (Lebesgue'scher Grenzwertsatz). *Gegeben sei eine Folge*

$$f_k : \mathbb{R}^n \rightarrow \mathbb{R}, \quad k = 0, 1, 2, \dots,$$

*von integrierbaren Funktionen, die **punktweise** gegen eine Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ konvergiert.*

Es existiere eine Funktion $h : \mathbb{R}^n \rightarrow \mathbb{R}$ mit den Eigenschaften

a) $|f_k| \leq h$ für $k = 1, 2, 3, \dots$,

b) $\int^- h(x) dx < \infty$.

Dann ist auch f integrierbar und es gilt

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

Beweis: Aus den Ungleichungen $|f_k| \leq h$ folgt $f \leq h$. Daher ist das äußere Integral von $|f|$ endlich:

$$\int^- |f(x)| dx < \infty.$$

Insbesondere sind daher äußeres und inneres Integral von f endlich ($\neq \pm\infty$). Aus $f \leq |f|$ und $-f \leq |f|$ folgt nämlich

$$\int^- f(x) dx < \infty \quad \text{und} \quad \int^- (-f(x)) dx < \infty,$$

also

$$-\infty < \int_- f(x) dx \leq \int^- f(x) dx < \infty.$$

Wir wollen den Lebesgue'schen Grenzwertsatz auf den Satz von Beppo Levi zurückführen und bilden hierzu

$$\begin{aligned} \mathfrak{g}_k(x) &= \text{Inf}\{f_\nu(x), \nu \geq k\}, \\ \bar{g}_k(x) &= \text{Sup}\{f_\nu(x), \nu \geq k\}. \end{aligned}$$

Dann gilt offenbar

$$\begin{aligned} \mathfrak{g}_1 &\leq \mathfrak{g}_2 \leq \mathfrak{g}_3 \leq \dots, \\ \bar{g}_1 &\geq \bar{g}_2 \geq \bar{g}_3 \geq \dots \end{aligned}$$

Aus $f = \lim_{k \rightarrow \infty} f_k$ folgert man leicht

$$\mathfrak{g}_k \uparrow f, \quad \bar{g}_k \downarrow f.$$

Als nächstes wird gezeigt, dass die \mathfrak{g}_k und \bar{g}_k integrierbare Funktionen sind. Dazu wird der Satz von Beppo Levi ausgenutzt. Bildet man nämlich

$$\underline{G}_{kj} = f_k \wedge f_{k+1} \wedge \dots \wedge f_{k+j}$$

und

$$\overline{G}_{kj} = f_k \vee f_{k+1} \vee \dots \vee f_{k+j}$$

so sind \underline{G}_{kj} und \overline{G}_{kj} integrierbar (A.21) und es gelten die Ungleichungen

$$-h \leq \underline{G}_{kj} \leq f_k \leq \overline{G}_{kj} \leq h.$$

Ferner gilt offensichtlich

$$\underline{G}_{kj} \downarrow \mathfrak{g}_k \quad (j \rightarrow \infty)$$

und

$$\overline{G}_{kj} \uparrow \bar{g}_k \quad (j \rightarrow \infty)$$

Daher sind nach dem Satz von Beppo Levi \mathfrak{g}_k und \bar{g}_k integrierbar.

Ferner gelten die Ungleichungen

$$h \leq \mathfrak{g}_k \leq f_k \leq \bar{g}_k \leq h,$$

also

$$-\int^- h(x)dx \leq \int_{\mathbb{R}^n} \mathfrak{g}_k(x)dx \leq \int_{\mathbb{R}^n} f_k(x)dx \leq \int_{\mathbb{R}^n} \bar{g}_k(x)dx \leq \int^- h(x)dx.$$

Da \mathfrak{g}_k und \bar{g}_k monoton gegen f konvergieren, ist f integrierbar und aus der letzten Ungleichung folgt

$$\begin{aligned} \lim_{k \rightarrow \infty} \int_{\mathbb{R}^n} \mathfrak{g}_k(x)dx &= \lim_{k \rightarrow \infty} \int_{\mathbb{R}^n} f_k(x)dx \\ &= \lim_{k \rightarrow \infty} \int_{\mathbb{R}^n} \bar{g}_k(x)dx = \int_{\mathbb{R}^n} f(x)dx. \end{aligned}$$

□

Bezeichnung: Sei f eine Funktion, deren Definitionsbereich die Menge $D \subseteq \mathbb{R}^n$ umfasse. Wir definieren

$$\chi_D \cdot f : \mathbb{R}^n \longrightarrow \mathbb{R}$$

durch

$$\chi_D \cdot f(x) = \begin{cases} f(x) & \text{für } x \in D, \\ 0 & \text{für } x \notin D. \end{cases}$$

Die Funktion f heißt über D integrierbar, wenn $\chi_D \cdot f$ integrierbar ist, und man definiert

$$\int_D f(X)dx := \int_{\mathbb{R}^n} \chi_D \cdot f(x)dx.$$

A.7 Integrierbarkeitskriterien

Eine Teilmenge $A \subseteq \mathbb{R}^n$ heißt **endlich messbar**, wenn die charakteristische Funktion

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

integrierbar ist und man nennt

$$v(A) := \int_{\mathbb{R}^n} \chi_A(x)dx$$

das (Euklidische) Volumen von A .

Satz A.33. *Die charakteristische Funktion*

$$\chi_U : \mathbb{R}^n \rightarrow \mathbb{R}$$

einer offenen Teilmenge $U \subseteq \mathbb{R}^n$ gehört der Klasse B^+ an.

Beweis: 1. *Schritt:* U ist ein offener Quader

$$U = \{x \in \mathbb{R}^n : a_\nu < x_\nu < b_\nu\} \quad (a_\nu \leq b_\nu).$$

Im Falle $n = 1$ approximiert man die charakteristische Funktion in naheliegender Weise von unten durch Trapeze. Im Fall $n > 1$ verfährt man ähnlich.

2. *Schritt.* U ist beliebig. Zunächst zeigen wir, dass U abzählbare Vereinigung von offenen Quadern ist:

$$U = U_1 \cup U_2 \cup U_3 \cup \dots, \quad U_\nu \text{ offene Quader.}$$

Man betrachte hierzu die Menge aller offenen Quader

$$Q \subseteq U; \quad Q = \{x : a_\nu < x_\nu < b_\nu, 1 \leq \nu \leq n\},$$

wobei die Zahlen $a_1, \dots, a_n, b_1, \dots, b_n$ rational sind. Es ist klar, dass die Menge dieser Quader U überdeckt. Außerdem ist diese Menge abzählbar, weil die rationalen Zahlen und damit auch die n -Tupel von rationalen Zahlen abzählbar sind.

Man hat jetzt eine monotone Approximation von χ_U :

$$\chi_{U_1}, \quad \chi_{U_1 \cup U_2}, \quad \chi_{U_1 \cup U_2 \cup U_3}, \dots$$

□

Satz A.34. *Jede stetige und beschränkte Funktion $f : D \rightarrow \mathbb{R}$ auf einem offenen und beschränkten Teil $D \subseteq \mathbb{R}^n$ ist integrierbar.*

Beweis: Man kann annehmen, dass f nirgends negativ ist. Die Funktion

$$f \cdot \chi_D(x) = \begin{cases} f(x) & \text{für } x \in D, \\ 0 & \text{für } x \notin D, \end{cases}$$

gehört dann sogar der Klasse $B^+(\mathbb{R}^n)$ an. Sie wird approximiert durch die Folge

$$f_k \cdot f \cdot \chi_D,$$

wobei $f_k \in C_c$ eine Folge ist, die χ_D monoton approximiert. □

Folgerung A.35.

- a) *Jede beschränkte offene Teilmenge des \mathbb{R}^n ist endlich messbar.*
- b) *Jede kompakte Teilmenge des \mathbb{R}^n ist endlich messbar.*

Beweis: Es ist nur noch b) zu beweisen.

Sei jetzt $K \subseteq \mathbb{R}^n$ kompakt. Wir werden eine Folge von offenen beschränkten Mengen

$$U_1 \supseteq U_2 \supseteq U_3 \supseteq \dots \supseteq K$$

konstruieren, so dass

$$\bigcap_{\nu=1}^{\infty} U_\nu = K$$

gilt. Dann können wir wieder den Grenzwertsatz (auf die Folge χ_{U_ν}) anwenden. Wir setzen

$$U_\nu = \left\{ x \in \mathbb{R}^n : \|x - y\| < \frac{1}{\nu} \text{ für mindestens ein } y \in K \right\}.$$

Es sei wieder dem Leser überlassen, die gewünschten Eigenschaften zu beweisen. \square

Das Lebesgue'sche Integral ist eine Verallgemeinerung des Reginintegrals

Sei

$$f : [a, b] \rightarrow \mathbb{R}$$

eine Regelfunktion. Dann ist f auch Lebesgue-integrierbar und das Regin- und Lebesgue-Integral stimmen überein. Dies ist für Treppenfunktionen einfach zu zeigen und folgt dann allgemein aus den Grenzwertsätzen.

Allgemeiner gilt:

Sei

$$f : D \rightarrow \mathbb{R}, \quad D \subseteq \mathbb{R} \quad \text{ein Intervall,}$$

eine Funktion auf einem nicht notwendigerweise abgeschlossenen Intervall. Die Einschränkung von f auf jedes geschlossene Intervall sei eine Regelfunktion.

Die Funktion f ist genau dann Lebesgue-integrierbar, wenn $|f|$ uneigentlich integrierbar ist im Sinne der [2], Abschnitt 4.3.

Das uneigentliche Integral und das Lebesgue-Integral stimmen dann überein. Der Beweis ergibt sich leicht aus dem Lebesgueschen Grenzwertsatz.

Die wichtigsten Integrierbarkeitskriterien liegen in den Grenzwertsätzen. Es gibt aber auch eine einfache direkte Charakterisierung der Integrierbarkeit, welche unabhängig von ihren Anwendungen von eigenem theoretischem Interesse ist.

Satz A.36. *Eine Funktion*

$$f : \mathbb{R}^n \rightarrow \mathbb{R}$$

ist genau dann integrierbar, wenn es zu jedem $\varepsilon > 0$ eine Funktion $g \in C_c$ mit

$$\int |f(x) - g(x)| dx < \varepsilon$$

gibt.

(Dieser Satz wird noch in A.8 kommentiert.)

Beweis: Es ist unmittelbar klar, dass die angegebene Bedingung notwendig für die Integrierbarkeit von f ist, denn man findet zunächst eine Bairesche Funktion, deren Integral sich beliebig wenig von dem von f unterscheidet und anschließend eine stetige Funktion mit kompaktem Träger, deren Integral beliebig nahe bei dem der Baireschen Funktion liegt. Wir müssen umgekehrt zeigen, dass diese Bedingung hinreichend ist.

Nach Voraussetzung findet man zu jedem $\varepsilon > 0$ eine stetige Funktion g mit kompaktem Träger und eine Bairesche Funktion h mit den Eigenschaften

$$|f(x) - g(x)| \leq h(x), \quad \int^- h(x) dx < \varepsilon.$$

Die Behauptung folgt nun aus der Charakterisierung A.26 (mit Hilfe der Baireschen Funktionen $h + g$ und $h - g$). \square

A.8 Nullmengen

Definition A.37.

(a) Eine Funktion $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ heißt Nullfunktion, wenn

$$\int^- |f(x)| dx = 0$$

gilt.

(b) Eine Teilmenge $A \subseteq \mathbb{R}^n$ heißt Nullmenge, wenn die charakteristische Funktion χ_A

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

eine Nullfunktion ist.

Die Ungleichung $\int_- \leq \int^-$ zeigt, dass jede Nullfunktion in der Tat integrierbar ist. Außerdem ist mit f auch h eine Nullfunktion, wenn $|h| \leq |f|$ gilt.

Satz A.38. Eine Funktion $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ ist dann und nur dann Nullfunktion, wenn die Menge $\{x \in \mathbb{R}^n : f(x) \neq 0\}$ eine Nullmenge ist.

Beweis: Sei f eine Nullfunktion. Man wende den Satz von Beppo Levi auf die Folge

$$h_\nu(x) = \begin{cases} \nu|f(x)|, & \text{falls } f(x) \neq \pm\infty \\ \nu, & \text{falls } f(x) = \infty \end{cases}$$

an und erhält

$$\int_{\mathbb{R}^n} h(x) dx = 0,$$

wobei

$$h(x) = \begin{cases} \infty & \text{für } f(x) \neq 0, \\ 0 & \text{für } f(x) = 0. \end{cases}$$

Aus der Ungleichung

$$\chi_A \leq h \text{ mit } A = \{x \in \mathbb{R}^n : f(x) \neq 0\}$$

folgt dann, dass A eine Nullmenge ist.

