Ausgewählte Kapitel aus der Analysis II

Skript zum zweiten Teil der Vorlesung
im Sommersemester 2000

H. v. Weizsäcker J. Voß

Fachbereich Mathematik
Universität Kaiserslautern

4. August 2000 - 1. vollständige Version
9. Oktober 2000 - Tippfehler korrigiert

http://www.mathematik.uni-kl.de/~wwwstoch/2000s/skript.html

Vorwort

Dieser Text entspricht dem zweiten Teil der Vorlesung am Fachbereich Mathematik der Universität Kaiserslautern im Sommersemester 2000. Die Vorlesung Analysis I im Wintersemester 1999/2000 und der erste Teil der Analysis II haben sich mit leichten Umstellungen weitgehend an den Lehrbüchern von O. Forster: Analysis I und Analysis II orientiert (unter Verzicht auf Fourier-Reihen), bis zur Diskussion lokaler Extrema ohne Nebenbedingungen.

Die Darstellung des daran anschließenden Stoffs war nicht an ein bestimmtes Lehrbuch angelehnt, daher schien es uns sinnvoll, den Hörern zur Nachbereitung eine schriftliche Grundlage zu bieten und die Dozenten der auf der Analysis aufbauenden Vorlesungen über den behandelten Stoff zu informieren.

Folgendes weicht ein wenig von der üblichen Darstellung im 2. Semester ab: Im Abschnitt über implizite Funktionen sind auch kurz das Newton-Verfahren sowie die Konstruktion von Karten und Tangentialräumen für Untermannigfaltigkeiten behandelt. Die Integrationstheorie beginnt zunächst nach dem Vorbild von Forsters Analysis III mit der Integration stetiger Funktionen über kompakten Quadern, und führt dann das Volumen zunächst nur für kompakte Mengen ein. Das erlaubt konkrete Volumenberechnungen mit dem Cavalierschen Prinzip noch vor der etwas langwierigen Fortsetzungstheorie. Diese wird mit dem Satz von Kysinski durchgeführt, der, wie im Anhang dokumentiert, flexibel genug ist, die klassische Carathéodory-Methode und den funktionalanalytischen Zugang zu verbinden. Die Einführung des Integrals für Maße geschieht dann wie üblich (vgl. Rudin), während die Integraltransformationsformel mit Hilfe des Vitalischen Überdeckungssatzes bewiesen wird, um den Zugang zu weiterführenden Fragen der reellen Analysis zu erleichtern. Beispiele solcher Anwendungen dieses Satzes finden sich ebenfalls im Anhang.

Nicht behandelt werden konnten dagegen insbesondere mehrdimensionale partielle Integration, Integration rotationsinvarianter Funktionen und der Gaußsche Integralsatz. Hierfür seien den Hörern bei späterem Bedarf die entsprechenden Abschnitte aus Forster III empfohlen.

Für die Ausarbeitung des ersten Kapitels war die Mitschrift von Hannah Markwig hilfreich. Vielen Dank! Auch sonst verdankt diese Ausarbeitung manche Details der Aufmerksamkeit von Hörern und ersten Lesern dieses Jahrgangs.

Inhaltsverzeichnis

Kapitel 1

Implizite Funktionen

Unser Hauptziel in diesem Kapitel ist es, Funktionen zu studieren, die nicht wie bisher durch eine Rechenvorschrift explizit gegeben sind, sondern implizit durch das Lösen einer Gleichung der Form $F(x,y)=0$. Gesucht ist also eine Funktion $y\colon x\mapsto y(x)$, so daß $F\big(x,y(x)\big)=0$ ist.

1.1 Inverse Funktionen

Zunächst wollen wir den Spezialfall $F(x,y)=y-f(x)$ behandeln, d.h. wir wollen die Gleichung $f(x)=y$ nach x auflösen. Unsere Frage ist hier also, ob zu der Funktion f eine Inverse existiert.

Das Vorbild im Eindimensionalen ist folgendes: Sei $f\colon(a,b)\to\mathbb{R}$ stetig differenzierbar. Wenn $f'(x_0)\neq 0$ ist, so gibt es, weil f' stetig ist, ein Intervall $I=(x_0-\eta,x_0+\eta)$, so daß $f'(x)>0$ für alle $x\in I$ oder $f'(x)<0$ für alle $x\in I$ gilt. Also ist f in I streng monoton und $f(I)$ ist ein Intervall mit $y_0=f(x_0)$ als innerem Punkt. Die Umkehrfunktion $g=f^{-1}\big|_I$ hat die Ableitung $g'(y_0)=1/f'(x_0)$.

Satz 1 *(über inverse Funktionen). Sei $U\subseteq\mathbb{R}^n$ offen, $f\colon U\longrightarrow\mathbb{R}^n$ eine stetig differenzierbare Abbildung und $x_0\in U$ so, daß $\det\big(Df(x_0)\big)\neq 0$. Dann gilt:*

a) Der Punkt $y_0=f(x_0)$ ist innerer Punkt der Menge $f(U)$.

b) Es gibt ein $\eta>0$ so daß $f|_{K_\eta(x_0)}$ injektiv ist.

Beweis. Zu Teil b: Wir betrachten zunächst die Gradienten der Komponenten f_i und setzen sie zu einer Funktion in n^2 Dimensionen zusammen. Sei also $x^i\in\mathbb{R}^n$ und damit $(x^1,\dots,x^n)\in\mathbb{R}^{n^2}$. Wir betrachten die Abbildung

$$(x^1,\dots,x^n)\mapsto\begin{pmatrix}\nabla f_1(x^1)\\ \vdots\\ \nabla f_n(x^n)\end{pmatrix}=\begin{pmatrix}\frac{\partial f_1}{\partial x_1}(x^1) & \dots & \frac{\partial f_1}{\partial x_n}(x^1)\\ \vdots & & \vdots\\ \frac{\partial f_n}{\partial x_1}(x^n) & \dots & \frac{\partial f_n}{\partial x_n}(x^n)\end{pmatrix}.$$

Definiere

$$\varphi(x^1,\dots,x^n):=\det\Big(\big(\frac{\partial f_i}{\partial x_k}(x^i)\big)_{\substack{i=1,\dots,n\\ k=1,\dots,n}}\Big).$$

Die Determinante hängt stetig von der Matrix ab. Daher ist die Funktion φ insbesondere an der Stelle $(x_0, \dots, x_0) \in \mathbb{R}^{n^2}$ mit $x_0 \in \mathbb{R}^n$ stetig. Es gilt aber $\varphi(x_0, \dots, x_0) = \det Df(x_0) \neq 0$. Also existiert ein $\eta > 0$ so, daß für alle $(x^1, \dots, x^n)$ mit $\|x^i - x_0\| < \eta$ für $i \in \{1, \dots, n\}$ auch $\varphi(x^1, \dots, x^n) \neq 0$ gilt.

Wir zeigen nun, daß f auf der Kugel $K_\eta(x_0)$ injektiv ist. Seien hierzu x, x' zwei Vektoren in dieser Kugel mit $f(x) = f(x')$ also $f_i(x) = f_i(x')$ für alle i. Wende den eindimensionalen Mittelwertsatz auf $h_i \colon t \mapsto f_i\big(x + t(x' - x)\big)$ an. Zu jedem $i \in \{1, \dots, n\}$ existiert ein $t^i \in (0,1)$, so daß $h_i'(t^i) = 0$. Mit $x^i = x + t^i(x' - x)$ folgt dann $0 = h_i'(t^i) = \nabla f_i(x^i) \cdot (x' - x)$. Zusammengenommen ergibt das

$$\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} \nabla f_1(x^1) \\ \vdots \\ \nabla f_n(x^n) \end{pmatrix} \cdot (x' - x).$$

Also hat entweder die Matrix keinen vollen Rang oder es ist $(x' - x) = 0$. Wegen $x^i \in K_\eta(x_0)$ folgt aber

$$\det \begin{pmatrix} \nabla f_1(x^1) \\ \vdots \\ \nabla f_n(x^n) \end{pmatrix} \neq 0,$$

d.h. die Matrix hat vollen Rang. Es muß stattdessen $x' = x$ gelten. Damit ist gezeigt, daß $f|_{K_\eta(x_0)}$ injektiv ist.

Zu Teil a: Sei η wie in Teil b. Wir betrachten die beiden Mengen $U_1 = \big\{x \;\big|\; \|x - x_0\| < \frac{\eta}{2}\big\}$ und $S = \big\{x \;\big|\; \|x - x_0\| = \frac{\eta}{2}\big\}$. Dann ist natürlich $\overline{U_1} = U_1 \cup S$. Die Abbildung f ist auf der kompakten Menge S stetig, also wird das Minimum $\delta = \min\{\|f(x) - y_0\| \mid x \in S\}$ angenommen. δ ist positiv, weil $f|_{\overline{U_1}}$ injektiv ist, also $\|f(x) - y_0\| > 0$ für alle $x \in S$ gilt.

Wir zeigen nun, daß jeder Punkt $y^* \in K_{\delta/2}(y_0)$ als Bildpunkt vorkommt. Zur Konstruktion des zugehörigen Urbilds betrachte die Hilfsfunktion $\rho(x) = \|f(x) - y^*\|^2$. Die Funktion ρ ist stetig auf der kompakten Menge $\overline{U_1}$. Also existiert eine Minimumstelle $x^* \in \overline{U_1}$. Für $x \in S$ gilt nach der umgekehrten Dreiecksungleichung

$$\big\|f(x) - y^*\big\| \geq \big|\|f(x) - y_0\| - \|y^* - y_0\|\big| > \frac{\delta}{2},$$

und damit $\rho(x) > (\delta/2)^2$. Andererseits ist $\rho(x^*) \leq \rho(x_0) = \|f(x_0) - y^*\|^2 = \|y^* - y_0\|^2 < (\delta/2)^2$, d.h. der Punkt x^* kann nicht auf dem Rand $S = \partial U_1$ liegen. Also liegt die Minimumsstelle x^* in der offenen Menge U_1 und es folgt $\nabla\rho(x^*) = 0$. Wegen $\nabla_y \|y - y^*\|^2 = 2(y - y^*)^T$ liefert die Kettenregel

$$0 \;=\; \nabla\rho(x^*) \;=\; 2(f(x^*) - y^*)^T Df(x^*).$$

Es folgt wieder $\det Df(x^*) = 0$ oder $f(x^*) - y^* = 0$. Wegen $\det Df(x^*) \neq 0$ ist $f(x^*) - y^* = 0$, d.h. es existiert ein x^* mit $f(x^*) = y^*$.