Umkehrung. Sei $A = \{x \in \mathbb{R}^n : f(x) \neq 0\}$ eine Nullmenge. Der erste Teil dieses Beweises zeigt dann, dass h eine Nullfunktion ist und wegen $|f| \leq h$ ist dann auch f eine Nullfunktion. \square

Bemerkung A.39. Wenn eine Funktion f integrierbar ist, dann ist die Menge der Punkte

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

eine Nullmenge.

Beweis: Es sei A die Menge der Unstetigkeitsstellen von f und χ_A die charakteristische Funktion, offenbar gilt

$$\chi_A(x) = f(x) + (-f(x)),$$

wann immer diese Summe definiert ist. Nicht definiert sind die Ausdrücke $\infty + (-\infty)$.

Nach A.29 gilt daher

$$\int_{\mathbb{R}^n} \chi_A(x) dx \leq \int_{\mathbb{R}^n} f(x) dx - \int_{\mathbb{R}^n} f(x) dx = 0.$$

\square

Satz A.40. Sei f eine integrierbare Funktion und g eine Funktion, welche nur auf einer Nullmenge von f verschieden ist. Dann ist auch g integrierbar und die Integrale von f und g stimmen überein.

Folgerung A.41. Man darf eine integrierbare Funktion in ihren Unendlichkeitsstellen beliebig abändern - etwa zu 0, ohne ihre Integrierbarkeit zu verlieren und den Wert des Integrals zu ändern.

Damit haben die Unendlichkeitsstellen ihre Bedeutung verloren!

Beweis: Man kann annehmen, dass die Funktion g in allen Punkten, in denen sie sich von f unterscheidet, verschwindet. Dann gilt aber $f = g + h$ mit einer Nullfunktion und man kann A.29 verwenden (A.28 genügt nicht, da wir noch Unendlichkeitsstellen zulassen). \square

Was haben wir nun gewonnen? Es wurde gezeigt, dass man integrierbare Funktionen in Nullmengen beliebig abändern kann ohne ihre Integrierbarkeit und den Wert des Integrals zu verändern.

Außerdem wissen wir, dass es auf die Unendlichkeitsstellen einer integrierbaren Funktion beim Integrieren nicht ankommt, denn diese bilden eine Nullmenge.

Satz A.42.

- (a) Jede Teilmenge einer Nullmenge ist eine Nullmenge.
- (b) Jeder Punkt ist eine Nullmenge.
- (c) Sind A_1, A_2, A_3, \dots Nullmengen, so ist auch $\bigcup_{\nu=1}^{\infty} A_{\nu}$ eine Nullmenge.

Beweis: a) und b) sind klar. c) folgert man leicht aus dem Satz von Beppo Levi. \square

Insbesondere ist \mathbb{Q} in \mathbb{R} eine Nullmenge, womit noch einmal sehr drastisch gezeigt ist, dass es irrationale Zahlen geben muss.

Weiteres Beispiel einer Nullmenge: Sei U ein Untervektorraum des \mathbb{R}^n mit $\dim U < n$. Dann ist

$$L = x + U \text{ für jedes } x \in \mathbb{R}^n$$

eine Nullmenge.

Übungsaufgabe. Man beweise dies wenigstens für einen achsenparallelen Unterraum, d.h.

$$U = \{(x_1, \dots, x_n) : x_k = \dots = x_n = 0\} \quad (1 \leq k \leq n).$$

(Den allgemeinen Fall kann man auf diesen speziellen mittels der Transformationsformel (A.10) zurückführen.)

A.9 Der Satz von Fubini

In diesem Kapitel soll Satz A.8 über die Vertauschbarkeit der Integrationsreihenfolge auf integrierbare Funktionen verallgemeinert werden. Gegeben sei eine integrierbare Funktion

$$(x, y) \mapsto f(x, y) = f(x_1, \dots, x_n, y_1, \dots, y_m)$$

in $n + m$ Veränderlichen. Wir setzen $z = (x, y)$ und bezeichnen die Volumenelemente im

$$\mathbb{R}^{n+m} \text{ mit } dz, \text{ im } \mathbb{R}^n \text{ mit } dx \text{ und im } \mathbb{R}^m \text{ mit } dy.$$

Es soll eine Formel der Art

$$\int_{\mathbb{R}^{n+m}} f(x, y) dz = \int_{\mathbb{R}^m} \left[\int_{\mathbb{R}^n} f(x, y) dx \right] dy = \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^m} f(x, y) dy \right] dx$$

bewiesen werden.

Das Problem ist zunächst, dass wir nicht wissen, ob f bei festgehaltenem y bzw. x integrierbare Funktion von x bzw. y ist.

Daher behelfen wir uns zunächst mit äußerem und innerem Integral, also beispielsweise mit

$$x \mapsto \int_{\mathbb{R}^m} f(x, y) dy \quad (\text{dies ist eine Funktion von } x).$$

Von dieser Funktion kann man sagen, dass sie integrierbar ist.

Satz A.43 (Fubini). Gegeben sei eine integrierbare Funktion

$$f : \mathbb{R}^{n+m} \rightarrow \mathbb{R}, \quad (x, y) \mapsto f(x, y).$$

Dann sind die beiden Funktionen

$$x \mapsto \int_{\mathbb{R}^m}^- f(x, y) dy, \quad y \mapsto \int_{\mathbb{R}^n}^- f(x, y) dx$$

ebenfalls integrierbar (im \mathbb{R}^n bzw. \mathbb{R}^m) und es gilt die Formel

$$\begin{aligned} \int_{\mathbb{R}^{n+m}} f(x, y) dz &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^m}^- f(x, y) dy dx \\ &= \int_{\mathbb{R}^m} \int_{\mathbb{R}^n}^- f(x, y) dx dy. \end{aligned}$$

Dieselbe Formel bleibt richtig, wenn man das äußere Integral durch das innere Integral ersetzt.

Es drängt sich natürlich die Frage auf, ob die Funktion $f(x, y)$ bei festem y (bzw. x) nicht sogar integrierbar ist. Aus A.43 folgt, dass dies bis auf eine Ausnahmemenge vom Maß 0, auf die es bei der Integration nicht ankommt, der Fall ist. Man muss nur beachten, dass aus A.43 beispielsweise

$$\int_{\mathbb{R}^n}^- \left[\int_{\mathbb{R}^m} f(x, y) dy - \int_{\mathbb{R}^m}^- f(x, y) dy \right] dx$$

folgt. Der Integrand ist also eine Nullfunktion. Man kann also die Fubini-Formel auch etwas unpräzise aber doch legitim in der Form

$$\int_{\mathbb{R}^{n+m}} f(x, y) dz = \int_{\mathbb{R}^m} \left[\int_{\mathbb{R}^n} f(x, y) dx \right] dy = \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^m} f(x, y) dy \right] dx$$

schreiben. Für die Anwendungen ist dieser Zusatz jedoch irrelevant. Man sieht die Integrierbarkeit meist direkt.

Beweis:

1. Schritt: $f \in C_c$: Die Behauptung folgt aus A.8.
2. Schritt: $f \in B^+(\mathbb{R}^{n+m})$. Es existiert eine Folge

$$f_k \uparrow f, \quad f_k \in C_c.$$

Bei festgehaltenem x gilt dann

$$\int_{\mathbb{R}^m} f(x, y) dy \uparrow \int_{\mathbb{R}^m} f(x, y) dy.$$

Nach A.6 ist die Folge dieser Integrale stetig, d.h. das Integral

$$\int_{\mathbb{R}^m} f(x, y) dy$$

liegt in B^+ (als Funktion von x) und außerdem gilt

$$\int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^m} f(x, y) dy \right] = \lim_{k \rightarrow \infty} \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^m} f_k(x, y) dy \right] dx.$$

Da die Behauptung für die stetige Funktion f mit kompaktem Träger f_k richtig ist, gilt sie auch für f .

3. *Schritt.* Sei $h : \mathbb{R}^{n+m} \rightarrow \mathbb{R}$ eine Funktion, $h \in B^+$ und $f \leq h$. Dann gilt

$$\begin{aligned} \int_{\mathbb{R}^{n+m}} h(x, y) dz &\stackrel{2.Schritt}{=} \int_{\mathbb{R}^m} \int_{\mathbb{R}^n} h(x, y) dx dy \\ &= \int_{\mathbb{R}^m} \left[\int_{\mathbb{R}^n} h(x, y) dx \right] dy \geq \int_{\mathbb{R}^m} \left[\int_{\mathbb{R}^n} f(x, y) dx \right] dy. \end{aligned}$$

Hieraus folgt mit Hilfe der Definition des Integrals

$$\begin{aligned} \int_{\mathbb{R}^{n+m}} f(x, y) dz &= \int_{h \in B^+, f \leq h} \left\{ \int_{\mathbb{R}^{n+m}} h(x, y) dz \right\} \\ &\geq \int_{\mathbb{R}^m} \left[\int_{\mathbb{R}^n} f(x, y) dx \right] dy. \end{aligned}$$

Diese Ungleichung kann man auch für $-f$ anstelle von f anwenden. Beachtet man

$$\int_{\mathbb{R}^n}^- (-f) dx = - \int_{\mathbb{R}^n}^- f dx,$$

so folgt

$$\int_{\mathbb{R}^m} \left[\int_{\mathbb{R}^n}^- f(x, y) dx \right] dy \geq \int_{\mathbb{R}^{n+m}} f(x, y) dz.$$

Aus der Ungleichung

$$\int_{\mathbb{R}^n}^- \leq \int_{\mathbb{R}^n}^-$$

folgt dann

$$\int^- \left[\int^- f(x, y) dx \right] dy \geq \int_- \left[\int^- f(x, y) dx \right] dy \geq \int_- \left[\int_- f(x, y) dx \right] dy.$$

Andererseits haben wir eben die Ungleichungen

$$\int_- \int_- \geq \int^- \int^-$$

erhalten, also folgt

$$\begin{aligned} \int_{\mathbb{R}^{n+m}} f(x, y) dz &= \int^- \left[\int^- f(x, y) dx \right] dy \\ &= \int_- \left[\int^- f(x, y) dx \right] dy. \end{aligned}$$

Da äußeres und inneres Integral von $\int^- f(x, y) dx$ somit übereinstimmen, ist diese Funktion integrierbar, und es gilt die Formel

$$\int_{\mathbb{R}^{n+m}} f(x, y) dz = \int \left[\int^- f(x, y) dx \right] dy.$$

Entsprechend werden auch die anderen Formeln bewiesen. □

Eine Anwendung:

Satz A.44. Sei

$$f : \mathbb{R}^n \rightarrow \mathbb{R}, \quad f \geq 0,$$

eine integrierbare Funktion, die keine negativen Werte annimmt.

Sei

$$M = \{(x, y) \in \mathbb{R}^n \times \mathbb{R} : 0 \leq y \leq f(x)\}.$$

Dann ist M endlich messbar und es gilt

$$\text{vol}(M) = \int_{\mathbb{R}^n} f(x) dx.$$

Beweis: Das Volumen ist definiert durch die Formel

$$\text{vol}(M) = \int_{\mathbb{R}^{n+1}} \chi_M(x, y) dz.$$

Man integriere zunächst bei festem x über y und wende den Satz von Fubini an. □

Beispiel: Hat man im \mathbb{R}^2 einen durch Ungleichungen beschriebenen Bereich

$$A = \{(y, z) : a(y) \leq z \leq b(y), c \leq y \leq d\}$$

mit integrierbaren Funktionen a und b , so gilt für eine über A integrierbare Funktion f

$$\int \int_A f(y, z) dy dz = \int_c^d \int_{a(y)}^{b(y)} f(y, z) dz dy.$$

Setzen wir speziell $f(x) = \chi_A(x)$, so bekommen wir für das Volumen von A die Formel

$$v(A) = \int_c^d (b(y) - a(y)) dy.$$

A.10 Die Transformationsformel

Wir erinnern uns daran, dass eine Funktion

$$f : D \rightarrow \mathbb{R}, \quad D \subseteq \mathbb{R}^n,$$

über D integrierbar heißt, wenn die Funktion

$$\tilde{f} : \mathbb{R}^n \rightarrow \mathbb{R}, \quad \tilde{f}(x) = \begin{cases} f(x) & \text{für } x \in D, \\ 0 & \text{für } x \notin D, \end{cases}$$

integrierbar ist, und wir setzen dann

$$\int_D f(x) dx = \int_{\mathbb{R}^n} \tilde{f}(x) dx.$$

Seien $A, B \subseteq \mathbb{R}^n$ offene Mengen. Unter einem **Diffeomorphismus**

$$\varphi : A \longrightarrow B$$

verstehen wir eine bijektive (=umkehrbare) stetig differenzierbare Abbildung mit nirgends verschwindender Funktionaldeterminante

$$j(\varphi, x) = \det J(\varphi; x) \neq 0 \quad \text{für alle } x \in A.$$

Nach dem Satz für implizite Funktionen ist dann auch φ^{-1} differenzierbar und es gilt

$$J(\varphi; x)^{-1} = J(\varphi^{-1}, \varphi(x)).$$

Satz A.45. Es sei $u : A \rightarrow B$ ein Diffeomorphismus zwischen offenen Mengen $A, B \subseteq \mathbb{R}^n$. Ist $f : B \rightarrow \mathbb{R}$ eine integrierbare Funktion, so gilt

$$\int_B f(y)dy = \int_A f(u(x))|j(u, x)|dx.$$

(Insbesondere wird also behauptet, dass $f(u(x))|j(u, x)|$ über A integrierbar ist.)

Beweis der Transformationsformel:

1. Schritt. Es sei $f : \mathbb{R}^n \rightarrow \mathbb{R}$ eine stetige Funktion mit kompaktem Träger. Dann gelten die Transformationsformeln

a)
$$\int_{\mathbb{R}^n} f(x_1, \dots, x_{i+1}, x_i + x_j, x_{i+1}, \dots, x_n)dx = \int_{\mathbb{R}^n} f(x)dx$$

b)
$$\int_{\mathbb{R}^n} f(a_1x_1, \dots, a_nx_n)dx = |a_1 \dots a_n| \int_{\mathbb{R}^n} f(x)dx$$

c)
$$\int_{\mathbb{R}^n} f(x_1, \dots, x_n)dx_{\sigma(1)} \dots dx_{\sigma(n)} = \int_{\mathbb{R}^n} f(x)dx_1 \dots dx_n$$

d)
$$\int_{\mathbb{R}^n} f(x + b)dx = \int_{\mathbb{R}^n} f(x)dx$$

Dabei sei $b = (b_1, \dots, b_n)$ ein festes n -Tupel.

Beweis: Man benutzt die Transformationsformel im Fall $n = 1$, sowie die Tatsache, dass das Integral für stetige Funktionen mit kompaktem Träger iterativ definiert ist, beispielsweise

$$\int_{\mathbb{R}^2} f(x_1 + x_2)dx = \int_{\mathbb{R}} \left[\int_{\mathbb{R}} f(x_1 + x_2, x_2)dx_1 \right] dx_2.$$

Im inneren Integral macht man die Substitution $t = x_1 + x_2$ und erhält für das innere Integral

$$\int_{\mathbb{R}} f(x_1, x_2)dx_1.$$

2. *Schritt.* Sei $f : \mathbb{R}^n \rightarrow \mathbb{R}$ stetig mit kompaktem Träger und

$$A = (a_{\nu\mu}) \quad 1 \leq \nu, \mu \leq n$$

eine feste Matrix mit von Null verschiedener Determinante, $b \in \mathbb{R}^n$ ein fester Vektor. Dann ist

$$\int_{\mathbb{R}^n} f(y) dy = \int_{\mathbb{R}^n} f(Ax + b) |\det A| dx.$$

Beweis: Man kann $b = 0$ annehmen (1. Schritt d)). Im Spezialfall, dass die Abbildung A zu den drei Typen

- a) Scherung
- b) $(x_1, \dots, x_n) \rightarrow (a_1 x_1, \dots, a_n x_n)$
- c) Permutation der Variablen

gehört, haben wir das im 1. Schritt erkannt.

Man muss jetzt nur aus der linearen Algebra wissen, dass sich jede lineare Abbildung A , $\det A \neq 0$, als Hintereinanderausführung von Transformationen des Typs a) - c) schreiben lässt. Dies ist nichts anderes als die Tatsache, dass sich jede Matrix A , $\det A \neq 0$, durch elementare Umformungen in die Einheitsmatrix überführen lässt.

Außerdem muss man benutzen, dass die Determinante multiplikativ ist, $\det(A_1 \dots, A_n) = \det A_1 \cdots \det A_n$.

3. *Schritt.* Sei $f : \mathbb{R}^n \rightarrow \mathbb{R}$ beliebig und $A = (a_{\nu,\mu})_{1 \leq \nu, \mu \leq n}$ eine Matrix mit $\det A \neq 0$, $b \in \mathbb{R}^n$. Dann ist

$$\int^- f(y) dy = \int^- f(Ax + b) |\det A| dx.$$

Beweis. Ist $f \in B^+$ und $f_k \uparrow f$ eine monotone Approximation durch Funktionen $f_k \in C_c$, so sind die Funktionen

$$g_k(x) = f_k(Ax + b)$$

ebenfalls stetig mit kompaktem Träger und approximieren monoton $g(x) = f(Ax + b)$. Damit ist die Behauptung für Funktionen $f \in B^+$ zurückgeführt auf den 2. Schritt. Ist f beliebig, so folgt die Behauptung unmittelbar aus der Definition des äußeren Integrals.

Die bisherigen Überlegungen zeigen:

1) Ist f integrierbar, so gilt

$$\int_{\mathbb{R}^n} f(y)dy = \int_{\mathbb{R}^n} f(Ax + b)|\det A|dx.$$

2) Jede Teilmenge $W_0 \subseteq \mathbb{R}^n$ der Form

$$W_0 = a + W, \quad W \subseteq \mathbb{R}^n \text{ ein Untervektorraum, } \dim W < n,$$

ist eine Nullmenge.

(Mit Hilfe einer linearen Transformation macht man W achsenparallel.)

4. *Schritt:* (A, B, u wie in A.45.) Die Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ sei stetig. Sei $a \in A$ fest, $b = u(a)$. Es gelte $f(b) > 0$. Außerdem sei $q > 1$ beliebig, aber fest gewählt. Dann existiert eine Umgebung $a \in U \subseteq A$,

$$U = U_{\mathbb{R}}(a) = \{x : \|x - a\| < R\} \quad (\text{hierbei sei } \|\cdot\| \text{ die Maximumsnorm}),$$

so dass für jeden Würfel

$$W = U_r(x_0) \subseteq U$$

gilt

$$q^{n+4} \int_{u(W)} f(y)dy \geq \int_W f(ux)|j(u, x)|dx.$$

(Es wird nicht gefordert, dass W den gleichen Mittelpunkt wie U hat!)

Beweis. Die Ungleichung wird auf den linearen Fall (3. Schritt) zurückgeführt, indem man $u(x)$ linear approximiert.

$$u(x) = u(A) + J(u, a)(x - a) + r(x).$$

Wir ersetzen u durch

$$u_0(x) = u(a) + J(u; a)(x - a)$$

und erhalten nach dem 3. Schritt

$$\int_{u_0(W)} f(y)dy = \int_W f(u_0x)|j(u, a)|dx.$$

Dabei sei r so klein gewählt, dass der abgeschlossene Würfel W noch in A enthalten ist, dann ist $f(u_0x)$ in W beschränkt und damit wegen A.34 über W integrierbar.

Wir denken uns R immer so klein gewählt, dass

$$f(u_0x) > 0 \quad \text{für } x \in U$$

gilt. (Beachte $u_0(a) = u(a) = b$ und $f(b) > 0$ nach Voraussetzung.)

In einer solchen Umgebung kann man

$$\frac{f(ux)}{f(u_0x)}$$

betrachten. Diese Funktion ist stetig und konvergiert gegen 1, wenn x nach a strebt. Daher gilt

$$q > \frac{f(ux)}{f(u_0x)} \quad \text{für } x \in U, \quad r \text{ genügend klein.}$$

Aus demselben Grund gilt

$$q|j(u, a)| \geq |j(u, x)| \quad \text{für } x \in U,$$

wenn man R genügend klein wählt.

Also gilt

$$q^2 \int_{u_0(W)} f(y) dy \geq \int_W f(ux) |j(u, x)| dx.$$

In dieser Ungleichung müsste $u(W)$ anstelle von $u_0(W)$ stehen. Wir werden daher

$$u(W) \quad \text{und} \quad u_0(W)$$

vergleichen.

Behauptung. Wählt man R genügend klein, so gilt $u_0(W^*) \subseteq u(W)$. Dabei sei

$$W^* = \{x : \|x - x^0\| < rq^{-1}\}$$

der Würfel, der aus W durch Schrumpfung um den Faktor q^{-1} entsteht.

Beweis. Dies bedeutet nichts anderes als

$$v(W^*) \subseteq W \quad \text{mit } v = u^{-1}u_0,$$

also

$$\|x - x^0\| < rq^{-1} \Rightarrow \|v(x) - v(x^0)\| < r.$$

Wir schätzen $v(x) - v(x^0)$ nach dem verallgemeinerten Mittelwertsatz der Differentialrechnung ab:

$$\|v(x) - v(x^0)\| \leq \|x - x^0\| \sum_{\mu=1}^n |\partial_{\mu} v_{\nu}(\xi^{(\nu)})|.$$

Dabei ist $\xi^{(\nu)}$ ein Punkt auf der Verbindungsstrecke zwischen x und x^0 . Hieraus folgt

$$\|v(x) - v(x^0)\| \leq \|x - x^0\| < M$$

mit

$$M = \sup_{\xi \in U} \sum_{\mu=1}^n |\partial_{\mu} v_{\nu}(\xi)|.$$

Wir wollen ja U so bestimmen, dass gilt

$$\|v(x) - v(x^0)\| < r \quad \text{falls } \|x - x^0\| < rq^{-1}.$$

Dazu benötigt man ersichtlich die Ungleichung $M \leq q$, d.h.

$$\sum_{\mu=1}^n |\partial_{\mu} v_{\nu}(\xi)| < q \quad \text{für alle } \xi \in U.$$

Nun beachte man, dass nach der Kettenregel für $v = u^{-1} \cdot u_0$ gilt:

$$j(v, x) = \text{Einheitsmatrix},$$

d.h. obige Ungleichung ist im Punkt $\xi = a$ erfüllt (wegen $1 < q$). Sie gilt dann aus Stetigkeitsgründen auch in einer vollen Umgebung von a .

Damit erhalten wir nun die Ungleichung

$$q^2 \int_{u(W)} f(y) dy \geq \int_{W^*} f(u(x)) |j(u, x)| dx.$$

Die Abbildung u_0 ist damit eliminiert, wir müssen allerdings noch die Integrale

$$\int_{W^*} g(x) dx \quad \text{und} \quad \int_W g(x) dx \quad \text{mit} \quad g(x) = f(u(x)) |j(u, x)|$$

vergleichen.

Wählt man R genügend klein, so gilt

$$g(a) \cdot q > g(x) > g(a)q^{-1} \quad \text{für } x \in U \quad (\text{beachte } q > 1 \text{ und } g(a) > 0).$$

Hieraus folgt

$$\int_{W^*} g(x) dx \geq g(a)q^{-1} = \text{vol}(W_{rq^{-1}}) = g(a)q^{-1} \cdot q^{-n}(2r)^n;$$

andererseits ist

$$\int_W g(x) dx \leq g(a) \cdot q \cdot \text{vol}(W) = g(a)q(2r)^n.$$

Aus den beiden Ungleichungen folgt

$$q^{n+2} \int_{W^*} g(x) dx \geq \int_W g(x) dx.$$

Wir erhalten die gewünschte Ungleichung

$$q^{n+4} \int_{u(W)} f(y) dy \geq \int_W f(ux) |j(u, x)| dx$$

für $W \subseteq U = U_R(A)$, R genügend klein.

5. Schritt, Konstruktion einer geeigneten Würfelüberdeckung.

Jede offene Teilmenge $U \subseteq \mathbb{R}^n$ lässt sich als abzählbare Vereinigung von abgeschlossenen Würfeln schreiben

$$U = \overline{W}_1 \cup \overline{W}_2 \cup \overline{W}_3 \cup \dots,$$

wobei die Würfel W_ν offen und paarweise disjunkt sind

$$\overline{W}_\nu = \{x : \|x - a\| \leq \varepsilon, \quad W_\nu = \{x : \|x - a\| < \varepsilon\}.$$

Beweis. Wir betrachten die Menge aller Würfel

$$W = \left\{ x : \frac{a_\nu u}{2^r} < x_\nu < \frac{a_\nu u + 1}{2^r}, \quad \nu = 1, \dots, n \right\},$$

wobei r alle natürlichen und a_ν alle ganzen Zahlen durchläuft.