Damit ist gezeigt, daß mit y_0 auch die gesamte Kugel $K_{\delta/2}(y_0)$ im Bild $f(U)$ liegt. ∎

Warnung. Selbst wenn $U = \mathbb{R}^2$ und $\det f(x) \neq 0$ für alle x ist, muß f nicht notwendigerweise global injektiv sein. Dies sieht man beispielsweise an der komplexen Exponentialfunktion $z \mapsto \mathrm{e}^z$. Indem man $\mathbb{C}$ mit $\mathbb{R}^2$ identifiziert erhält man wegen $\exp(x+iy) = e^x(\cos y + i \sin y)$ die Abbildung

$$\begin{array}{rcl} \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ (x,y) & \longmapsto & (\mathrm{e}^x \cos y, \mathrm{e}^x \sin y) \end{array}$$

Die Jacobimatrix Df läßt sich hier leicht ausrechnen: es gilt

$$\begin{aligned} Df(x,y) &= \begin{pmatrix} \frac{\partial}{\partial x} e^x \cos y & \frac{\partial}{\partial y} e^x \cos y \\ \frac{\partial}{\partial x} e^x \sin y & \frac{\partial}{\partial y} e^x \sin y \end{pmatrix} \\ &= \begin{pmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{pmatrix} \end{aligned}$$

und folglich $\det Df(x,y) = e^{2x}(\cos^2 y + \sin^2 y) = e^{2x} > 0$. Andererseits ist wegen der Periodizität von Sinus und Cosinus $f(x,y) = f(x, y+2\pi)$, d.h. f ist nicht global injektiv.

Bemerkung. Sei $U \subseteq \mathbb{R}^2$ offen und $f\colon U \to \mathbb{R}^2$ differenzierbar. Weiter sei $\tilde{U} = \{\, x+iy \mid (x,y) \in U \,\} \subseteq \mathbb{C}$ und $\tilde{f}\colon \tilde{U} \to \mathbb{C}$ die entsprechende komplexe Funktion. Dann ist $\tilde{f}$ an der Stelle $z_0 = x_0 + iy_0$ komplex differenzierbar mit Ableitung $\tilde{f}'(z_0) = a + ib$, genau dann wenn

$$Df(x_0, y_0) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

gilt.

Der folgende Zusatz zum Satz über inverse Funktion formuliert das Analogon zum Satz über die Ableitung der Umkehrfunktion aus Analysis I. Das Ergebnis dort war $\big(f^{-1}\big)'\big(f(x)\big) = 1/f'(x)$. Hier erhält man eine völlig analoge Aussage:

Satz 2 *Sei $U \subseteq \mathbb{R}^n$ offen und $f\colon U \to \mathbb{R}^n$ stetig differenzierbar. Weiter sei $V \subseteq U$ offen, so daß $f|_V$ injektiv ist und $\det Df(x) \neq 0$ für alle $x \in V$ gilt. Dann ist die Umkehrfunktion $f^{-1}\colon f(V) \to V$ differenzierbar mit Ableitung*

$$Df^{-1}\big(f(x)\big) = \big(Df(x)\big)^{-1} \qquad \textit{für alle } x \in V.$$

Beweis. Wenn wir schon wüßten, daß f^{-1} differenzierbar ist, dann folgt die Formel aus der Kettenregel, denn es gilt $\mathrm{id}\colon V \to V$ ist $f^{-1} \circ f$, daher $id = D\big(\mathrm{id}\big)(x) = D\big(f^{-1} \circ f\big)(x) = Df^{-1}(f(x))Df(x)$ und somit folgt die Behauptung $Df^{-1}\big(f(x)\big) = \big(Df(x)\big)^{-1}$.

Sei $x^* \in V$ und $y^* = f(x^*)$. Wir müssen nur noch zeigen, daß f^{-1} an der Stelle y^* differenzierbar ist. Sei für $y \in f(V)$ der Vektor $s(y)$ definiert durch

$$x - x^* \;=\; f^{-1}(y) - f^{-1}(y^*) \;=\; \big(Df(x^*)\big)^{-1}(y - y^*) + s(y). \tag{1.1}$$

Wir wollen zeigen, daß es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so daß $\|s(y)\| \leq \varepsilon \|y - y^*\|$ für alle $y \in K_\delta(y^*)$ gilt. Dies sieht man folgendermaßen.

Zunächst hat die lineare Abbildung $\big(Df(x^*)\big)^{-1}$ endliche Operatornorm, d.h. es gibt eine Zahl $C < \infty$, so daß

$$\left\| \big(Df(x^*)\big)^{-1} z \right\| \;\leq\; C \,\|z\| \qquad \text{für alle } z \in \mathbb{R}^n \tag{1.2}$$

gilt. Sei $r(x)$ das durch

$$f(x) - f(x^*) \;=\; Df(x^*)(x - x^*) + r(x) \tag{1.3}$$

definierte Fehlerglied. Es gibt ein $\eta > 0$ so daß für alle $x \in K_\eta(x^*)$ gilt

$$\|r(x)\| \leq \min\Big(\frac{1}{2C}, \frac{\varepsilon}{2C^2}\Big)\|x - x^*\|.$$

Wähle $\delta > 0$ nach Satz 1, so daß $x = f^{-1}(y) \in K_\eta(x^*)$ ist für alle $y \in K_\delta(y^*)$. Wenn man auf (1.3) $Df^{-1}(x^*)$ anwendet und mit (1.1) vergleicht, ergibt sich $s(y) = Df^{-1}(x^*)r(x)$ und damit

$$\|s(y)\| \leq C\|r(x)\| \leq C\frac{1}{2C}\|x - x^*\| = \frac{1}{2}\|x - x^*\|.$$

Mit (1.1) folgt n

$$\|x - x^*\| \leq C\|y - y^*\| + \frac{1}{2}\|x - x^*\|$$

oder $\|x - x^*\| \leq 2C\|y - y^*\|$ und schließlich

$$\|s(y)\| \leq C\|r(x)\| \leq C \cdot \frac{\varepsilon}{2C^2}\|x - x^*\| \leq \varepsilon\|y - y^*\|$$

für alle y. ■

1.2 Das Newtonverfahren

Bisher haben wir gezeigt, daß unter den obigen Voraussetzungen zu der Funktion f eine Umkehrfunktion f^{-1} existiert, und wir kennen sogar die Ableitung der Umkehrfunktion. Noch offen ist allerdings folgende Frage: Wie kann man bei gegebenem y^* den Wert x^* bestimmen? Da sich die Gleichung $y = f(x)$ im allgemeinen nicht explizit nach x auflösen läßt, lösen wir dieses Problem durch ein Approximationsverfahren, das sogenannte **Newton-Verfahren**. Dieses wollen wir im Folgenden kurz darstellen.

Das Verfahren beruht auf folgender Idee. Sei x ein Punkt in der Nähe von x^*. Dann ist, wie in Gleichung (1.3) oben, $Df(x^*)(x - x^*) \approx y - y^*$. Diese Approximationsgleichung kann man explizit auflösen und mit $Df(x^*) \approx Df(x)$ erhält man

$$x^* \;\approx\; x - \big(Df(x)\big)^{-1}(y - y^*).$$

Satz 3 *(Newton-Verfahren) Sei $U \subseteq \mathbb{R}^n$ offen, $f\colon U \to \mathbb{R}^n$ stetig differenzierbar, $x^* \in U$ so, daß $Df(x^*) \neq 0$ ist, und $y^* = f(x^*)$. Für $x \in U$ definiere rekursiv eine Folge $(x_n)_{n\in\mathbb{N}}$ mit Startwert x durch die Vorschrift*

$$\begin{aligned} x_0 &= x \\ x_{n+1} &= x_n + \big(Df(x_n)\big)^{-1}(y^* - f(x_n)). \end{aligned}$$

Dann existiert eine Zahl $\eta > 0$, so daß diese Folge für alle Startwerte $x \in K_\eta(x^*)$ gegen den Wert x^* konvergiert.

Beweis. Wir beginnen mit einer Vorbemerkung. Sei E die Menge aller invertierbaren $n \times n$-Matrizen, also $E = \{ A \in \mathbb{R}^{n\times n} \mid \det A \neq 0 \}$. Dann ist $A \mapsto A^{-1}$ eine stetige Abbildung von E nach E, da die Cramersche Regel die Einträge von A^{-1} in stetiger Abhängigkeit der Einträge von A berechnet. Daher ist der Abstand $\|Df(x)^{-1} - Df(x^*)^{-1}\|$ klein, falls y nahe an y^* und damit x nahe an x^* liegt. Anders ausgedrückt bedeutet dies $\lim_{x\to x^*} \|Df(x)^{-1} - Df(x^*)^{-1}\| = 0$.

Mit $y_n = f(x_n)$ erhält man

$$x_{n+1} - x^* \;=\; \big(Df(x_n)\big)^{-1}(y^* - y_n) - (x^* - x_n)$$

und damit

$$\begin{aligned}
\|x_{n+1} - x^*\| \;&\leq\; \Big\|\big(Df(x_n)\big)^{-1}(y^* - y_n) - \big(Df(x^*)\big)^{-1}(y^* - y_n)\Big\| \\
&\qquad + \Big\|\big(Df(x^*)\big)^{-1}(y^* - y_n) - (x^* - x_n)\Big\| \\
&=\; \|r_1^n\| + \|r_2^n\|.
\end{aligned}$$

Wir wollen zeigen, daß die rechte Seite gegen Null konvergiert und schätzen dazu die einzelnen Fehlerterme r_1^n und r_2^n ab.

Sei $\varepsilon > 0$. Falls $\|x_n - x^*\|$ so klein ist, daß $\|Df(x)^{-1} - Df(x^*)^{-1}\| \leq \varepsilon$ ist, so gilt zunächst

$$\|r_1^n\| \;=\; \Big\|\big([Df(x_n)]^{-1} - [Df(x^*)]^{-1}\big)(y^* - y_n)\Big\| \;\leq\; \varepsilon\|y^* - y_n\|$$

Weiter gibt es dann eine Konstante C, so daß auch $\|r_1^n\| \leq C\varepsilon\|x^* - x_n\|$ gilt.

Nach den obigen Sätzen ist f^{-1} an der Stelle y^* differenzierbar mit Ableitung $\big(Df(x^*)\big)^{-1}$. Der zweite Fehlerterm r_2^n ist daher gerade der Fehler der Taylorapproximation erster Ordnung um die Stelle y^*. Für hinreichend kleines $\|x^* - x_n\|$ ist $\|y^* - y_n\| \leq C\|x^* - x_n\|$ ebenfalls klein. Daher kann man auch

$$\|r_2^n\| \;\leq\; \varepsilon\|y^* - y_n\| \;\leq\; \varepsilon C\|x^* - x_n\|$$

erreichen.