Die Menge dieser Würfel ist abzählbar. (Die Menge $\mathbb{N} \times \mathbb{Z}^n$ ist abzählbar, weil allgemein das kartesische Produkt von abzählbaren Mengen abzählbar ist.) Man überlegt sich nun (dies im einzelnen durchzuführen sei dem Leser überlassen):

a) Sind W und W' zwei der beschriebenen Würfelmengen, so gilt

$$W \cap W' \neq \emptyset \Rightarrow W \subseteq W' \text{ oder } W' \subseteq W.$$

b) Jeder Punkt $a \in \mathbb{R}^n$ ist enthalten in einem der Würfel W , wobei noch r beliebig groß gewählt werden kann (und daher die Kantenlänge beliebig klein).

Aus a) und b) konstruiert man nun leicht eine Würfelaufteilung der gewünschten Art.

6. Schritt. Die Funktion

$$f : \mathbb{R}^n \longrightarrow \mathbb{R}, \quad f \geq 0,$$

sei stetig mit kompaktem Träger in B . Es gilt

$$\int_B f(y) dy \geq \int_A f(u(x)) |j(u, x)| dx.$$

Beweis indirekt: Die Ungleichung sei falsch. Dann kann jedenfalls f nicht identisch Null sein, aus Stetigkeitsgründen sind beide Integrale nicht Null und man kann eine Zahl $q > 1$ finden, so dass sogar

$$q^{n+4} \int_B f(y) dy < \int_A f(u(x)) |j(u, x)| dx$$

gilt.

Wir betrachten nun die Menge

$$A_0 = \{x \in \mathbb{R}^n : f(u(x)) \neq 0\} \subseteq A.$$

Diese Menge ist offen, da f stetig ist. Nach dem 5. Schritt existiert eine Überdeckung

$$A_0 = \overline{W}_1 \cup \overline{W}_2 \cdots,$$

wobei die Würfel W_ν offen und paarweise disjunkt sind.

Behauptung. Für mindestens einen dieser Würfel, nennen wir ihn $W^{(1)}$, gilt die Ungleichung

$$q^{n+4} \int_{u(W^{(1)})} f(y) dy < \int_{W^{(1)}} f(u(x)) |j(u, x)| dx.$$

Beweis. Würde für alle Würfel W die Ungleichung „ \geq “ gelten, so könnte man mit Hilfe des Lebesgueschen Grenzwertsatzes sogar

$$\begin{aligned} q^{n+4} \int_B f(y) dy &\geq \sum_{\nu=1}^{\infty} q^{n+4} \int_{u(W_\nu)} f(y) dy \\ &\geq \sum_{\nu=1}^{\infty} \int_{W_\nu} f(u(x)) |j(u, x)| dx \\ &= \int_A f(u(x)) |j(u, x)| dx \end{aligned}$$

schließen. (Man beachte, dass das Integral über den Rand eines Würfels Null ergibt.)

Die Existenz eines Würfels $W^{(1)}$ mit

$$q^{n+4} \int_{u(W^{(1)})} f(y) dy < \int_{W^{(1)}} f(u(x)) |j(u, x)| dx$$

ist damit gesichert. Den Würfel $W^{(1)}$ kann man durch Halbieren der Kantenlänge in 2^n Würfel aufteilen. Mit der gleichen Schlussweise folgt die Existenz eines Würfels $W^{(2)}$ mit der obigen Ungleichung.

So fortfahrend erhält man eine Folge von Würfeln

$$W^{(1)} \supseteq W^{(2)} \supseteq W^{(3)} \supseteq \dots$$

mit den Eigenschaften

a) Kantenlänge von $W^{(k)} \rightarrow 0$ für $k \rightarrow \infty$.

b) $q^{n+4} \int_{u(W^{(k)})} f(y) dy < \int_{W^{(k)}} f(u(x)) |j(u, x)| dx$.

Nach dem verallgemeinerten Intervallschachtelungsprinzip existiert ein Punkt

$$a \in \overline{W^{(1)}} \cap \overline{W^{(2)}} \cap \overline{W^{(3)}} \dots$$

Zu diesem a betrachten wir die im 4. Schritt konstruierte Umgebung U . Wählt man k genügend groß, so gilt

$$W^{(k)} \subseteq U$$

und wir haben einen Widerspruch zwischen den Ungleichungen des 4. und 6. Schritts erhalten.

7. Schritt, Beweis von Theorem A.45 für stetige Funktionen mit kompaktem Träger f .

Man kann annehmen, dass $f \geq 0$ gilt. Dies liegt an der Möglichkeit des Aufspaltens

$$f = f^+ - f^-$$

mit

$$f^+ = f \vee 0, \quad f^- = (-f)^+.$$

Nach dem 6. Schritt gilt die Ungleichung

$$\int_B f(y) dy \geq \int_A f(u(x)) |j(u, x)| dx.$$

Wendet man diese Ungleichung an auf

$$\begin{aligned}
& A \text{ anstelle von } B \\
& B \text{ anstelle von } A \\
& u^{-1} \text{ anstelle von } u \\
& f(u(x))|j(u, x)| \text{ anstelle von } f,
\end{aligned}$$

so resultiert die umgekehrte Gleichung.

8. Schritt, Beweis von A.45. Sei

$$f : \mathbb{R}^n \rightarrow \mathbb{R}, \quad f \geq 0$$

eine halbstetige Funktion. Dann ist auch $f \chi_B$ halbstetig. Man wähle

$$f_k \uparrow f$$

und

$$g_k \uparrow \chi_B; \quad g_k \geq 0,$$

wobei f_k und g_k stetig mit kompaktem Träger sei. Dann gilt

$$f_k \cdot g_k \uparrow f \cdot \chi_B.$$

Die Gleichung

$$\int_B f(y) dy = \int_A f(u(x)) |j(u, x)| dx$$

folgt nun aus dem 8. Schritt. Jetzt folgt die Transformationsformel unmittelbar für das äußere Integral und dann erst recht für das Integral. \square

A.11 Messbarkeit

Man kann jede Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ als Limes einer Folge von beschränkten Funktionen mit kompaktem Träger schreiben, nämlich

$$f = \lim f_k$$

mit

$$f_k(x) = \begin{cases} f_k(x), & \text{falls } \|x\| \leq k \text{ und } |f(x)| \leq k, \\ 0, & \text{sonst.} \end{cases}$$

Definition A.46. Eine Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ heißt messbar, falls die Funktionen f_k , $k = 1, 2, \dots$ alle integrierbar sind.

Es gilt offenbar

$$f_k = (\varphi_k \wedge f) \vee (-\varphi_k),$$

wenn φ_k die Funktion

$$\varphi_k(x) = \begin{cases} k & \text{für } \|x\| \leq k \\ 0 & \text{sonst} \end{cases}$$

bezeichnet. Da φ_k integrierbar ist, folgt

Bemerkung A.47. Jede integrierbare Funktion ist messbar.

Diese und die folgende Bemerkung folgen unmittelbar aus dem Lebesgueschen Grenzwertsatz.

Bemerkung A.48. Sei $f : \mathbb{R}^n \rightarrow \mathbb{R}$ eine messbare Funktion. Es existiere eine integrierbare Funktion h mit der Eigenschaft

$$|f(x)| \leq |h(x)| \quad \text{für alle } x.$$

Dann ist auch f integrierbar.

Die messbaren Funktionen haben alle wünschenswerten Stabilitätseigenschaften.

Satz A.49.

- 1) Summe und Produkt von messbaren Funktionen sind messbar, konstante Funktionen sind messbar.
- 2) Seien f, g messbare Funktionen. Dann sind auch $f \wedge g$, $f \vee g$, insbesondere $|f|$ messbar.
- 3) Sei f_n eine punktweise konvergente Folge messbarer Funktionen. Dann ist auch ihr Grenzwert messbar.
- 4) Man darf eine messbare Funktion auf einer Nullmenge abändern, ohne ihre Messbarkeit zu verlieren.

Beweis: Es genügt zu zeigen, dass das Produkt messbarer Funktionen messbar ist, da alle anderen Stabilitätseigenschaften aus entsprechenden Eigenschaften integrierbarer Funktionen folgen.

Wegen der Formel

$$4(f \cdot g) = (f + g)^2 - (f - g)^2$$

braucht man nur zu beweisen, dass mit f auch f^2 messbar ist. Da man f durch f_k ersetzen darf, genügt es zu zeigen:

Das Quadrat einer *beschränkten* integrierbaren Funktion ist integrierbar. Dies folgt aus der Charakterisierung A.36 unter Verwendung der Formel $f^2 - h^2 = (f + h)(f - h)$. \square

Als Anwendung des Begriffs der messbaren Funktion dient:

Ergänzung zum Satz von Fubini. Sei $f : \mathbb{R}^{n+m} \rightarrow \mathbb{R}$ eine messbare Funktion, so dass

$$\int \int |f(x, y)| dy dx$$

endlich ist. Dann ist f integrierbar und es gilt infolgedessen der Satz von Fubini A.43.

Der Beweis ist einfach und wird übergangen.

Als weitere Anwendung des Begriffs der messbaren Funktion führen wir die \mathcal{L}^p - und die L^p -Räume ein. Dabei kann X ein beliebiger lokal kompakter metrischer Raum sein, welcher abzählbar im Unendlichen ist und auf welchem ein Radonsches Maß ausgezeichnet ist. Für uns ist der Fall $X = \mathbb{R}^n$ mit dem Standardmaß ausreichend.

Den Fall $p = 1$ haben wir bereits behandelt (A.2). Auf Beweise gehen wir hier nicht ein. Man findet sie in der Standardliteratur über Maßtheorie. Hier geben wir nur einen Ausblick.

$\mathcal{L}^p(X)$, $p > 0$, bestehe aus allen messbaren Funktionen f , so dass $|f|^p$ integrierbar ist. Wir definieren

$$\|f\|_p = \sqrt[p]{\int_{\mathbb{R}^n} |f|^p dv} \quad \text{für } f \in \mathcal{L}^p(X).$$

Es gilt

- a) $\|cf\| = |c| \|f\|_p$,
- b) $\|f\|_p \geq 0$,
- c) Mit f und g ist auch $f + g$ in $\mathcal{L}^p(X)$ enthalten und es gilt

$$d) \|f + g\|_p \leq \|f\|_p + \|g\|_p.$$

Identifiziert man zwei Funktionen, wenn sie sich nur auf einer Nullmenge unterscheiden, so erhält man den Raum

$$L^p(X) = \{[f], f \in \mathcal{L}^p(X)\}, \quad [f] = \{g \in \mathcal{L}^p(X), \|f - g\|_p = 0\}.$$

Die Definitionen

$$[f] + [g] = [f + g], \quad C[f] = [Cf], \quad \|[f]\|_p = \|f\|_p$$

hängen nicht von der Wahl der Repräsentanten ab.

In $L^p(X)$ gilt

$$\|[f]\|_p = 0 \implies [f] = [0],$$

d.h. $L^p(X)$ ist ein normierter Vektorraum. Aus dem Lebesgueschen Grenzwertsatz lässt sich (in nicht offensichtlicher Weise) folgern:

Satz A.50. Die Räume $(L^p(X), \|\cdot\|_p)$ sind für $p > 0$ **Banachräume**. Der Unterraum der Funktionen $[f], f \in C_c(X)$, ist dicht in $L^p(X)$.

Übrigens: Wenn f und g stetige Funktionen mit kompaktem Träger sind, so gilt $[f] = [g] \implies f = g$. Wir können also (leicht unpräzise aber suggestiv) $C_c(X) \subseteq L^p(X)$ schreiben.

Der Fall $p = 2$ ist besonders wichtig. Man kann zeigen, dass durch

$$\langle [f], [g] \rangle := \int_X f(x) \overline{g(x)} dx$$

ein Skalarprodukt auf $L^2(X)$ definiert wird und erhält somit:

Satz A.51. Der Raum $L^2(X)$ ist ein **Hilbertraum** (wenn er mit obigem Skalarprodukt versehen wird).

(Dabei versteht man unter einem **Hilbertraum** einen vollständigen, mit einem Skalarprodukt versehenen \mathbb{R} - oder \mathbb{C} -Vektorraum. Die Forderung der Vollständigkeit bezieht sich stets auf die durch $\|x\| := \sqrt{\langle x, x \rangle}$ definierte Norm.)

B Integralrechnung von Funktionen mehrerer Veränderlicher

B.1 Jordanscher Inhalt und Lebesguesches Maß

Wir beginnen unsere Betrachtungen damit, dass wir versuchen, möglichst vielen Teilmengen $A \subseteq \mathbb{R}^n$ in anschaulicher Weise ein sogenanntes **Maß** zuzuordnen (das Lebesgue-Maß). Dies dient uns dann im nächsten Paragraphen als Grundlage für die Integration von Funktionen $f : A \rightarrow \mathbb{R}^m$, wobei $A \subseteq \mathbb{R}^n$ eine messbare Menge ist.

Wir definieren zunächst den elementareren Begriff des **Inhalts**. Sei stets $\mathfrak{M} \subseteq \mathcal{P}(\mathbb{R}^n)$, d.h. \mathfrak{M} sei eine Menge von Teilmengen des \mathbb{R}^n . Sei $\mathbb{R}_+ = \{x \in \mathbb{R} : x \geq 0\}$.

Definition B.1. Eine Mengenfunktion $m : \mathfrak{M} \rightarrow \mathbb{R}_+ \cup \{\infty\}$ mit den Eigenschaften

$$(M1) \quad A, B \in \mathfrak{M} \Rightarrow A \cup B, A \cap B, A \setminus B = \{x \in A : x \notin B\} \in \mathfrak{M}$$

$$(M2) \quad A \cap B = \emptyset \Rightarrow m(A \cup B) = m(A) + m(B)$$

heißt **Inhalt** (auf \mathfrak{M}).

Folgerungen:

- (1) Sind die Mengen A_1, \dots, A_k Elemente von \mathfrak{M} mit $A_i \cap A_j = \emptyset$ für $i \neq j$, so gilt

$$m \left(\bigcup_{i=1}^k A_i \right) = \sum_{i=1}^k m(A_i).$$

- (2) Für beliebige $A, B \in \mathfrak{M}$ gilt $m(A \cup B) + m(A \cap B) = m(A) + m(B)$.

- (3) $m(A \cup B) \leq m(A) + m(B)$.

- (4) $m(A) \leq m(B)$, falls $A \subseteq B$.

Definition B.2. Ein Inhalt m auf \mathfrak{M} heißt **Maß**, wenn anstelle von (M2) folgende schärfere Bedingung gilt:

(M3) Ist A_1, A_2, \dots eine Folge paarweise disjunkter Mengen aus \mathfrak{M} , so ist $\cup_{i=1}^{\infty} A_i \in \mathfrak{M}$ und es gilt $m(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} m(A_i)$.

Bemerkung. (M3) ist äquivalent zu

(M3') Ist B_1, B_2, \dots eine aufsteigende Folge von Mengen aus \mathfrak{M} , d.h. $B_i \subseteq B_{i+1}$ für $i \in \mathbb{N}$, so ist $\cup_{i=1}^{\infty} B_i \in \mathfrak{M}$ und es gilt:

$$m(\cup_{i=1}^{\infty} B_i) = \lim_{n \rightarrow \infty} m(B_n).$$

Folgerung B.3 (aus (M3)). Sei A_1, A_2, \dots eine Folge beliebiger Mengen aus \mathfrak{M} . Dann gilt

$$\cup_{i=1}^{\infty} A_i \in \mathfrak{M} \quad \text{und} \quad m(\cup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} m(A_i).$$

Folgerung B.4 (duale Aussage zu (M3')). Sei A_1, A_2, \dots eine absteigende Folge (d.h. $A_1 \supseteq A_2 \supseteq A_3 \supseteq \dots$) von Mengen aus \mathfrak{M} , mit $m(A_k) < \infty$ für alle k . Dann gehört $\cap_{k=1}^{\infty} A_k$ zu \mathfrak{M} , und es gilt

$$m\left(\bigcap_{k=1}^{\infty} A_k\right) = \lim_{k \rightarrow \infty} m(A_k).$$

Bemerkung. Die Bedingung $m(A_k) < \infty$ ist wesentlich, wie etwa das Beispiel $A_k = \{x \in \mathbb{R} : x \geq k\} \subseteq \mathbb{R}^1$ zeigt.

Wir werden nun an das Lebesgue-Maß eine zusätzliche Forderung (M4) stellen, und es wird sich zeigen, dass diese bei geeigneter Menge \mathfrak{M} ein Maß eindeutig definiert.

Gegeben seien reelle Zahlen $a_i < b_i$ für $i = 1, \dots, n$. Unter einem **offenen Quader** verstehen wir eine Menge

$$Q_1 = \{x : a_i < x_i < b_i \text{ für } i = 1, \dots, n\}$$

und unter einem **abgeschlossenen Quader** eine Menge der Form $Q_2 = \overline{Q_1}$. Jede Menge Q mit $Q_1 \subseteq Q \subseteq Q_2$ heie **Quader**, insbesondere nennen wir

$$Q = Q(a_i, b_i) = \{x : a_i \leq x_i < b_i \text{ fur } i = 1, \dots, n\}$$

einen **halboffenen Quader**.

Forderung:

(M4) Ist $Q = Q(a_i, b_i)$ ein halboffener Quader, so ist $Q \in \mathfrak{M}$ und es gilt $m(Q) = \prod_{i=1}^n (b_i - a_i)$.

Problem: Existieren m und \mathfrak{M} , so dass (M1) bis (M4) gelten?

Wir setzen zunchst

$$\mathfrak{M}_0 := \{\text{Vereinigungen endlich vieler halboffener Quader}\}.$$

Diese Menge \mathfrak{M}_0 erfllt fur $A, B \in \mathfrak{M}_0$ offensichtlich $A \cup B \in \mathfrak{M}_0$. Ebenso ist fur $A, B \in \mathfrak{M}_0$ auch (mit Quadern Q_i, R_j)

$$A \cap B = (\cup_{i=1}^r Q_i) \cap (\cup_{j=1}^s R_j) = \cup_{i,j} (Q_i \cap R_j) \in \mathfrak{M}_0,$$

da $Q_i \cap R_j$ ein Quader oder die leere Menge ist. Es bleibt zu zeigen, dass mit $A, B \in \mathfrak{M}_0$ auch $A \setminus B \in \mathfrak{M}_0$ gilt.

Hilfssatz B.5. *Jede endliche Vereinigung $A = \cup_{i=1}^r Q_i$ von halboffenen Quadern ist darstellbar als endliche Vereinigung von paarweise punktfremden Quadern.*

Analog:

Zu $A = \cup_{i=1}^r Q_i$ und $B = \cup_{j=1}^s R_j$ gibt es endlich viele paarweise punktfremde Quader $\{S_k\}$ ($k \in I$) derart, dass $A = \cup_{k \in K} S_k$ und $B = \cup_{n \in L} S_n$ gilt mit gewissen Teilmengen K, L der Indexmenge I . Daraus ergibt sich

$$A \setminus B = \bigcup_{k \in K \setminus L} S_k,$$

und somit liegt $A \setminus B$ in \mathfrak{M}_0 , was zu zeigen war.

Definition B.6.

(a) Fur einen halboffenen Quader $Q = Q(a_k, b_k)$ sei $m(Q) = \prod_{k=1}^n (b_k - a_k)$.

(b) Ist $A = \cup_{i=1}^r Q_i$ eine Vereinigung punktfremder Quader Q_i , so sei $m(A) = \sum_{i=1}^r m(Q_i)$.

Hier ist zu zeigen, dass die Definition (b) nicht von der Wahl der Zerlegung von A in punktfremde Quader abhängt. D.h.: ist $A = \cup_{i=1}^r Q_i = \cup_{j=1}^s R_j$ mit paarweise disjunkten Quadern Q_i bzw. R_j , so ist zu zeigen

$$\sum_{i=1}^r m(Q_i) = \sum_{j=1}^s m(R_j).$$

Hat man dies gezeigt, so hat man: Die Einschränkung von m auf \mathfrak{M}_0 erfüllt (M1), (M2) und (M4). Die Eigenschaft (M3) ist nicht erfüllt, da nicht jede Vereinigung abzählbar vieler Quader in \mathfrak{M}_0 liegt.

Daher definieren wir nun

$$\mathfrak{M}_1 := \{\text{Vereinigungen abzählbar vieler halboffener Quader}\}.$$

Es ist leicht zu sehen, dass \mathfrak{M}_1 zugleich die Menge aller Vereinigungen aufsteigender Folgen aus \mathfrak{M}_0 ist. Für $A, B \in \mathfrak{M}_1$ gilt stets $A \cup B \in \mathfrak{M}_1$ wie auch $A \cap B \in \mathfrak{M}_1$. Hingegen gehört $A \setminus B$ im allgemeinen nicht zu \mathfrak{M}_1 (Beispiel siehe unten), d.h. \mathfrak{M}_1 erfüllt (M1) nicht vollständig.

Hilfssatz B.7. *Jede offene Menge gehört zu \mathfrak{M}_1 .*

Beweis: klar, Übungsaufgabe!

Insbesondere gehören alle **offenen** Kreisscheiben zu \mathfrak{M}_1 . Ist aber beispielsweise $A = K(0, 1)$ und $B = K(0, 2)$, so liegt $B \setminus A$ nicht in \mathfrak{M}_1 , da der Kreisrand überabzählbar ist. Wir erweitern nun m auf \mathfrak{M}_1 :

Definition B.8. *Sei A_1, A_2, \dots eine aufsteigende Folge von Mengen aus \mathfrak{M}_0 und $A = \cup_{i=1}^{\infty} A_i$. Dann sei*

$$m(A) := \lim_{i \rightarrow \infty} m(A_i).$$

Dieser Grenzwert existiert oder ist $+\infty$, da $(m(A_i))$ eine monoton steigende Folge ist. Wieder ist zu zeigen, dass die Definition (B.8) sinnvoll ist, also nicht von der Wahl der gegen A aufsteigenden Folge abhängt.

Wenn man dies gezeigt hat, lässt sich weiter zeigen, dass die Mengenfunktion m auf \mathfrak{M}_1 die Bedingungen (M2) und (M3) erfüllt. Weiter gelten, obwohl (M1) nicht vollständig erfüllt ist, die zu Beginn als Folgerungen festgehaltenen Regeln

- (1) A_1, \dots, A_k paarweise disjunkt $\Rightarrow m(\cup_{i=1}^k A_i) = \sum_{i=1}^k m(A_i)$.
- (2) $m(A \cup B) + m(A \cap B) = m(A) + m(B)$.
- (3) $m(A \cup B) \leq m(A) + m(B)$.
- (4) $m(A) \leq m(B)$, falls $A \subseteq B$.

Bemerkung. Obwohl $m|\mathfrak{M}_1$ die meisten geforderten Eigenschaften hat, ist (M1) nicht vollständig erfüllt, und vielen anschaulichen Mengen ist auf diese Weise noch kein Inhalt zugeordnet — beispielsweise allen abgeschlossenen Kreisscheiben. Idee: man versucht, solche Mengen durch Mengen aus \mathfrak{M}_1 zu approximieren.

Definition B.9. Für $A \subseteq \mathbb{R}^n$ heißt

$$\bar{m}(A) := \inf\{m(A') : A' \in \mathfrak{M}_1 \text{ und } A \subseteq A'\}$$

äußeres Lebesguesches Maß und

$$\bar{i}(A) = \inf\{m(A') : A' \in \mathfrak{M}_0 \text{ und } A \subseteq A'\}$$

äußerer Jordanscher Inhalt. In beiden Fällen lassen wir beliebige Teilmengen A des \mathbb{R}^n zu.

Allerdings kann man für die Elemente von \mathfrak{M} nicht alle Teilmengen des \mathbb{R}^n zulassen, da man in diesem Fall zeigen kann, dass (M2) nicht gilt.

Definition B.10. A heißt **messbar im Sinne von Lebesgue (oder kurz L-messbar bzw. messbar)**, d.h. $A \in \mathfrak{M}$, wenn es zu jedem $\varepsilon > 0$ Mengen A' und A'' aus \mathfrak{M}_1 gibt mit

$$A \subseteq A', \quad A' \setminus A \subseteq A'', \quad m(A'') < \varepsilon.$$

Ist A messbar, so heißt $m(A) := \bar{m}(A)$ das **Lebesgue-Maß (L-Maß)** von A .

Definition B.11. A heißt **messbar im Sinne von Jordan (kurz J-messbar)**, wenn es zu jedem $\varepsilon > 0$ Mengen A' und A'' aus \mathfrak{M}_0 gibt mit $A \subseteq A'$, $A' \setminus A \subseteq A''$ und $m(A'') < \varepsilon$. Ist A J-messbar, so heißt $i(A) := \bar{i}(A)$ der **Jordansche Inhalt** von A .