Nun wählen wir ε so klein, daß $L = 2\varepsilon C < 1$ gilt, und $\eta > 0$ so, daß die obigen Abschätzungen für $x_0 \in K_\eta(x^*)$ richtig sind. Durch Induktion folgt $x_n \in K_\eta(x^*)$ für alle n, d.h. für $\gamma_n = \|x_n - x^*\|$ gilt die Abschätzung $\gamma_{n+1} \leq L\|x^* - x_n\| = L\gamma_n$ für alle n. Daher konvergiert $\gamma_n \to 0$ für $n \to \infty$. ■

Die Newtonregel ist für eine Vielzahl konkreter Rechnungen nützlich. Im Folgenden werden wir z.B. näherungsweise den Wert $\sqrt{5}$ bestimmen. Dazu setzen wir $f(x) = x^2$ und $y^* = 5$. Die Iterationsvorschrift wird hier wegen $f'(x) = 2x$ zu

$$x_{n+1} \;=\; x_n + \frac{5 - x_n^2}{2x_n} \;=\; \frac{x_n}{2} + \frac{5}{2x_n}.$$

Mit $x_0 = 1$ erhält man die Werte

$$\begin{aligned}
x_1 \;&=\; \frac{1}{2} + \frac{5}{2} = 3 \\
x_2 \;&=\; \frac{3}{2} + \frac{5}{6} = \frac{14}{6} = \frac{7}{3} \approx 2,33
\end{aligned}$$

Der wahre Wert ist $\sqrt{5} \approx 2.24$.

Warnung. Wenn der Startwert nicht nahe bei x^* ist, kann das Verfahren divergieren. Dies wird durch folgendes Beispiel illustriert: $f(x) = \begin{cases} 1 - e^{-x} & x \geq 0 \\ e^x - 1 & x \leq 0 \end{cases}$
Dann ist $f'(x) = e^{-|x|}$ stetig, also f stetig differenzierbar. Gesucht x^* mit $f(x^*) = 0 = y^*$, also kennen wir die Lösung $x^* = 0$. Das Newton-Verfahren liefert die Iterationsgleichung

$$x_{n+1} - x_n = -\frac{f(x_n)}{f'(x_n)} = e^{|x_n|} - 1$$

und diese Differenz ist $\geq 3|x_n|$, falls $|x_n|$ groß genug ist z.B. $|x_n| \geq 2$. Damit wird $sgn(x_{n+1} - x_n) = -sgn(x_n)$ und $|x_n| \to \infty$.

1.3 Implizite Funktionen

Nachdem wir in der ersten Hälfte des Kapitels das Problem behandelt haben, Gleichungen der Form $y = f(x)$ nach x aufzulösen, kommen wir nun auf das eingangs geschilderte Problem zurück. Wir wollen also Gleichungen der Form $F(x, z) = 0$ nach z auflösen, indem wir Funktionen $g\colon x \mapsto g(x)$ finden, so daß $F\big(x, g(x)\big) = 0$ ist.

Bei der Funktion $F(x, z) = x^2 + z^2 - 1$, beispielsweise, wird die Gleichung $F(x, z) = 0$ genau durch die Punkte (x, z), die auf dem Einheitskreis liegen, gelöst. Wenn wir z als Funktion von x ausdrücken wollen, so müssen wir drei Fälle unterscheiden:

a) Für $|x| > 1$ gibt es kein passendes z.

b) Für $|x| < 1$ ist die Lösung nicht eindeutig bestimmt: mit z ist zugleich auch $-z$ Lösung. Die Lösung ist aber lokal eindeutig, denn wenn g stetig ist, und $F(x, g(x)) = 0$ in einer Umgebung von $x = x^*$ gilt, so ist die Lösung in der Nähe des Punktes $(x^*, f(x^*)) \in \mathbb{R}^2$ eindeutig. Man wählt einfach den oberen oder unteren Halbkreisbogen.

c) Für $|x| = 1$ sind die Lösungen an der Stelle x zwar eindeutig, aber in jeder Umgebung von x mehrdeutig.

Wir werden folgendes Ergebnis beweisen: g ist eindeutig und eine differenzierbare Funktion, falls

$$\frac{\partial F}{\partial z}(x^*, z^*) \neq 0$$

gilt. Im Beispiel ist $\frac{\partial F}{\partial z}(x^*, z^*) = 2z^* \neq 0$ außer für die Punkte $(1, 0)$ und $(-1, 0)$. In der Umgebung dieser Ausnahmepunkte ist die Lösung nicht eindeutig.

Nun zum analogen Problem in höheren Dimensionen: Sei im folgenden $F : U \to \mathbb{R}^m$ stetig differenzierbar wobei $U \subseteq \mathbb{R}^n \times \mathbb{R}^m$ offen ist. Sei $(x^*, z^*) \in U$ ein Punkt mit $F(x^*, z^*) = 0$. Wir wollen die Gleichung $F(x, z) = 0$ in einer Umgebung von (x^*, z^*) nach z auflösen. Beachte, daß die Gleichung sich auch in Form des Gleichungssystems

$$F_i(x_1, \ldots, x_n,\ z_1, \ldots, z_m) = 0 \quad \text{für } i = 1, \ldots, m$$

schreiben läßt, wobei die Anzahl m der Unbekannten z_j gleich der Anzahl der Gleichungen ist.
Die Jacobi-Matrix von F ist eine $m \times (n+m)$-Matrix:

$$DF = \left(\begin{array}{ccc:ccc} \frac{\partial F_1}{\partial x_1} & \cdots & \frac{\partial F_1}{\partial x_n} & \frac{\partial F_1}{\partial z_1} & \cdots & \frac{\partial F_1}{\partial z_m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial F_m}{\partial x_1} & \cdots & \frac{\partial F_m}{\partial x_n} & \frac{\partial F_m}{\partial z_1} & \cdots & \frac{\partial F_m}{\partial z_m} \end{array} \right) = (\underset{m\times n}{D_x F} \vdots \underset{m\times m}{D_z F})$$

Die oben im eindimensionalen erwähnte Bedingung, daß die partielle Ableitung von F nach z nicht verschwindet, wird jetzt durch eine entsprechende Bedingung an die Determinante der quadratischen Matrix $D_z F$ ersetzt. Der folgende Satz ist das Herzstück dieses Kapitels.

Satz 4 (implizite Funktionen) *Sei $U \subseteq \mathbb{R}^n \times \mathbb{R}^m$ offen und $F : U \to \mathbb{R}^m$ stetig differenzierbar. Sei $(x^*, z^*) \in U$, so daß $F(x^*, z^*) = 0$. Wenn*

$$\det D_z F(x^*, z^*) \neq 0$$

ist, dann gibt es Umgebungen V von x^* und W von z^*, so daß zu jedem $x \in V$ genau eine Lösung $z \in W$ von $F(x, z) = 0$ existiert und so daß die zugehörige Funktion $g : x \longmapsto z$ $g : V \longmapsto W$ stetig differenzierbar ist mit

$$Dg(x^*) = -D_z F(x^*, z^*)^{-1} D_x F(x^*, z^*).$$

Bemerkungen 1. Die letzte Gleichung ist also eine Gleichung zwischen $m \times n$-Matrizen. Im eindimensionalen Fall ist folgt sie daraus, daß wegen der Kettenregel und $F(x, g(x)) = 0$ gilt

$$0 = \frac{d}{dx} F(x, g(x)) = \frac{\partial F}{\partial x}(x, g(x)) + \frac{\partial F}{\partial z}(x, g(x)) g'(x).$$

Im Beispiel ist dann $g'(x^*) = -\frac{\partial F}{\partial x}(x^*, z^*) / \frac{\partial F}{\partial z}(x^*, z^*) = \frac{-2x^*}{2z^*} = \frac{-x^*}{z^*}$.
2. Der Vektor x kann als Parameter aufgefaßt werden, der die genaue Gestalt der Gleichung $F(x, z) = 0$ festlegt. Der Satz sagt also: Wenn die Gleichung für ein spezielles x^* eine Lösung z^* hat und die Determinantenbedingung erfüllt ist, dann lassen sich auch die ähnlichen Gleichungen $F(x, z) = 0$ lösen, falls x nahe genug an x^* liegt. Außerdem gibt es in der Nähe von z^* gibt es keine andere Lösungen. Schließlich hängen die Lösungen differenzierbar von x ab.

Beweis. 1. Wir wollen den Satz über inverse Abbildungen anwenden. Definiere $f : U \longrightarrow \mathbb{R}^n \times \mathbb{R}^m : (x, z) \longmapsto (x, F(x, z))$. Die Jacobi-Matrix von f hat die Gestalt

$$Df(x, z) = \begin{pmatrix} E_n & 0 \\ D_x F & D_z F \end{pmatrix},$$

wobei E_n die n-dimensionale Einheitsmatrix ist. Also ist

$$\det Df(x^*, z^*) = \det(E_n) \cdot \det(D_z F(x^*, z^*)) = \det(D_z F(x^*, z^*)) \neq 0.$$

Nach dem Satz über inverse Funktionen gibt es Zahlen $\eta > 0$ und $\delta > 0$, so daß $f|_{K_\eta(x^*, z^*)}$ injektiv ist und $f^{-1} : K_\delta(f(x^*, z^*)) \longrightarrow K_\eta(x^*, z^*)$ wohldefiniert

und stetig differenzierbar ist. Wähle Umgebungen $V \subset K_\delta(x^*)$ von x^* und W von z^*, so daß $V \times W \subset K_\eta(x^*, z^*)$.
2. Für jedes $x \in V$ existiert genau ein $z \in W$ so daß $F(x, z) = 0$. Betrachte hierzu den Vektor $y = (x, 0) \in \mathbb{R}^{n+m}$. Dann ist $\delta > \|x - x^*\| = \|y - (x^*, 0)\| = \|y - (x^*, F(x^*, z^*))\| = \|y - f(x^*, z^*)\|$. Also ist $y \in K_\delta(f(x^*, z^*))$. Der Urbildbektor $f^{-1}(y) \in U$ hat die Form $(\tilde{x}, \tilde{z})$. Es gilt nach Definition von f dann $(x, 0) = y = f(\tilde{x}, \tilde{z}) = (\tilde{x}, F(\tilde{x}, \tilde{z}))$. Wir erhalten durch Komponentenvergleich $\tilde{x} = x$ und $F(x, \tilde{z}) = 0$, dh. $\tilde{z}$ ist eine Lösung in W. Sei jetzt $\hat{z}$ eine andere Lösung mit $\hat{z} \in W$. Dann ist das Paar $(x, \hat{z})$ in $K_\eta(x^*, z^*)$ und es ist auch $f(x, \hat{z}) = (x, F(x, \hat{z})) = (x, 0) = y$. Da f auf $K_\eta(x^*, z^*)$ injektiv ist, gilt $\hat{z} = \tilde{z}$.
3. Wir definieren jetzt die Funktion g, die dem Punkt $x \in V$ den eindeutig bestimmten Punkt $z \in W$ zuordnet mit $F(x, z) = 0$. Dann ist $g(x)$ gleich der zweiten Komponente von $f^{-1}(x, 0)$, also ist $g(x) = \mathrm{pr}_2 \circ f^{-1} \circ (id, 0)(x)$, insbesondere ist g stetig differenzierbar. Die Funktion $h : x \longmapsto F(x, g(x))$ verschwindet identisch auf V. Also ist