Bemerkung B.12. Eine Menge $A \subseteq \mathbb{R}^n$ ist genau dann J -messbar, wenn sie sich von innen und außen durch endliche Quadersummen approximieren läßt, d.h. wenn zu $\varepsilon > 0$ Mengen $B, C \in \mathfrak{M}_0$ existieren mit $B \subseteq A \subseteq C \Rightarrow m(C \setminus B) \leq \varepsilon$. (Denn man kann $C = A'$, $B = A' \setminus A''$ wählen.)

Bemerkung B.13. Es existieren keine hierzu analogen Aussagen (Approximation mit abzählbaren Quadersummen) für L -messbare Mengen. Grund: $A' \setminus A''$ liegt im allgemeinen nicht in \mathfrak{M}_1 .

Man kann nun zeigen: Das durch (B.10) definierte Lebesgue-Maß $m|\mathfrak{M}$ erfüllt alle Eigenschaften (M1)-(M4). Das Ziel dieses Paragraphen ist dann erreicht. Analog gilt: Der oben definierte Jordansche Inhalt i erfüllt (M1), (M2) und (M4).

Weiter gilt

1. Jede kompakte Menge ist messbar und hat endliches Maß.
2. Jede Menge A mit $\overline{m}(A) = 0$ ist messbar. Eine solche Menge heißt **Nullmenge**.

Jede Teilmenge einer Nullmenge ist Nullmenge, und jede abzählbare Vereinigung von Nullmengen ist wieder eine Nullmenge. Nullmengen sind insbesondere alle einpunktigen Mengen und damit auch alle abzählbaren Mengen. Daraus folgt

3. Die Punkte des Einheitsintervalls (und damit auch die einer Kreislinie) sind nicht abzählbar. Sonst hätte das Einheitsintervall im \mathbb{R}^1 das Maß 0, was einen Widerspruch zu (M4) bildet.

Was sind die Mengen aus \mathfrak{M} mit endlichem Maß?

Satz B.14. Die Menge A gehört zu \mathfrak{M} und hat endliches Maß genau dann, wenn es zu jedem $\varepsilon > 0$ eine Menge A_0 aus \mathfrak{M}_0 gibt mit

$$\overline{m}((A_0 \setminus A) \cup (A \setminus A_0)) \leq \varepsilon.$$

B.2 Das Lebesguesche Integral

Wir gehen ähnlich wie bei der Definition des Riemannsches Integrals im \mathbb{R}^1 vor. Stets sei $A \subseteq \mathbb{R}^n$ und $f : A \rightarrow \mathbb{R}^m$ (wobei alle Zielmengen \mathbb{R} , \mathbb{C} , \mathbb{R}^k , \mathbb{C}^k zugelassen sind).

Die **Schwankung** $s(f, A)$ von f auf A sei definiert durch

$$s(f, A) := \sup_{x, y \in A} |f(y) - f(x)|.$$

Ein Mengensystem $\{A_i\} = \mathcal{Z}$ heißt **Zerlegung** von A , wenn A die *disjunkte* Vereinigung $A = \cup_{i=1}^{\infty} A_i$ ist mit $A_i \in \mathfrak{M}$ und $m(A_i) < \infty$. Gilt jeweils $s(f, A_i) < \infty$, so heißt

$$S(f, \mathcal{Z}) = \sum_{i=1}^{\infty} s(f, A_i) m(A_i)$$

die **Schwankungssumme** von f bezüglich \mathcal{Z} .

Hilfssatz B.15. *Sei A messbar. Folgende Aussagen sind äquivalent:*

- (i) *Zu jedem $\varepsilon > 0$ existiert eine Zerlegung \mathcal{Z} von A mit $S(f, \mathcal{Z}) \leq \varepsilon$.*
- (ii) *Für jede offene Teilmenge $C \subseteq \mathbb{R}^m$ ist $f^{-1}(C) = \{x \in \mathbb{R}^n : f(x) \in C\}$ messbar.*

Definition B.16. *Eine Funktion f heißt auf A **messbar**, wenn A messbar ist und die äquivalenten Bedingungen (i), (ii) von (B.15) erfüllt sind.*

Sei nun f auf A messbar und $\mathcal{Z} = \{A_i\}$ eine Zerlegung von A , so dass jedes A_i ein endliches Maß hat. Wähle nun jeweils ein $x_i \in A_i$ und setze

$$R(f, \mathcal{Z}) = \sum_{i=1}^{\infty} f(x_i) m(A_i).$$

Problem: Diese Summe konvergiert nicht für jedes f und beliebige Zerlegungen \mathcal{Z} von A .

Hilfssatz B.17. Sei f auf A messbar und $\mathcal{Z} = \{A_i\}$ eine Zerlegung von A mit $S(f, \mathcal{Z}) \leq \varepsilon$. In jedem A_i seien zwei Punkte x_i, x'_i gewählt. Wenn von den beiden Reihen

$$\sum_{i=1}^{\infty} f(x_i)m(A_i), \quad \sum_{i=1}^{\infty} f(x'_i)m(A_i)$$

eine absolut konvergiert, so auch die andere, und es gilt

$$\left| \sum_{i=1}^{\infty} f(x_i)m(A_i) - \sum_{i=1}^{\infty} f(x'_i)m(A_i) \right| \leq \varepsilon.$$

Definition B.18. Eine auf A messbare Funktion f heißt **über A integrierbar**, wenn für jede Zerlegung $\mathcal{Z} = \{A_i\}$ von A mit $S(f, \mathcal{Z}) < \infty$ die Reihe $R(f, \mathcal{Z}) = \sum_{i=1}^{\infty} f(x_i)m(A_i)$ mit $x_i \in A_i$ absolut konvergiert.

Definition B.19. Die Zerlegung $\mathcal{Z}' = \{B_j\}$ von A heißt eine **Verfeinerung** der Zerlegung $\mathcal{Z} = \{A_i\}$ von A (oder **feiner als \mathcal{Z}**), wenn jedes B_j in einem der A_i enthalten ist.

Hilfssatz B.20. Ist die Zerlegung $\mathcal{Z}' = \{B_j\}$ von A feiner als $\mathcal{Z} = \{A_j\}$, dann gilt $S(f, \mathcal{Z}') \leq S(f, \mathcal{Z})$.

Hilfssatz B.21. Sei f auf A messbar und seien $\mathcal{Z} = \{A_i\}$, $\mathcal{Z}' = \{B_j\}$ zwei Zerlegungen von A mit $S(f, \mathcal{Z}) < \infty$ und $S(f, \mathcal{Z}') < \infty$. Konvergiert von zwei Näherungssummen $R(f, \mathcal{Z})$ und $R(f, \mathcal{Z}')$ die eine absolut, so auch die andere, und es gilt

$$|R(f, \mathcal{Z}) - R(f, \mathcal{Z}')| \leq S(f, \mathcal{Z}) + S(f, \mathcal{Z}').$$

Definition B.22. Sei f über A integrierbar und $\mathcal{Z}_1, \mathcal{Z}_2, \dots$ eine Folge von Zerlegungen von A mit $\lim_{i \rightarrow \infty} S(f, \mathcal{Z}_i) = 0$. Dann heißt

$$\int_A f(x)dx = \int_A f(x_1, \dots, x_n)dx_1 \dots dx_n := \lim_{i \rightarrow \infty} R(f_i, \mathcal{Z}_i)$$

das **Lebesguesche Integral** von f über A .

Hierbei folgt aus (B.21), dass der Limes existiert und unabhängig von der Wahl der Zerlegungsfolge ist.

Integrationsregeln:

- (1) Seien f, g über A integrierbar und a, b Konstanten. Dann ist auch $af+bg$ über A integrierbar, und es gilt

$$\int_A af(x) + bg(x) dx = a \int_A f(x)dx + b \int_A g(x)dx.$$

- (2) Seien A, B messbare Mengen mit $A \cap B = \emptyset$. Genau dann ist f über $A \cup B$ integrierbar, wenn f über A und über B integrierbar ist. Dann gilt

$$\int_{A \cup B} f(x)dx = \int_A f(x)dx + \int_B f(x)dx.$$

- (3) Sind f und g über A integrierbar und $f \leq g$, so gilt

$$\int_A f(x)dx \leq \int_A g(x)dx.$$

Dies folgt aus der entsprechenden Ungleichung für Näherungssummen.

- (4) Ist f auf A messbar (bzw. integrierbar), so ist auch $|f|$ auf A messbar (bzw. integrierbar).

Zum Beweis verwendet man, dass $s(|f|, B) \leq s(f, B)$ für beliebige $B \subseteq A$ gilt und dass die Definition der Integrierbarkeit auf absoluter Konvergenz der Näherungssummen beruht.

Bemerkung. Aus der Messbarkeit von $|f|$ folgt i.a. nicht die Messbarkeit von f .

- (5) Ist f über A integrierbar, so gilt

$$\left| \int_A f(x)dx \right| \leq \int_A |f(x)|dx.$$

Dies folgt für reellwertige Funktionen aus Regel (3) und lässt sich im allgemeinen Fall mit Hilfe der Dreiecksungleichung zeigen.

- (6) Sei f auf A messbar und g über A integrierbar. Gilt $|f(x)| \leq g(x)$ auf A , so ist f über A integrierbar.

(7) Sei A messbar mit $m(A) < \infty$. Dann ist jede konstante Funktion $f(x) \equiv c$ über A integrierbar und es gilt $\int_A f(x)dx = c \cdot m(A)$. (Wähle die Zerlegung $\mathcal{Z} = \{A\}$; dann ist $S(f, \mathcal{Z}) = 0$.)

Satz B.23. *Jede auf einer messbaren Menge A stetige Funktion ist messbar.*

(Zum Beweis verwende man (B.15) und die Tatsache, dass stetige Urbilder offener Mengen offen sind.)

Satz B.24. *Ist A messbar mit $m(A) < \infty$ und f auf A messbar und beschränkt, so ist f über A integrierbar.*

(Zum Beweis verwende man Regel (6).)

Satz B.25. *Ist A messbar mit $m(A) < \infty$ und ist f stetig und beschränkt auf A , so ist f über A integrierbar.*

Speziell gilt: Ist A kompakt, so ist A auch messbar mit $m(A) < \infty$, und jede stetige Funktion f auf A ist beschränkt. Daraus folgt

Satz B.26. *Eine auf einer kompakten Menge A stetige Funktion ist über A integrierbar.*

Satz B.27. *Ist f über $[a, b]$ Riemann-integrierbar (im eigentlichen Sinn), so ist f über $[a, b]$ auch Lebesgue-integrierbar und es gilt*

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

Aber: es ist möglich, dass das uneigentliche Riemann-Integral existiert, ohne dass das entsprechende Lebesgue-Integral existiert. Ein Beispiel ist

$$\int_1^\infty \frac{\sin x}{x} dx.$$

B.3 Grenzwertsätze

Wir interessieren uns für die Frage, in welchen Fällen die Limesbildung mit dem Lebesgue-Integral vertauschbar ist, ohne dabei gleichmäßige Konvergenz vorauszusetzen.

Satz B.28. *Ist f_1, f_2, \dots eine Folge von über A messbaren Funktionen $f_i : A \rightarrow \mathbb{R}^m$ und existiert für alle $x \in A$ der Grenzwert $\lim_{n \rightarrow \infty} f_n(x) =: f(x)$, so ist f über A messbar.*

Zum Beweis ist zu zeigen: Ist $C \subseteq \mathbb{R}^n$ offen, so ist $f^{-1}(C)$ messbar. Urbilder offener Mengen kann man nach oben durch messbare Mengen abschätzen, Urbilder abgeschlossener Mengen sind entsprechend nach unten durch messbare Mengen abschätzbar. Das wird benutzt, um durch einen Kunstgriff $f^{-1}(C)$ selbst durch Vereinigungen bzw. Durchschnitte messbarer Mengen darzustellen. Man setzt

$$C_n := \{x \in C : \exists r > 1/n \text{ mit } K(x, r) \subseteq C\}$$

und zeigt als erste Feststellung, dass alle C_n offen sind. Als zweite Feststellung ergibt sich sodann $C = \bigcup_{n=1}^{\infty} C_n$, als dritte Feststellung dann $\overline{C_n} =: D_n \subseteq C_{n+1}$. Damit folgt $f^{-1}(C) = \bigcup_{n=1}^{\infty} f^{-1}(C_n) \subseteq \bigcup_{n=1}^{\infty} f^{-1}(D_n) \subseteq \bigcup_{n=1}^{\infty} f_n^{-1}(C_{n+1}) = f^{-1}(C)$.