$$0 = Dh(x^*) = D_x F(x^*, z^*) + D_z F(x^*, z^*) Dg(x^*)$$

und somit

$$Dg(x^*) = -(D_z F(x^*, z^*))^{-1} D_x F(x^*, z^*).$$

■

Beispiel: Wir betrachten die Abhängigkeit von Nullstellen von Polynomen von den Koeffizienten des Polynoms. Sei $a_0^* + a_1^* z + \ldots + a_n^* z^n = p_{a^*}(z)$ gegeben und sei z^* eine einfache Nullstelle. Dann ist $p_{a^*}(z^*) = 0$ und $p'_{a^*}(z^*) \neq 0$. Die Funktion $F : (a_0, \ldots, a_n, z) \longmapsto \sum_{i=0}^n a_i z^i$ erfüllt daher $\frac{\partial F}{\partial z}(a^*, z^*) = p'_{a^*}(z^*) \neq 0$. Wir können den Satz anwenden: Es gibt also eine stetig differenzierbare Funktion $g : (a_0, \ldots, a_n) \longmapsto g(a) = z$ so daß $p_a(z) = 0$. Sie liefert in einer Umgebung von (a^*, z^*) die einzige Nullstelle.

1.4 Karten, Tangentialräume und Extrema mit Nebenbedingungen

Wir wollen nun noch als Anwendung drei Sätze beweisen, die alle die folgende Situation betreffen:

Generalvoraussetzung Sei $m < k$, $U \subseteq \mathbb{R}^k$ offen und $f_1, \ldots, f_m : U \longrightarrow \mathbb{R}$ stetig differenzierbare Funktionen. Sei M die gemeinsame Lösungsmenge der Gleichungen $f_i = 0,\ i = 1, \ldots, m$ dh.

$$M = \bigcap_{i=1}^m \{y \in U : \ f_i(y) = 0\}.$$

(Sehr viele Mengen lassen sich auf diese Weise darstellen.) Sei $a \in M$ so daß die Gradienten $\nabla f_i(a), \ldots, \ \nabla f_m(a)$ linear unabhängig sind.

Satz 5 (Existenz von „Karten".) *Unter der obigen Voraussetzung sei $n = k - m$. Es gibt es eine Umgebung $\tilde{V}$ von a, eine offene Menge V in $\mathbb{R}^n$ und eine stetig differenzierbare bijektive Funktion $\psi : V \longrightarrow \tilde{V} \cap M$. Dabei kann man erreichen, daß rg $D\psi(x^*) = n$ ist, wobei $x^* = \psi^{-1}(a)$.*

Beweis. Sei N der lineare Unterraum von $\mathbb{R}^k$, der von $\nabla f_1(a), \ldots, \nabla f_m(a)$ aufgespannt wird (Normalenraum). Da diese Vektoren linear unabhängig sind, ist $\dim N = m$.
1. Fall: Wir identifizieren $\mathbb{R}^k$ mit dem Produktraum $\mathbb{R}^n \times \mathbb{R}^m$ und nehmen zusätzlich an, daß N die spezielle Gestalt

$$N = \{(0, \ldots, 0,\ z_1, \ldots, z_m) \in \mathbb{R}^n \times \mathbb{R}^m\}$$

hat. Dann definieren wir die Funktion $F : U \longrightarrow \mathbb{R}^m$ durch

$$F(x, z) = (f_1(x, z), \ldots, f_m(x, z)).$$

Damit wird $M = \{(x, z) \in U : F(x, z) = 0\}$. Mit (x^*, z^*) bezeichnen wir die Komponenten des Vektors a. Die partielle Ableitungen von $F_i = f_i$ nach den z_j ergeben sich durch

$$\frac{\partial F_i}{\partial z_j} = \langle \nabla F_i(a), e_{n+j} \rangle = \langle \nabla f_i(a), e_{n+j} \rangle.$$

Da nach Voraussetzung die ersten n Komponenten der Gradienten $\nabla f_i(a)$ verschwinden, ist

$$\nabla f_i(a) = (0, \ldots, 0,\ \frac{\partial F_i}{\partial z_1}(a), \ldots, \frac{\partial F_i}{\partial z_m}(a))$$

und

$$D_z F(a) = \left((\nabla f_i(a))_{n+j} \right)_{\substack{i=1,\ldots,m \\ j=1,\ldots,m}}.$$

Die Zeilen dieser Matrix sind nach Voraussetzung linear unabhängig. Also ist der Satz über implizite Funktionen anwendbar. Es gibt damit Umgebungen V von x^* und W von z^* und ein stetig differenzierbares $g \colon V \longrightarrow W$ derart daß für alle $x \in V$ der Wert $z = g(x)$ die einzige Lösung von $F(x, z) = 0$ in W ist. Sei $\tilde{V} = \{(x, z) : x \in V, z \in W\}$. Dann ist $\tilde{V}$ offen und ein Punkt $(x, z) \in \tilde{V}$ ist genau dann in M, wenn $F(x, z) = 0$ ist, dh. wenn (x, z) von der Form $(x, g(x))$ ist. Also ist

$$\tilde{V} \cap M = \{(x, g(x)) : x \in V\} = \psi(V),$$

wobei $\psi(x) = (x, g(x))$. Damit ist $\psi : V \longrightarrow \tilde{V} \cap M$ bijektiv und stetig differenzierbar.

Die Funktionalmatrix von ψ hat n Spalten und als ersten Teil die Einheitsmatrix in n Dimensionen, also hat sie den Rang n.

2. Den allgemeinen Fall führen wir mittels Koordinatentransformation auf Fall 1 zurück.

Wähle eine ON-Basis $(b_1, \ldots, b_n)$ von $N^\perp$. Wähle eine ON-Basis $(b_{n+1}, \ldots, b_{n+m})$ von N und betrachte den Vektorraum-Isomorphismus $B : \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}^k$ mit $B(e^i) = b_i$ für $1 \leq i \leq k$. Dann ist $N^\perp$ der Raum der Vektoren der Form $B(x, 0)$ und N der Raum der Vektoren $B(0, z)$. Wir ersetzen jetzt U durch $\tilde{U} = B^{-1}(U)$, und die Funktionen f_i durch $\tilde{f}_i = f_i \circ B$. Dann ist $\tilde{M} = B^{-1}(M)$ die Lösungsmenge der Gleichungen $\tilde{f}_i = 0$. Ferner ist $\nabla \tilde{f}_i(y) = \nabla f_i(By) \circ B$ nach der Kettenregel, also

$$\begin{aligned} \nabla \tilde{f}_i(x, z) \cdot e^j &= \nabla (f_i \circ B)(x, z) \cdot e^j \;=\; \nabla f_i(B(x, z)) \cdot B e^j \\ &= \nabla f_i\big(B(x, z)\big) \cdot b_j. \end{aligned}$$

Sei nun (x^*, z^*) so daß $B(x^*, z^*) = a$. Dann verschwindet an dieser Stelle für $j \leq n$ dieses Skalarprodukt wegen $b_j \in N^\perp$. Also ist $\nabla \tilde{f}_i(x^*, z^*)$ ein Vektor der Form $\{(0, \ldots, 0, z_1, \ldots, z_m)\}$. Ferner ergibt sich, daß die Funktionalmatrix $(\frac{\partial \tilde{f}_i}{\partial z_j}(x^*, z^*))$ gerade die Koeffizientenmatrix der linear unabhängigen Vektoren $\nabla f_1(a), \ldots, \nabla f_m(a))$ bezüglich der Basis $(b_{n+1}, \ldots, b_{n+m})$ von N ist. Damit ist die Matrix nicht singulär und die Gradienten $\nabla \tilde{f}_i(x^*, z^*)$ sind linear unabhängig. Also sind alle Voraussetzungen aus Fall 1 für die $\tilde{f}_i$ erfüllt.
Nach Fall 1 gibt es eine stetig differenzierbare, bijektive Abbildung $\psi_0 : V \longrightarrow \tilde{M} \cap \tilde{V}$. Die Abbildung $\psi = B \circ \psi_0 : V \longrightarrow B(\tilde{M} \cap \tilde{V})$ ist dann die gesuchte Funktion. Die Aussage über den Rang von $D\psi$ ergibt sich mit der Kettenregel, da Multiplikation mit der zu B gehörenden Matrix den Rang einer Matrix erhält. ∎

Satz 6 (Charakterisierung des Tangentialraums) *Sei N der von den Vektoren $\nabla f_1(a), \ldots, \nabla f_m(a)$ aufgespannte Raum. Ein Vektor v ist genau dann in $N^\perp$, wenn es eine Kurve $\gamma : (-\varepsilon, \varepsilon) \longrightarrow M$ gibt mit $\gamma(0) = a$ und $\gamma'(0) = v$.*

Beweis. 1. Sei $v = \gamma'(0)$ für eine stetig differenzierbare Kurve $\gamma : (-\varepsilon, \varepsilon) \to M$ mit $\gamma(0) = a$. Dann gilt $f_i(\gamma(t)) = 0$ für alle t, also $0 = (f_i \circ \gamma)'(0) = \langle \nabla f_i(\gamma(0)), \gamma'(0) \rangle = \langle \nabla f_i(a),\ v \rangle$. Daher ist $v \in N^\perp$.

2. Sei $v \in N^\perp$. Zu zeigen ist die Existenz einer Kurve γ mit $\gamma(0) = a,\ \gamma'(0) = v,\ \gamma : (-\varepsilon, \varepsilon) \longrightarrow M$. Sei ψ wie im vorigen Satz. Dann ist $f_i \circ \psi(x) = 0$ für alle $x \in V$. Also ist $\nabla(f_i \circ \psi) = 0$ und $\langle \nabla f_i(a),\ D\psi(x^*)w \rangle = 0$ für alle $w \in \mathbb{R}^n$. Also $D\psi(x^*)\ w \in N^\perp$ für alle $w \in \mathbb{R}^n$. Wegen $\dim N^\perp = n$ und Rang $D\psi(x^*) = n$ ist also $D\psi(x^*) : \mathbb{R}^n \longrightarrow N^\perp$ surjektiv, dh. es existiert zu $v \in N^\perp$ ein Vektor $w \in \mathbb{R}^n$ so daß $D\psi(x^*)w = v$. Die Kurve $\gamma(t) := \psi(x^* + t \cdot w)$ erfüllt also $\gamma'(0) = D\psi(x^*) \cdot w = v$. ∎

Definition 1 *Der lineare Raum*

$$T_aM = \{\gamma'(0) \mid \gamma \colon (-\varepsilon, \varepsilon) \longrightarrow M \text{ stetig differenzierbar mit } \gamma(0) = a\}$$

heißt **Tangentialraum** *oder* **Raum der Tangentialvektoren** *von M im Punkt a. Ein Vektor aus N der Länge 1 heißt auch* **Normale** *oder* **Normalenvektor** *an M in a.*

Schließlich erhalten wir eine einfache Begründung der Technik der „Lagrange-Mulitplikatoren".

Satz 7 (Lokale Extrema unter Nebenbedingungen) *Sei $h : U \longrightarrow \mathbb{R}$ eine weitere stetig differenzierbare Funktion, die im Punkt $a \in M$ ein lokales Extremum unter den Nebenbedingungen $f_i = 0$ hat (dh. es gibt eine Umgebung W von a so, daß $h(a) \geq h(y)$ für alle $y \in M \cap W$ oder $h(a) \leq h(y)$ für alle $y \in M \cap W$). Dann existieren Zahlen $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$ („***Lagrange-Multiplikatoren***") so daß $\nabla h(a) = \sum_{i=1}^m \lambda_i \nabla f_i(a)$.*

Beweis. Sei $v \in N^\perp$ beliebig. Wähle eine Kurve γ nach dem letzten Satz mit $\gamma(0) = a,\ \gamma'(0) = v$, Dann hat $h \circ \gamma$ ein lokales Extremum für $t = 0$ weil $\gamma(t) \in M$ ist für alle t. Daraus folgt $0 = (h \circ \gamma)'(0) = \langle \nabla h(a), \gamma'(0) \rangle = \langle \nabla h(a), v \rangle$. Dies gilt für alle $v \in N^\perp$. Daher ist $\nabla h(a) \in N^{\perp\perp} = N$, dh. $\nabla h(a)$ ist eine Linearkombination von den $\nabla f_i(a)$. ∎

Kapitel 2

Integration stetiger Funktionen mit kompaktem Träger

In weiteren Verlauf dieses Skripts haben wir folgende Ziele:

a) Definiere $\int_{\mathbb{R}^n} f(x)dx$ für möglichst viele Funktionen $f : \mathbb{R}^n \longrightarrow \mathbb{R}$

b) Definiere $\mathrm{vol}_m(A)$ für möglichst viele Mengen A

c) Berechnungsregeln, Beziehung zur Differentialrechnung

Die Darstellung ist so angelegt, daß Verallgemeinerungen (Ausdehnung auf andere Grundräume und andere Inhaltsbegriffe („Maße“)) leicht möglich werden.

2.1 Ein Satz über parameterabhängige Integrale

Der folgende Satz über eindimensionale Integrale benutzt die Tatsache, daß man bei gleichmäßiger Konvergenz Integral und Grenzwert vertauschen kann. Später werden wir den Satz von der majorisierten Konvergenz beweisen, der dann auch zu entsprechenden Verschärfungen dieses Abschnitts führt.

Satz 8 *Seien $I = [a,b]$ ein Intervall und $Q \subset \mathbb{R}^n$ eine kompakte Menge. Sei $f : I \times Q \longrightarrow \mathbb{R}$ stetig. Definiere für $t \in Q$*

$$F(t) := \int_a^b f(x,t)\ dx.$$

Dann gilt:

a) Die Funktion F ist stetig auf Q.

b) Wenn $Q = [c,d]$ auch ein Intervall ist und $\frac{\partial}{\partial t} f(x,t) : I \times I_2 \longrightarrow \mathbb{R}$ existiert und stetig ist, dann ist F stetig differenzierbar und es gilt

$$F'(t) = \int_a^b \frac{\partial}{\partial t} f(x,t)\ dx.$$

Beweis. a) Weil $I \times Q$ kompakt ist, ist f gleichmäßig stetig, d.h. zu jedem $\varepsilon > 0$ existiert ein $\delta > 0$, so daß

$$|f(x,t) - f(x^*,t')| < \varepsilon \quad \text{für} \quad \begin{array}{ll} \text{alle} & x \in I,\ t, t' \in Q \\ \text{mit} & \|(x,t) - (x,t')\| > \delta \end{array}$$

Sei jetzt (t_n) eine Folge mit $t_n \longrightarrow t$. Setze

$$\varphi_n(x) := f(x, t_n), \ \ \varphi(x) = f(x,t).$$

Dann gilt für $|t_n - t| < \delta$ auch $\sup_{x \in I} |\varphi_n(x) - \varphi(x)| \leq \varepsilon$. Also konvergiert die Funktionsfolge (φ_n) gleichmäßig gegen φ. Es folgt

$$F(t_n) = \int_a^b f(x,t_n)dx = \int_a^b \varphi_n(x)dx \longrightarrow \int_a^b \varphi(x)dx = F(t).$$

b) Wähle analog zu $\varepsilon > 0$ ein $\delta > 0$ mit

$$|\frac{\partial}{\partial t} f(x,t) - \frac{\partial}{\partial t} f(x,t')| < \varepsilon \quad \text{für alle } x, t, t' \text{ mit } \|(x,t) - (x,t')\| < \delta.$$

Dies ist möglich, weil $\frac{\partial}{\partial t} f$ auf $I \times Q$ gleichmäßig stetig ist. Sei wieder (t_n) eine Folge mit $t_n \longrightarrow t$. Sei $|t_n - t| < \delta$ für alle $n \geq N$. Wir wollen zeigen, daß $\Delta_n F = \frac{F(t_n) - F(t)}{t_n - t}$ konvergiert. Es ist

$$\begin{aligned} \Delta_n F &= \frac{1}{t_n - t} \left(\int_a^b f(x,t_n)dx - \int_a^b f(x,t)dx \right) \\ &= \int_a^b \frac{f(x,t_n) - f(x,t)}{t_n - t} dx \\ &= \int_a^b \frac{\partial}{\partial t} f(x, t_n^*(x))dx \end{aligned}$$

nach dem Mittelwertsatz, wobei $|t_n^*(x) - t| \leq |t_n - t| < \delta$, also

$$\begin{aligned} |\Delta_n F &- \int_a^b \frac{\partial}{\partial t} f(x,t)dx| \\ &\leq \int_a^b |\frac{\partial}{\partial t} f(x, t_n^*(x) - \frac{\partial}{\partial t} f(x,t)|dx \\ &< \varepsilon(b-a) \end{aligned}$$

für alle $n \geq N$ und damit

$$\lim_{n \to \infty} \Delta_n F = \int_a^b \frac{\partial}{\partial t} f(x,t)dt.$$

■

2.2 Mehrfache Integrale für stetige Funktionen

Seien $I_1, \ldots, I_m$ Intervalle, $I_j = [a_j, b_j]$, $f : \mathbb{R}^m \longrightarrow \mathbb{R}$ stetig.

Dann definieren wir

$$\int_{a_m}^{b_m} \ldots \int_{a_1}^{b_1} f(x_1, \ldots, x_m)\, dx_1 \ldots dx_m \tag{2.1}$$

rekursiv durch sukzessives Integrieren über die einzelnen Variablen.

Ausführlicher: Bei festgehaltenem $(x_2, \ldots, x_m) \in I_2 \times \ldots \times I_m$ betrachten wir das eindimensionale Integral bzgl. x_1 :

$$F_1(x_2, \ldots, x_m) = \int_{a_1}^{b_1} f(x_1, x_2, \ldots, x_m)\, dx_1.$$

Die neue Funktion F_1 ist nach dem Satz über parameterabhängige Integrale stetig in $(x_2, \cdots, x_m)$. Also kann man

$$F_2(x_3, \ldots, x_m) = \int_{a_2}^{b_2} F_1(x_2, x_3, \ldots, x_m) dx_2$$

für alle $(x_3, \ldots, x_m) \in I_3 \times \ldots \times I_m$ berechnen. So fährt man fort, bis man durch Integrieren bzgl. x_m die Zahl (1) erhält.

Eine Menge der Form $Q = I_1 \times \ldots \times I_m$ heißt achsenparalleler kompakter Quader. Man schreibt auch einfach

$$\int_Q f(x_1, \ldots, x_m)\, dx_1 \ldots dx_m$$

oder noch kürzer $\int_Q f(x) d^n x$ bzw. $\int_Q f(x) dx$.

Satz 9 *Es kommt nicht auf die Reihenfolge der Integrationen an, d.h. für $m = 2$ gilt*

$$\int_c^d \int_a^b f(x,y) dx\ dy = \int_a^b \int_c^d f(x,y)\ dy\ dx$$

und allgemein für jede Permutation $\sigma \in S_m$

$$\int_{I_m} \ldots \int_{I_1} f\ dx_1 \ldots dx_m = \int_{I_{\sigma(m)}} \ldots \int_{I_{\sigma(1)}} f\ dx_{\sigma(1)} \ldots dx_{\sigma(m)}$$

Beweis. Für $m = 2$: Setze $\varphi(x,t) = \int_c^t f(x,y) dy$.
Dann ist $\varphi(x,c) = 0$ und

$$\frac{\partial \varphi}{\partial t}(x,t) = f(x,t).$$

Sei weiter

$$\Phi(t) = \int_a^b \varphi(x,t) dx.$$

Dann gilt

$$\Phi'(t) = \int_a^b \frac{\partial\varphi}{\partial t}(x,t)dx = \int_a^b f(x,t)dt$$

und daher

$$\begin{aligned}\int_a^b \int_c^d f(x,y)\,dy\,dx &= \int_a^b \varphi(x,d)\ dx \ = \ \Phi(d) \ = \ \Phi(d) - \Phi(c) \\ &= \int_c^d \Phi'(t)\,dt \ = \ \int_c^d \int_a^b f(x,t)\,dx\,dt.\end{aligned}$$

Den allgemeinen Fall erhält man, indem zunächst benachbarte Komponenten vertauscht und dann σ als sukzessive Vertauschung geeigneter benachbarter Paare darstellt. (Dies ist immer möglich.) ∎

Der **Träger** einer reellwertigen Funktion auf einem metrischen Raum ist die kleinste abgeschlossene Menge, außerhalb derer die Funktion verschwindet:

$$\text{supp } f = \overline{\{x | f(x) \neq 0\}}.$$

Wenn eine Funktion auf dem $\mathbb{R}^m$ außerhalb einer beschränkten Menge verschwindet, ist ihr Träger also kompakt. Sei

$$C_c(\mathbb{R}^m) := \{f \in C(\mathbb{R}^m) \mid \text{supp}\, f \,\text{kompakt}\}.$$

Für $f \in C_c(\mathbb{R}^m)$ sei $Q = I_1 \times \ldots \times I_m$ ein Quader mit supp $f \subseteq Q$. Wenn dann $\tilde{Q} = \tilde{I}_1 \times \ldots \times \tilde{I}_m$ ein größerer Quader ist (d.h. $I_k \subseteq \tilde{I}_k$), so ist

$$\int_Q f(x)d^n x = \int_{\tilde{Q}} f(x)d^n x$$

Denn bei eindimensionalen Integralen kann man die Teilintervalle, auf denen der Integrand verschwindet, beim Integrationsbereich weglassen. Aus der sukzessiven Definition des Integrals folgt dann auch die analoge Aussage für Quader.