Wir führen nun folgende Sprechweise ein: Eine Eigenschaft $E(x)$ gilt für **fast alle** x aus A , bzw. **fast überall** in A , wenn es eine Nullmenge $B \subseteq A$ gibt, so dass $E(x)$ für alle $x \in A \setminus B$ gilt. Ein Beispiel: Seien f, g Funktionen $A \rightarrow \mathbb{R}^m$ und sei $E(x)$ die Eigenschaft $f(x) = g(x)$. Gilt $E(x)$ für fast alle $x \in A$, so folgt

$$\int_A f(x) dx = \int_A g(x) dx,$$

falls eines der beiden Integrale existiert. Dies folgt aus

Bemerkung. Ist B eine Nullmenge und f eine beliebige Funktion auf B , so ist f über B integrierbar und $\int_B f(x) dx = 0$.

Hier nun (in diesem Exzerpt ohne Beweis) zwei wichtige und oft gebrauchte Grenzwertsätze:

Satz B.29 (B. Levi, Satz von der monotonen Konvergenz). Sei A messbar, f_1, f_2, \dots auf A definierte reelle Funktionen mit $0 \leq f_1 \leq f_2 \leq \dots$. Die f_n seien über A integrierbar mit $\int_A f_n(x) dx < s$. Dann ist $\{f_n(x)\}$ für fast alle $x \in A$ beschränkt, die Grenzfunktion $f(x) = \lim_{n \rightarrow \infty} f_n(x)$ ist über A integrierbar, und es gilt:

$$\int_A f(x) dx = \int_A \lim_{n \rightarrow \infty} f_n(x) dx = \lim_{n \rightarrow \infty} \int_A f_n(x) dx.$$

Bemerkung. Die Voraussetzung $f_1(x) \geq 0$ ist überflüssig, da man sonst f_i durch $f_i - f_1$ ersetzen kann. Ebenso ist der Satz richtig für monoton fallende Funktionenfolgen $f_i \leq f_{i-1}$, denn man kann f_i durch $-f_i$ ersetzen.

Satz B.30 (H. Lebesgue, Satz von der majorisierten Konvergenz). Die Funktionen f_i ($i \in \mathbb{N}$) und g seien über A integrierbar mit $|f_i(x)| \leq g(x)$ für alle $x \in A$. Für fast alle $x \in A$ konvergiere $(f_n(x))$ gegen den Grenzwert $f(x)$. Dann ist f über A integrierbar, und es gilt

$$\int_A f(x) dx = \lim_{n \rightarrow \infty} \int_A f_n(x) dx.$$

Anwendungen der Sätze von Levi und Lebesgue:

Satz B.31. Sei $f(x, y)$ definiert für $x \in A \subseteq \mathbb{R}^n$, $y \in B \subseteq \mathbb{R}^m$, und es existiere $\lim_{y \rightarrow y_0} f(x, y) = h(x)$ für alle $x \in A$. $f(x, y)$ sei für jedes feste y integrierbar über A , und für alle $x \in A$ und alle $y \in B$ gelte $|f(x, y)| \leq g(x)$ mit einer über A integrierbaren Funktion $g(x)$. Dann ist $h(x)$ über A integrierbar und es gilt

$$\int_A h(x) dx = \lim_{y \rightarrow y_0} \int_A f(x, y) dx.$$

Beispiel: Wir betrachten die Gammafunktion

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt, \quad 0 < x < \infty.$$

Behauptung: Γ ist stetig und sogar differenzierbar.

Dazu wählen wir $0 < a < b < \infty$. Für $x \in [a, b]$ ist dann

$$|t^{x-1} e^{-t}| \leq e^{-t} \cdot (t^{a-1} + t^{b-1}) =: g(t).$$

Für beliebige $t \in [0, \infty)$ ist der Integrand stetig in x , und aus (B.31) folgt daher die Stetigkeit von $\Gamma(x)$ auf $[a, b]$ und damit in $(0, \infty)$. Zum Beweis der Differenzierbarkeit:

$$\left| \frac{t^{y-1} - t^{x-1}}{y-x} \cdot e^{-t} \right| \leq (t^{a-1} + t^{b-1})e^{-t} \cdot \log t =: g(t).$$

Satz B.32. *Ist $A \in \mathfrak{M}$ und $A = \cup_{i=1}^{\infty} A_i$ die disjunkte Vereinigung von $A_i \in \mathfrak{M}$, so ist f genau dann über A integrierbar, wenn f für jedes n über A_n integrierbar ist und $\sum_{n=1}^{\infty} \int_{A_n} |f(x)| dx$ konvergiert. Es gilt dann*

$$\int_A f(x) dx = \sum_{n=1}^{\infty} \int_{A_n} f(x) dx.$$

Wir interessieren uns nun für den Zusammenhang des Lebesgue-Integrals mit dem uneigentlichen (Riemann-)Integral. Sei $f : [a, \infty) \rightarrow \mathbb{R}$ über $[a, b]$ eigentlich integrierbar für jedes b . Dann gilt: f ist über $[a, \infty)$ L-integrierbar genau dann, wenn $\int_a^{\infty} f(x) dx$ absolut konvergiert. In diesem Fall ist

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

Die analoge Aussage gilt für nach links unbeschränkte Intervalle, ebenso für beschränkte Intervalle und unbeschränkte Funktionen f .

Aus Satz (B.32) folgt nun: Das Lebesgue-Integral $\int_{[a, \infty)} f(x) dx$ existiert und ist gleich $\int_a^{\infty} f(x) dx$. Daher auch die bei der Gammafunktion verwendete Schreibweise $\int_0^{\infty} f(x) dx$.

Bemerkung. Falls $\int_a^{\infty} f(x) dx$ existiert im Riemannschen Sinne, nicht aber $\int_a^{\infty} |f(x)| dx$, so existiert das Lebesgue-Integral $\int_{[a, \infty)} f(x) dx$ nicht. Aber: Für endliche Intervalle liefert das Lebesgue-Integral mehr als das eigentliche Riemann-Integral (vgl. (B.27)).

B.4 Der Satz von Fubini und das Cavalierische Prinzip

Sei zunächst

$$Q = \{x : a_i \leq x_i \leq b_i \text{ für } i = 1, 2\} \subseteq \mathbb{R}^2$$

ein Quader im \mathbb{R}^2 . Dann läßt sich Q als Produkt $I_1 \times I_2$ von reellen Intervallen I_1, I_2 schreiben. Das Maß von Q ist dann das Produkt der Maße dieser beiden Intervalle.

Allgemeiner: Sei $Q \subseteq \mathbb{R}^n = \mathbb{R}^{r+s}$ ein Quader und sei

$$x = (y_1, \dots, y_r, z_1, \dots, z_s) = (y, z)$$

mit $y \in \mathbb{R}^r$ und $z \in \mathbb{R}^s$. Dann ist

$$Q = \{(y, z) : y \in Q', z \in Q''\}$$

mit Quadern $Q' \subseteq \mathbb{R}^r$, $Q'' \subseteq \mathbb{R}^s$. Wir haben

$$m(Q) = m_r(Q') \cdot m_s(Q''), \quad (\text{B.1})$$

wobei m_r das Lebesgue-Maß im \mathbb{R}^r und m_s das Lebesgue-Maß im \mathbb{R}^s bezeichnet.

Definition B.33. Sei $A \subseteq \mathbb{R}^n$ eine beliebige Punktmenge und $x \in \mathbb{R}^n$. Dann heißt

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

die **charakteristische Funktion** oder **Indikatorfunktion** von A . Genau dann ist A messbar mit endlichem Maß, wenn χ_A integrierbar ist. Dann gilt

$$m(A) = \int \chi_A(x) dx.$$

(B.1) ist gleichbedeutend mit

$$\int \chi_A(x) dx = \int \left(\int \chi_A(y, z) dy \right) dz. \quad (\text{B.2})$$

Zum Beweis verwendet man, dass $\int \chi_A(y, z) dy$ für fast alle $z \in \mathbb{R}^s$ definiert und integrierbar ist, wobei Integration die rechte Seite von (B.1) ergibt. Allgemeiner gilt:

Satz B.34 (Satz von Fubini). Sei $f(x)$ integrierbar über \mathbb{R}^n . Dann existiert das Integral $\int f(y, z)dy$ für fast alle $z \in \mathbb{R}^s$, und es gilt

$$\int f(x)dx = \int \left(\int f(y, z)dy \right) dz.$$

Beweis. Der Beweis erfolgt in mehreren Schritten. Zunächst ist zu zeigen, dass (B.34) für charakteristische Funktionen von Quadern im \mathbb{R}^n gilt. Anschließend zeigt man die Richtigkeit für Funktionen des Typs

$$f(x) = \sum_{i=1}^k a_i \chi_{Q_i}(x),$$

wobei diese Funktionen auch **Elementarfunktionen** genannt werden. Schließlich approximiert man dann beliebige Funktionen durch Elementarfunktionen. Hierzu ein

Hilfssatz B.35. Sei f_1, f_2, \dots eine Folge von Elementarfunktionen, und die Reihe $\sum_{i=1}^{\infty} \int |f_i(x)|dx$ konvergiere. Dann gilt

- (a) Für fast alle $x \in \mathbb{R}^n$ konvergiert die Reihe $f(x) = \sum_{i=1}^{\infty} f_i(x)$ absolut.
- (b) Es gibt eine Nullmenge $B \subseteq \mathbb{R}^s$ und für jedes $z \notin B$ eine Nullmenge $C(z) \subseteq \mathbb{R}^n$ derart, dass

$$f(y, z) = \sum_{i=1}^{\infty} f_i(y, z) \quad \text{für } z \notin B, y \notin C(z)$$

absolut konvergiert und $\int f(x)dx = \int \left(\int f(y, z)dy \right) dz$ gilt.

Dieser Hilfssatz wird mittels (B.29) bewiesen. Aus ihm kann man zunächst folgern, dass der Satz von Fubini für Funktionen gilt, die außerhalb einer Nullmenge $A \subseteq \mathbb{R}^n$ verschwinden. Der letzte Schritt ist dann der Beweis folgender Aussage: Zu jeder über \mathbb{R}^n integrierbaren Funktion g existiert eine Folge von Elementarfunktionen f_i mit den Voraussetzungen von (B.35) und mit der Eigenschaft, dass $f = \sum_{i=1}^{\infty} f_i$ fast überall mit g übereinstimmt. Dann folgt: (B.34) gilt nach (B.35) für f und damit auch für $g - f$, also auch für g . \square

Hilfssatz B.36. Sei f integrierbar und $\varepsilon > 0$. Dann existiert eine Elementarfunktion $h(x)$ mit

$$\int |f(x) - h(x)| dx \leq \varepsilon.$$

(Zum Beweis verwendet man (B.14).) Dies sichert die Existenz einer Folge $h_i(x)$ mit $\int |f(x) - h_i(x)| dx \rightarrow 0$ für $i \rightarrow \infty$. Insbesondere gilt, da man die Rollen von \mathbb{R}^r und \mathbb{R}^s vertauschen kann,

$$\int \int f(y, z) dy dz = \int \int f(y, z) dz dy.$$

Bemerkung.

- (a) Aus der Existenz der iterierten Integrale folgt i.a. noch nicht die Integrierbarkeit von f , d.h. die Existenz von $\int f(x) dx$.
- (b) Es gibt Beispiele nicht integrierbarer Funktionen, bei denen die iterierten Integrale existieren, aber voneinander verschieden sind!