Also definieren wir einfach

$$\int_{\mathbb{R}^m} f(x)dx := \int_Q f(x)dx,$$

wobei Q irgendein Quader ist mit supp $f \subset Q$ und die spezielle Wahl von Q keine Rolle spielt.

Spezialfall: Funktionen von Produktform. Sei

$$f(x_1, \ldots, x_m) \ = \ f_1(x_1) \cdots f_m(x_m) \quad \text{mit} \quad f_k \in C_c(\mathbb{R}).$$

Dann gilt

$$\int_{\mathbb{R}^n} f\ dx = \Big(\int_{\mathbb{R}} f_1(x_1)dx_1\Big) \cdots \Big(\int_{\mathbb{R}} f_m(x_m)dx_m\Big).$$

Beweis. Für $m = 2$: Weil man Konstanten aus eindimensionalen Integralen herausziehen kann, gilt

$$\begin{aligned}\int_{\mathbb{R}} \int_{\mathbb{R}} f(x,y)\ dx\ dy &= \int_{\mathbb{R}} f_2(y) \int_{\mathbb{R}} f_1(x)\ dx\ dy \\ &= (\int_{\mathbb{R}} f_1(x)\ dx)\ (\int_{\mathbb{R}} f_2(y)\ dy).\end{aligned}$$

Allgemeiner Fall: sukzessives Wiederholen des Arguments. ■

Wir notieren die folgenden einfachen Eigenschaften des Integrals

Satz 10 *Für alle* $a, b \in \mathbb{R}, \quad f, g \in C_c(\mathbb{R})$ *gilt*

a) $\int_Q (af + bg)(x)\ dx = a \int_Q f(x)\ dx + b \int_Q g(x)\ dx$ *(Linearität)*

b) Aus $f(x) \le g(x)\, \forall\, x \in Q$ *folgt*

$$\int_Q f(x)\ dx \le \int_Q g(x)\ dx \qquad (\textit{Isotonie})$$

c) Insbesondere gilt

$$(\inf_{x\in Q} f(x)) \cdot |I_1| \cdot \ldots \cdot |I_n| \le \int_Q f(x)\ dx \le (\sup_{x\in Q} f(x))\ |I_1| \cdot \ldots \cdot |I_n|$$

Beweis. Dies folgt aus den entsprechenden Eigenschaften des eindimensionalen Integrals. ■

Satz 11 *(Dini). Sei* (f_n) *eine monotone Folge stetiger Funktionen auf einer kompakten Menge* K *in einem metrischen Raum mit* $f_n(x) \longrightarrow f(x) \quad \forall x \in K$. *Ist dann auch* $f : K \longrightarrow \mathbb{R}$ *stetig, so konvergiert* $f_n \longrightarrow f$ *gleichmäßig auf* K.

Beweis. 1. Fall: (f_n) monoton fallend. Es sei $\varepsilon > 0$. Betrachte die Menge

$$K_n = \{x | f_n(x) - f(x) \ge \varepsilon\}.$$

Diese Menge ist abgeschlossen und in supp $f \cup$ supp f_n enthalten. Also ist sie kompakt. Ferner gilt $K_{n+1} \subseteq K_n$, weil $f_{n+1}(x) \le f_n(x)$. Wegen $f_n(x) - f(x) \longrightarrow 0$ ist jedes x nur in endlich vielen K_n enthalten, d.h. es ist

$$\bigcap_{n=1}^{\infty} K_n = \emptyset.$$

Da die K_n kompakt sind gibt es sogar ein $n_0 \in \mathbb{N}$ mit

$$\bigcap_{n=1}^{n_0} K_n = K_{n_0} = \emptyset.$$

Daher gilt

$$f(x) \le f_n(x) \le f(x) * \varepsilon \qquad \forall\, n \ge n_0,$$

$$\text{d.h. } ||f_n - f||_\infty \le \varepsilon \qquad \text{für alle } n \ge n_0$$

2. Fall: analog (f_n) wachsend. ■

Folgerung 1 *Sei* Q *ein kompakter Quader. Seien* $f_n, f \in C_c(\mathbb{R}^m)$, *sei* (f_n) *monoton fallend und* $\lim_{n\to\infty} f_n(x) \le f(x)$ *für alle* $x \in Q$. *Dann gilt*

$$\lim_{n\to\infty} \int_Q f_n(x)\ dx \le \int_Q f(x)\ dx.$$

Beweis. Wir betrachten die Funktion $\tilde{f}_n = \max(f_n, f)$. Offenbar ist $\tilde{f}_n$ stetig und es gilt $\tilde{f}_n(x) \downarrow f(x)$ für alle $x \in Q$. Nach dem Satz von Dini folgt
$\sup_{x\in Q} \tilde{f}_n(x) - f(x) \underset{n\to\infty}{\longrightarrow} 0$ und daher nach dem vorletzten Satz

$$
\begin{aligned}
\lim_{n\to\infty} \int_Q f_n dx &\leq \lim_{n\to\infty} \int_Q \tilde{f}_n(x) dx \\
&= \int_Q f(x)\ dx + \lim_{n\to\infty} \int \tilde{f}_n - f\ dx \\
&\leq \int_Q f(x)\ dx + \lim_{n\to\infty} \left(\sup_x (f_n - f)(x) \right) (|I_1| \dots |I(n|) \\
&= \int_Q f(x)\ dx.
\end{aligned}
$$

■

2.3 Das Verhalten des Integrals bei affinen Transformationen

Satz 12 *Das n-dimensionale Integral is* **translationsinvariant***: Für jeden Vektor $a \in \mathbb{R}^m$ und jede Funktion $f \in C_c(\mathbb{R}^n)$ gilt*

$$\int_{\mathbb{R}^m} f(x+a)dx = \int_{\mathbb{R}^n} f(x)dx$$

Beweis. Betrachte die Funktion $\tau_a f : x \longmapsto f(x+a)$. Dann ist auch $\tau_a f$ stetig mit kompakten Träger (ist supp $f \subset Q$, so ist supp $\tau_a f \subset Q - a$). Die Behauptung folgt wieder aus der entsprechenden 1-dimensionalen Eigenschaft:
Wir führen den Beweis zunächst für $m = 2$.

$$
\begin{aligned}
&\int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x+a_1, y+a_2)dx \right) d\ dy \\
= &\int_{\mathbb{R}} \underbrace{\left(\int_{\mathbb{R}} f(x, y+a_2)dx \right)}_{F(y+a_2)} dy \\
= &\int_{\mathbb{R}} F(y+a_2)dy = \int_{\mathbb{R}} F(y)dy \int_{\mathbb{R}} \int_{\mathbb{R}} f(x,y)dxdy
\end{aligned}
$$

Allgemeine m analog durch sukzessives verschieben der Komponenten. ■

Satz 13 *Sei $A\colon \mathbb{R}^m \longrightarrow \mathbb{R}^m$ linear und nicht singulär. Dann gilt für $f \in C_c(\mathbb{R}^m)$ $\int_{\mathbb{R}^n} f(Ax)dx = \frac{1}{|\det A|} \int_{\mathbb{R}^n} f(x)dx$*

Beweis. 1. Wenn A nicht singulär ist, ist auch A^{-1} linear und stetig. Weil f außerhalb von supp f verschwindet, verschwindet $f \circ A$ außerhalb $A^{-1}\mathrm{supp} f$ (kompakt), also hat $f \circ A$ auch kompakten Träger.

2. Wenn die Behauptung für A und für B stimmt, dann auch für AB : Sei $g(x) = f(Ax)$. Dann ist

$$\begin{aligned}
\int_{\mathbb{R}^m} f(ABx)\,dx &= \int_{\mathbb{R}^m} g(Bx)\,dx \frac{1}{|\det B|} \int g(x)\,dx \\
&= \frac{1}{|\det B|} \int f(Ax)\,dx = \frac{1}{|\det A|} \cdot \frac{1}{|\det B|} \int f(x)\,dx \\
&= \frac{1}{|\det AB|} \int f(x)\,dx.
\end{aligned}$$

3. Wir bemerken, daß jede Matrix Produkt von Elementarmatrizen ist. Die den Elementarmatrizen entsprechenden Transformationen sind von folgender Art:

a) Zwei Komponenten werden vertauscht:

$$P(x_1, \dots, x_i, \dots, x_j, \dots, x_m) = (x_1, \dots, x_j, \dots, x_i, \dots, x_m).$$

b) Eine Komponente wird mit einer reellen Zahl multipliziert

$$A_{\lambda,k}(x_1, \dots, x_k, \dots, x_m) = (x_1, \dots, \lambda x_k, \dots, x_m)$$

c) Ein Vielfaches einer Komponente wird zu einer anderen addiert.

$$A_{k,l,\lambda}(x_1, \dots, x_k, \dots, x_m) = (x_1, \dots, x_k + \lambda x_l, \dots, x_m)$$

a) Für Vertauschungs-Abbildungen P gilt $\det P = -1$ und daher

$$\int_{\mathbb{R}^m} f(Px)dx = \int_{\mathbb{R}^m} f(x)dx = \frac{1}{|\det P|} \int_{\mathbb{R}^m} f(x)dx,$$

weil es nicht auf die Integrationsreihenfolge ankommt.
b) Im Fall b) können wir durch zusätzliche Komponentenvertauschungen auf den Fall beschränken, daß die erste Komponente multipliziert wird. Dann ist

$$\begin{aligned}
\int f(Ax)dx &= \int\int \left(\int f(\lambda x_1, x_2, \dots, x_m) dx_1 \right) dx_2 \dots dx_m \\
&= \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \left(\frac{1}{|\lambda|} \int_{\mathbb{R}} f(x_1, \dots, x_m dx_1 \right) \dots dx_m \\
&= \frac{1}{|\det A|} \int f(x)dx \text{ wegen } \det A = \lambda.
\end{aligned}$$