Man kann nun durch mehrfache Anwendung des Satzes von Fubini die Integration im \mathbb{R}^n auf n einfache Integrationen zurückführen. Hierzu benutzt man z.B. im \mathbb{R}^2 die Schreibweise

$$\int f(x) dx = \int \int f(x_1, x_2) dx_1 dx_2.$$

Es ist auch nicht erforderlich, dass f über ganz \mathbb{R}^n definiert ist, sondern man kann f als eine Funktion auf irgendeiner messbaren Teilmenge $A \subseteq \mathbb{R}^n$ betrachten. Das Integral schreibt sich dann als

$$\int_A f(x) dx,$$

oder, mit der Funktion

$$g(x) = \begin{cases} f(x) & \text{falls } x \in A \\ 0 & \text{falls } x \notin A \end{cases}$$

gilt $\int g(x) dx = \int_A f(x) dx$. Auf g kann man den Satz von Fubini anwenden und erhält

$$\int_A f(x) dx = \int g(x) dx = \int \int g(y, z) dy dz.$$

Für $z \in \mathbb{R}^s$ betrachten wir nun die **Schnittmenge**

$$A_z = \{y \in \mathbb{R}^r : (y, z) \in A\}.$$

Dann ist leicht zu zeigen, dass A_z für fast alle z messbar als Teilmenge des \mathbb{R}^r ist. Ist $m(A) < \infty$, so zeigt man die Messbarkeit von $A_z = \{y : \chi_{A_z}(y) > 0\}$. Ist $m(A) = \infty$, so stellt man A_z als Vereinigung einer aufsteigenden Folge von Schnittmengen $(A_i)_z$ dar und folgert daraus, dass A_z für fast alle z messbar ist. Damit folgt

$$\int g(y, z) dy = \int_{A_z} g(y, z) dy = \int_{A_z} f(y, z) dy$$

und

$$\int_A f(x) dx = \int \int_{A_z} f(y, z) dy dz = \int_B \int_{A_z} f(y, z) dy dz. \quad (\text{B.3})$$

Dabei ist B irgendeine messbare Teilmenge des \mathbb{R}^s mit

$$\{z : (y, z) \in A \text{ für mindestens ein } y\} \subseteq B.$$

Beispiel B.37. Sei

$$A = \{(y, z) : a(y) \leq z \leq b(y), c \leq y \leq d\} \subseteq \mathbb{R}^2$$

mit integrierbaren Funktionen a, b und sei f über A integrierbar. Dann gilt

$$\int_A \int f(y, z) dy dz = \int_c^d \int_{a(y)}^{b(y)} f(y, z) dz dy. \quad (\text{B.4})$$

Ist speziell $f = \chi_A$, so erhält man

$$\int_A f(x) dx = m(A) = \int_c^d (b(y) - a(y)) dy, \quad (\text{B.5})$$

und dies ist der Flächeninhalt von A .

Beispiel B.38. Sei $A \subseteq \mathbb{R}^3$ messbar und $A_z = \{(x, y) : (x, y, z) \in A\}$. Dann gilt

$$\int_A f(x, y, z) dx dy dz = \int \left(\int_{A_z} f(x, y, z) dx dy \right) dz$$

und speziell für $f = \chi_A$

$$m_3(A) = \int m_2 A_z dz. \quad (\text{B.6})$$

Daraus ergibt sich das

Cavalierische Prinzip: Sind A und B zwei messbare Mengen im \mathbb{R}^3 und gilt $m_2 A_z = m_2 B_z$ für alle z , so haben A und B das gleiche Maß.

Ist beispielsweise A der Rotationskörper

$$\{(x, y, z) : c \leq z \leq d, x^2 + y^2 \leq f(z)^2\},$$

so gilt $m_2 A_z = \pi(f(z))^2$ und wegen (B.6) daher

$$m_3(A) = \pi \int_c^d f(z)^2 dz. \quad (\text{B.7})$$

Beispiel B.39. Sei

$$A = \{(x, y, z) \in \mathbb{R}^3 : (x, y) \in B, a(x, y) \leq z \leq b(x, y)\}.$$

Der Satz von Fubini liefert

$$\int_A f(x, y, z) dx dy dz = \int_B \left(\int_{a(x, y)}^{b(x, y)} f(x, y, z) dz \right) dx dy. \quad (\text{B.8})$$

Ist nun $f = \chi_A$, so gilt

$$m(A) = \int_B (b(x, y) - a(x, y)) dx dy. \quad (\text{B.9})$$

Anwendungsbeispiel: Zu berechnen ist das iterierte Integral

$$\int e^{-(x_1^2 + \dots + x_n^2)} dx_1 \dots dx_n.$$

Man setzt zunächst

$$A = \{(x_0, x_1, \dots, x_n) : 0 < x_0 < e^{-(x_1^2 + \dots + x_n^2)} \subseteq \mathbb{R}^{n+1}\}$$

und

$$A_N := \{(x_0, \dots, x_n) \in A : |x_i| < N \text{ für } i = 1, \dots, n\}.$$

Alle A_N sind offen und beschränkt mit dem Maß

$$m(A_N) = \left(\int_{-N}^N e^{-x^2} dx \right)^n$$

(vgl. (B.8) und Satz von Fubini). Das uneigentliche Riemann-Integral

$$\int_{-\infty}^{\infty} e^{-x^2} dx =: F$$

konvergiert. Nun ist $m(A_N) \leq F^n$ und $m(A) = F^n$.

Andererseits kann man $m(A)$ wie folgt berechnen:

$$\int \left(\int \chi_A(x_0, x_1, \dots, x_n) dx_1 \dots dx_n \right) dx_0$$

wobei der Integrand für $0 < x_0 < 1$ das Volumen des Körpers ist, der durch $x_1^2 + \dots + x_n^2 < -\log x_0$ gegeben ist. D.h.

$$\int \chi_A(x_0, x_1, \dots, x_n) dx_1 \dots dx_n = I_n \cdot (-\log x_0)^{n/2}$$

mit I_n als Volumen der n -dimensionalen Einheitskugel. Es folgt

$$F^n = m(A) = I_n \cdot \int_0^1 (-\log x)^{n/2} dx,$$

und mit der Substitution $x = e^{-y}$ ergibt sich die rechte Seite zu

$$I_n \cdot \int_0^{\infty} y^{n/2} \cdot e^{-y} dy = I_n \cdot \Gamma\left(\frac{n}{2} + 1\right)$$

(mit der Gammafunktion $\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt$).

Ist nun n gerade, also $n = 2m$, so liefert $\Gamma(m + 1) = m!$ eine Beziehung zwischen I_{2m} und F . Etwa für $m = 1$ ist $I_2 = \pi$ und damit

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

Es folgt

$$I_{2m} = \frac{\pi^m}{m!}.$$

Ist andererseits $n = 2m + 1$, so ergibt sich mit $m = 0$

$$\Gamma(3/2) = \frac{1}{2}\sqrt{\pi}$$

und

$$\Gamma(1/2) = \int_0^{\infty} e^{-t} \frac{dt}{t} = \sqrt{\pi}.$$

Per Induktion erhält man

$$I_{2m+1} = \frac{2^{2m+1} \cdot \pi^m}{(m+1)(m+2) \dots (2m+1)}.$$

B.5 Transformation von Integralen

Wir erinnern zunächst an die Substitutionsregel im \mathbb{R}^1

$$\int_{g(a)}^{g(b)} f(y) dy = \int_a^b f(g(x)) \cdot g'(x) dx$$

etwa unter den Voraussetzungen, dass f stetig und g stetig differenzierbar ist. Falls weiter $g'(x) \neq 0$ auf dem Intervall $[a, b] = A$ gilt, ist

$$\int_{g(A)} f(y) dy = \int_A f(g(x)) \cdot |g'(x)| dx. \quad (\text{B.10})$$

Wenn g' keine Nullstelle hat, ist g umkehrbar eindeutig. Wir wollen nun diese Überlegungen auf den allgemeinen Fall des \mathbb{R}^n übertragen.

Dazu verlangen wir die Ein-Eindeutigkeit von g und das Nichtverschwinden der Funktionaldeterminante der Abbildung $g : A \rightarrow \mathbb{R}^n$ ($A \subseteq \mathbb{R}^n$). Für f genügt als Voraussetzung die Integrierbarkeit, und so erhält man auch für den eindimensionalen Fall eine erhebliche Verallgemeinerung.

Satz B.40. *Sei $A \subseteq \mathbb{R}^n$ offen, $g : A \rightarrow g(A) \subseteq \mathbb{R}^n$ eine umkehrbar eindeutige, stetig differenzierbare Abbildung, deren Funktionaldeterminante nirgends verschwindet, und f eine auf $g(A)$ definierte Funktion. Wenn eines der beiden Integrale*

$$\int_{g(A)} f(y) dy \quad \text{und} \quad \int_A f(g(x)) \left| \det \left(\frac{\partial g_i}{\partial x_j} \right) \right| dx$$

existiert, so auch das andere und beide Integrale sind gleich.

Zum **Beweis** wendet man für stetiges f Induktion nach n an und setzt dabei zunächst weitere Eigenschaften von g und A voraus (zunächst ist A ein offener Quader, g stetig differenzierbar auf ganz \overline{A}), die dann später schrittweise abgebaut werden. Das entscheidende Hilfsmittel in diesem Beweis ist der Satz von Fubini, ferner die eindimensionale Substitutionsregel. Um den Fall zu behandeln, dass f nicht stetig ist, wird noch benötigt:

Lemma B.41. *Sei $g : A \rightarrow g(A)$ ($A, g(A) \subseteq \mathbb{R}^n$) eine umkehrbar eindeutige, stetig differenzierbare Abbildung, deren Funktionaldeterminante nirgendwo verschwindet.*

(i) Ist B eine offene Teilmenge von A , so gilt

$$m(g(B)) = \int_B \left| \det \left(\frac{\partial g_i}{\partial x_j} \right) \right| dx.$$

(ii) Ist $C \subseteq A$ eine Nullmenge, so ist auch $g(C)$ eine Nullmenge.

Damit kann nun gezeigt werden, dass (B.40) richtig ist für den Fall $f = \chi_Q$, $Q \subseteq A$ offen, und insbesondere für Quader Q . Der Satz gilt damit für Elementarfunktionen. Der letzte Schritt ist dann die Approximation der Funktion f durch eine Folge von Elementarfunktionen.

Zwei Spezialfälle von Satz (B.40):

- (1) Sei g eine affine Transformation, d.h. $g(x) = Cx + b$ mit einer quadratischen Matrix C . Dann ist $\det C$ die Funktionaldeterminante. Ist dann $f(y) = \chi_A(y)$, so gilt

$$m(g(A)) = |\det C| \cdot m(A).$$

Ist nun speziell g eine orthogonale Transformation, d.h. $|\det C| = 1$, so erhalten wir:

Korollar B.42. *Das Lebesgue-Maß ist bewegungsinvariant.*

- (2) Wir betrachten die Abbildung g , die im \mathbb{R}^2 den Polarkoordinaten r, φ die kartesischen Koordinaten $x, y \in \mathbb{R}^1$ zuordnet. Dann gilt

$$\frac{\partial(x, y)}{\partial(r, \varphi)} = r.$$

Sei $A \subseteq \mathbb{R}^2$ und $g : A \rightarrow B$. Dann folgt aus (B.40)

$$\int \int_B f(x, y) dx dy = \int \int_A f(r \cos \varphi, r \sin \varphi) r dr d\varphi.$$

Die Eineindeutigkeit erzeugt man häufig dadurch, dass man eine Nullmenge C wählt und B durch $B \setminus C$ ersetzt. Etwa

$$B = \{(x, y) : x^2 + y^2 < R^2\}, \quad C = \{(x, y) : y = 0, 0 \leq x < R\},$$

$$A = \{(r, \varphi) : 0 < r < R, 0 < \varphi < 2\pi\}.$$

In diesem Fall folgt

$$\int \int_B f(x, y) dx dy = \int_0^{2\pi} \int_0^R f(r, \cos \varphi, r \sin \varphi) r dr d\varphi.$$

Literatur

- [1] V.I. Arnold: Ordinary Differential Equations, 2. Aufl. 1980
- [2] H.-J. Bartels Analysis I, Vorlesungsskriptum
- [3] F. Erwe: Gewöhnliche Differentialgleichungen, BI Mannheim 1964
- [4] O. Forster: Analysis 2, Differentialrechnung im \mathbb{R}^n , Gewöhnliche Differentialgleichungen, Vieweg-Verlag, 7. Aufl. 2006
- [5] O. Forster: Analysis 3, Integralrechnung im \mathbb{R}^n mit Anwendungen, Vieweg-Verlag, 4. Aufl. 2007
- [6] E. Freitag: Vorlesungen über Analysis, Teil II, Vorlesungsskriptum, Heidelberg 2001
- [7] I.M. Gelfand/G.E. Silov: Verallgemeinerte Funktionen, Bd. 1-4, Akademie Verlag Berlin 1964-1969
- [8] H. Grauert/W. Fischer: Differential- und Integralrechnung II, Springer 1973
- [9] H. Heuser: Gewöhnliche Differentialgleichungen, Teubner Stuttgart 1991, 2. Auflage
- [10] H. Heuser: Lehrbuch der Analysis, Teil 2, Teubner Stuttgart 1992
- [11] E. Kamke: Differentialgleichungen/ Lösungsmethoden und Lösungen I, 10. Aufl., Teubner Stuttgart 1983
- [12] M. Kneser: Differential- und Integralrechnung II und III, Vorlesungsskriptum, Göttingen 1977/1978
- [13] K. Meyberg/P. Vachenauer: Höhere Mathematik 2, Springer Verlag 1991
- [14] L. Schwartz, Theorie des Distributions I und II, Hermann Paris 1957/1959.