Im dritten Fall können wir durch Komponentenvertauschungen erreichen, daß $A(x_1, \dots, x_m) = (x_1 + \lambda x_2, x_2, \dots, x_m)$. Dann ist $\det A = 1$ und wegen der Translationsinvarianz des innersten Integrals

$$\begin{aligned}
\int f(Ax)dx &= \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x_1 + \lambda x_2, x_2, \dots, x_m) dx_1 \right) \dots dx_m \\
&= \int \dots \int \left(\int_{\mathbb{R}} f(x_1, x_2, \dots, x_m) dx_1 \right) \dots dx_m \\
&= \frac{1}{|\det A|} \int_{\mathbb{R}^n} f(x)dx
\end{aligned}$$

■

Folgerung 2 *Das Integral ist rotationsinvariant, d.h. für orthogonale Transformationen U gilt*

$$\int_{\mathbb{R}^m} f(Ux)dx = \int_{\mathbb{R}^m} f(x)dx.$$

Beweis. Aus $U^T = U^{-1}$ folgt $1 = \det U \det U^T = (\det U)^2$ und daher $|\det U| = 1$. ■

Kapitel 3

Das Volumen für kompakte Mengen

Wenn man das m-dimensionale Volumen von Teilmengen des $\mathbb{R}^m$ definieren will, soll natürlich ein kompakter Quader $Q = I_1 \times \cdots \times I_m$ das Volumen $|I_1| \cdots |I_m|$ erhalten. Man könnte dann, von den Quadern ausgehend, zunächst Vereinigungen von Quadern betrachten und dann zu komlizierteren Mengen übergehen. Die folgende Definition geht einen etwas anderen Weg. Sie führt das Volumen für kompakte Mengen direkt auf das Integral von stetigen Funktionen mit kompaktem Träger zurück. Wir können dann mit Hilfe des sogenannten Cavalieri-schen Prinzips das Volumen einfacher geometrischer Objekte ausrechnen, bevor wir die etwas kniffligere Definition des Volumens für allgemeinere Mengen im nächsten Paragraphen diskutieren.

Definition 2 *Sei K eine kompakte Teilmenge des $\mathbb{R}^m$. Dann definieren wir ihr Volumen $vol_m(K)$ durch*

$$vol_m(K) = \inf\left\{ \int_{\mathbb{R}^m} f(x)dx \ : \ f \in \mathcal{C}_c(\mathbb{R}^m), \mathbf{1}_K \leq f \right\}.$$

Bemerkungen. 1. Wir werden weiter unten sehen, daß dies speziell für Quader mit dem oben angegebenen Ausdruck übereinstimmt.
2. Die folgende Bemerkung ist nicht wesentlich für das Verständnis des Rest des Skripts. Die Beweise der im folgenden ersten allgemeinen Teil dieses Kapitels hergeleiteten Eigenschaften des Volumens benutzen vom Integral nur die (schon bewiesene) Eigenschaft, daß die Abbildung $I : \mathcal{C}_c(\mathbb{R}^m) \to \mathbb{R}, f \mapsto \int_{\mathbb{R}^m} f(x)\ dx$ linear und isoton ist. Im Anhang verwenden wir diese Tatsache für einen Beweis des sogenannten Rieszschen Darstellungssatzes.

3.1 Allgemeine Eigenschaften

Natürlich folgt aus der obigen Definition, daß für jede kompakte Menge eine Folge (f_n) in $\mathcal{C}_c(\mathbb{R}^m)$ existiert mit $\mathbf{1}_K \leq f_n$ für alle n, deren Integrale von oben gegen $vol_m(K)$ konvergieren. Wir zeigen jetzt mit Hilfe des Satzes von Dini, daß dies sogar für *jede* Folge (f_n), die monoton fallend gegen $\mathbf{1}_K$ konvergiert, gilt.

Wir schreiben $f_n \downarrow \mathbf{1}_K$, wenn $(f_n(x))$ für jedes x monoton fallend ist und für $x \in K$ gegen 1 und für $x \notin K$ gegen 0 konvergiert.

Satz 14 *Sei (f_n) eine Folge von Funktionen aus $\mathcal{C}_c(\mathbb{R}^m)$ mit $f_n \downarrow \mathbf{1}_K$. Dann gilt*

$$vol_m(K) = \lim_{n\to\infty} \int_{\mathbb{R}^m} f_n(x)\ dx.$$

Beweis. 1. Aus $f_n \geq \mathbf{1}_K$ folgt, daß das Integral der Funktion f_n als Element der geschweiften Klammer in der Definition vorkommt. Also ist $vol_m(K) \leq \lim_{n\to\infty} \int_{\mathbb{R}^m} f_n(x)\ dx$.
2. Umgekehrt sei $g \in \mathcal{C}_c(\mathbb{R}^m)$ irgendeine Funktion mit $g \geq \mathbf{1}_K$. Wir müssen zeigen, daß

$$\lim_{n\to\infty} \int_{\mathbb{R}^m} f_n(x)\ dx \leq \int_{\mathbb{R}^m} g(x)\ dx,$$

denn dann ist der Grenzwert auch kleiner oder gleich dem Infimum in der Definition von $vol_m(K)$ also auch kleiner oder gleich $vol_m(K)$ selber. Sei Q ein kompakter Quader, der die Menge $supp\ f_1 \cup supp\ g$ enthält. Dann verschwinden alle beteiligten Funktionen außerhalb von Q. Wegen der Stetigkeit können wir die Folgerung des Satzes von Dini anwenden und erhalten die Behauptung.

■

Als nächstes stellt sich die Frage, ob zu jeder gegebenen kompakten Menge eine solche Funktionenfolge (f_n) existiert.

Satz 15 *Für jede kompakte Menge $K \subset \mathbb{R}^m$ gibt es eine Folge (f_n) in $\mathcal{C}_c(\mathbb{R}^m)$ mit $f_n \downarrow \mathbf{1}_K$.*

Beweis. Wir wählen ein $\varepsilon > 0$ und irgendeine Folge (φ_n) stetiger Funktionen von $\mathbb{R}_+$ in sich, so daß $\varphi_n(0) = 1$ und $\varphi(\alpha) = 0$ für $\alpha > \varepsilon$ und schließlich $\varphi_n(\alpha) \downarrow 0$ für alle $\alpha > 0$. Dann setzen wir einfach $f_n(x) = \varphi_n(dist(x, K))$. Die Folge (f_n) fällt tatsächlich monoton. Die f_n nehmen auf der Menge K den Wert 1 an und verschwinden für alle x mit $dist(x, K) > \varepsilon$, also sind sie in $\mathcal{C}_c(\mathbb{R}^m)$. Für $x \notin K$ haben wir wegen der Abgeschlossenheit von K auch $dist(x, K) > 0$, d.h. $f_n(x) \downarrow 0$. Zusammen ergibt sich die Behauptung. ■

Wir studieren nun einige einfache Eigenschaften der Volumenfunktion auf den kompakten Mengen. Offensichtlich ist $vol_m(\emptyset) = 0$, denn $1_\emptyset = 0 \in C_c(\mathbb{R}^m)$, also kommt $\int 0 dx = 0$ ab Element der Klammer in der Definition von $\mathrm{vol}_m(\emptyset)$ vor. Einfach ist

Satz 16 *Das Volumen vol_m ist* monoton*: Sind K und L zwei kompakte Mengen mit $K \subset L$, dann ist $vol_m(K) \leq vol_m(L)$.*

Denn je größer die kompakte Menge, desto weniger Funktionen gibt es, die die zugehörige Indikatorfunktion majorisieren, und desto größer ist daher das Infimum über die Integrale über diese Funktionen in der Definition von vol_m.

Wichtig sind die beiden folgenden Sätze.

Satz 17 *Das Volumen ist* modular*, d.h. für zwei beliebige kompakte Mengen* K, L *gilt*

$$vol_m(K \cup L) + vol_m(K \cap L) = vol_m(K) + vol_m(L).$$

Insbesondere ist vol_m additiv*, d.h. für je zwei* disjunkte *kompakte Mengen gilt*

$$vol_m(K \cup L) = vol_m(K) + vol_m(L).$$

Beweis. Wir wählen zwei Folgen (f_n) und (g_n) in $\mathcal{C}_c(\mathbb{R}^m)$ mit $f_n \downarrow \mathbf{1}_K$ und $g_n \downarrow \mathbf{1}_L$. Dann gilt

$$\max(f_n, g_n) \downarrow \max(\mathbf{1}_K, \mathbf{1}_L) = \mathbf{1}_{K \cup L}$$

und analog

$$\min(f_n, g_n) \downarrow \min(\mathbf{1}_K, \mathbf{1}_L) = \mathbf{1}_{K \cap L}.$$

Ferner ist für je zwei reelle Zahlen a, b stets $\max(a, b) + \min(a, b) = a + b$, also auch $\max(f_n, g_n) + \min(f_n, g_n) = f_n + g_n$. Damit kann man die Modularitätsformel auf die Linearität des Integrals zurückführen.
Wenn die beiden Mengen disjunkt sind, folgt nun die Additivitätsformel aus $vol_m(K \cap L) = vol_m(\emptyset) = 0$. ■

Ferner gilt die folgende Stetigkeitseigenschaft des Volumens.

Satz 18 *Die Volumenfunktion* vol_m *auf den kompakten Mengen ist* nach unten σ-stetig*, d.h. für jede monoton fallende Folge* (K_l) *kompakter Mengen gilt*

$$vol_m\left(\bigcap_{l=1}^{\infty} K_l\right) = \lim_{l \to \infty} vol_m(K_l).$$

Beweis. Sei $(K_l)_l$ monoton fallend und K der Durchschnitt $\bigcap_{l=1}^{\infty} K_l$. Aus der Monotonie folgt $vol_m(K) \leq \lim_{l \to \infty} vol_m(K_l)$.
Der folgende Trick für den Beweis der umgekehrten Ungleichung wird uns ähnlich später wieder begegnen. Es gilt $\mathbf{1}_{K_l} \downarrow \mathbf{1}_K$. Nach Satz 15 können wir für jedes l eine Folge $(f_n^l)_n$ in $\mathcal{C}_c(\mathbb{R}^m)$ finden, so daß $f_n^l \downarrow \mathbf{1}_{K_l}$ für $n \to \infty$. Für jedes $n \in \mathbb{N}$ betrachten wir die Funktion

$$f_n = \min(f_n^1, \cdots, f_n^n).$$

Einerseits gilt $f_n^l \geq \mathbf{1}_{K_l} \geq \mathbf{1}_{K_n}$ für alle $l \leq n$ und damit ist auch $f_n \geq \mathbf{1}_{K_n}$ und

$$\int_{\mathbb{R}^m} f_n(x)\, dx \geq vol_m(K_n)$$

für alle n. Andererseits hängen für festes l die f_n^l monoton fallend vom unteren Index ab, es gilt also $f_{n+1} \leq f_n$. Wegen $f_n \leq f_n^l$ für alle $n \geq l$ gilt $\mathbf{1}_K \leq \lim_n f_n^l = \mathbf{1}_{K_l}$ für jedes l, also $f_n \downarrow \mathbf{1}_K$ und somit $vol_m(K) = \lim_n \int_{\mathbb{R}^m} f_n(x)\, dx \geq \lim_n vol_m(K_n)$. ■

3.2 Konkrete Volumenberechnungen

Das einfachste Verfahren zur konkreten Volumenberechnung beruht wie die Definition des Integrals am Anfang des Kapitels auf sukzessiver Integration der einzelnen Variablen. Der folgende Satz gibt eine Variante dieses Prinzips. Er ist ein Spezialfall des später zu beweisenden Satzes von Fubini.

Satz 19 *(Cavalieri-sches Prinzip) Sei* $[a,b]$ *ein kompaktes reelles Intervall und* $K \subset [a,b] \times \mathbb{R}^m$ *eine kompakte Menge. Für jedes* $t \in [a,b]$ *betrachte die (automatisch kompakte) Schnittmenge* $K_t = \{x \in \mathbb{R}^m : (t,x) \in K\}$*. Wenn die Funktion* $t \mapsto vol_m(K_t)$ *stetig auf dem Intervall* $[a,b]$ *ist, dann gilt*

$$vol_{m+1}(K) = \int_a^b vol_m(K_t)\, dt.$$

Beweis. Sei $\varepsilon > 0$. Indem wir die Funktion $t \mapsto vol_m(K_t)$ außerhalb des Intervalls schnell genug gegen 0 abfallen lassen, können wir sie zu einer stetigen Funktion $\varphi \in \mathcal{C}_c(\mathbb{R})$ fortsetzen mit

$$\int_a^b vol_m(K_t)\ dt \le \int_{\mathbb{R}} \varphi(t)\ dt \le \int_a^b vol_m(K_t)\ dt + \varepsilon.$$

Wir wissen, daß es eine Folge (f_n) von Funktionen in $\mathcal{C}_c(\mathbb{R}^{m+1})$ gibt mit $f_n(t,x) \downarrow \mathbf{1}_K(t,x)$. Dann gilt für jedes einzelne t, daß die Schnittfunktionen $f_n(t,\cdot) : x \mapsto f_n(t,x)$ auch monoton gegen $\mathbf{1}_{K_t}$ fallen, wobei $K_t = \emptyset$ für $t \notin [a,b]$. Wir setzen $\varphi_n(t) = \int_{\mathbb{R}^m} f_n(t,x)\, dx$ für alle reellen t. Also gilt $\lim_n \varphi_n(t) \downarrow vol_m(K_t) = \varphi(t)$ für alle $t \in [a,b]$ und $\varphi_n(t) \downarrow vol_m(\emptyset) = 0 \le \varphi(t)$ für $t \notin [a,b]$. Nach Definition des Integrals von f_n sind die Funktionen φ_n stetig und es gilt

$$vol_{m+1}(K) = \lim_{n\to\infty} \int_{\mathbb{R}^{m+1}} f_n(t,x)\, dx\, dt = \lim_{n\to\infty} \int_{\mathbb{R}} \varphi_n(t)\ dt$$

Nach der Folgerung zum Satz von Dini kann man diesen Grenzwert nach oben abschätzen durch Integration über $[a,b]$

$$\lim_{n\to\infty} \int_a^b \varphi_n(t)\ dt \le \int_a^b \varphi(t)\ dt = \int_a^b vol_m(K_t) dt$$

und nach unten durch

$$\int_{\mathbb{R}} \varphi(t)\ dt \le \lim_{n\to\infty} \int_a^b vol_m(K_t) dt + \varepsilon.$$

Daraus folgt die Behauptung. ∎

Als triviales Beispiel ergibt sich hieraus induktiv die eingangs erwähnte Formel $vol_m(I_1 \times \cdots \times I_m) = |I_1| \cdots |I_m|$. Für andere Anwendungen dieses Prinzips ist ein zusätzliches Hilfsmittel nützlich: Aus dem Transformationsverhalten von Integralen bei affinen Abbildungen können wir natürlich auch eine analoge Aussage für Volumina kompakter Mengen folgern.

Satz 20 *Sei* $A : \mathbb{R}^m \to \mathbb{R}^m$ *eine nichtsinguläre lineare Transformation und* $y \in \mathbb{R}^m$ *ein Vektor. Dann gilt für jede kompakte Menge* K *die Beziehung*

$$vol_m(AK + y) = |detA|\ vol_m(K).$$

Beweis. 1. Wir wählen wieder eine Folge (f_n) in $\mathcal{C}_c(\mathbb{R}^m)$ mit $f_n \downarrow \mathbf{1}_K$. Dann gilt auch $f_n \circ A^{-1}(x) = f_n(A^{-1}x) \downarrow \mathbf{1}_K(A^{-1}x) = \mathbf{1}_{AK}(x)$ für alle $x \in \mathbb{R}^m$. Ferner ist $f_n \circ A^{-1} \in \mathcal{C}_c(\mathbb{R}^m)$. Also folgt

$$\begin{aligned} vol_m(AK) &= \lim_{n\to\infty} \int_{\mathbb{R}^m} f_n(A^{-1}x)\ dx = \lim_{n\to\infty} \frac{1}{|detA^{-1}|} \int_{\mathbb{R}^m} f_n(x)\ dx \\ &= \lim_{n\to\infty} |detA| \int_{\mathbb{R}^m} f_n(x)\ dx = |detA| vol_m(K). \end{aligned}$$

2. Es bleibt die Aussage $vol_m(L+y) = vol_m(L)$ für alle Kompakta L und alle y zu beweisen, was analog zum ersten Beweisschritt auf die Translationsinvarianz des Integrals zurückgeführt werden kann. ■

Als Beispiel ergibt sich etwa das Volumen einer Kugel.

Satz 21 *Für $r > 0$ sei $K_m(r) = \{x \in \mathbb{R}^m : \|x\| \le r\}$ die Kugel mit Radius r. Sei $\tau_m = vol_m(K_m(1))$. Dann gilt $vol_m(K_m(r)) = r^m \tau_m$ und es ist für alle $k \ge 1$*

$$\tau_{2k} = \frac{\pi^k}{k!},\ \textit{und}\ \tau_{2k+1} = \frac{2^{k+1}}{1\cdot 3\cdots(2k+1)}\pi^k$$

oder gleichbedeutend für alle $m \ge 1$

$$\tau_m = \frac{\pi^{m/2}}{\Gamma(\frac{m}{2}+1)}.$$

Beweis. 1. Nach dem vorigen Satz ist

$$vol_m(K_m(r)) = vol_m(rK_m(1)) = |det(rI)| vol_m(K_m(1)) = r^m vol_m(K_m(1)).$$

2. Nach dem Cavalierischen Prinzip ist

$$\begin{aligned} \tau_m &= \int_{-1}^{1} vol_{m-1}(\{x \in \mathbb{R}^{m-1} : (t,x) \in K_m(1)\})\ dt \\ &= \int_{-1}^{1} vol_{m-1}(\{x \in \mathbb{R}^{m-1} : x_1^2 + \cdots + x_{m-1}^2 \le 1 - t^2\})\ dt \\ &= \int_{-1}^{1} vol_{m-1}(K_{m-1}(\sqrt{1-t^2}))\ dt = \tau_{m-1} \int_{-1}^{1} \left(\sqrt{1-t^2}\right)^{m-1}\ dt. \end{aligned}$$

Das letzte Integral bezeichnen wir mit c_m. Es läßt sich mit der Variablentransformation $t = \sin x$ wie folgt umformen:

$$\begin{aligned} c_m &= \int_{-1}^{1} (\sqrt{1-t^2})^{m-1}\ dt = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} (1-\sin^2 x)^{\frac{m-1}{2}} \cos x\ dx \\ &= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} (\cos^2 x)^{\frac{m-1}{2}} \cos x\ dx = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^m x\ dx \\ &= 2\int_{0}^{\frac{\pi}{2}} \cos^m x\ dx = 2\int_{0}^{\frac{\pi}{2}} \sin^m x\ dx. \end{aligned}$$

Nach Analysis I ist dies gleich $\pi \prod_{l=1}^{k} \frac{2l-1}{l}$ für $m = 2k$ und gleich $2\prod_{l=1}^{k} \frac{2l}{2l+1}$ für $m = 2k+1$. Insbesondere ist $c_m c_{m-1} = \frac{2\pi}{m}$ und somit

$$\tau_m = c_m \tau_{m-1} = c_m c_{m-1} \tau_{m-2} = \frac{2\pi}{m}\tau_{m-2}.$$

Außerdem ist der Beginn der Folge gegeben durch $\tau_1 = 2, \tau_2 = c_2\tau_1 = 2\pi\frac{1}{2} = \pi$. Damit erhalten wir die in der Behauptung angegebenen Formeln für τ_{2k} und τ_{2k+1}. Die Vereinheitlichung mit Hilfe der Gamma-Funktion ergibt sich dann aus den Beziehungen $\Gamma(k+1) = k!$ (dies liefert Übereinstimmung für $m = 2k$) und

$$\Gamma(k+\frac{3}{2}) = (k+\frac{1}{2})\Gamma(k+\frac{1}{2}) = \cdots = \prod_{l=0}^{k} \frac{2l+1}{2}\sqrt{\pi}.$$

Hier haben wir verwendet, daß $\Gamma(\frac{1}{2}) = \sqrt{\pi}$. Dies liefert die Behauptung für ungerade m. ■

Als Spezialfall ergibt sich also für das gewöhnliche Volumen der dreidimensionalen Einheitskugel

$$\tau_3 = \frac{\pi^{3/2}}{\Gamma(3/2+1)} = \frac{\pi\sqrt{\pi}}{3/4\sqrt{\pi}} = \frac{4}{3}\pi.$$

Als Motivation für den nächsten Paragraphen betrachten wir zum Abschluß noch ein eindimensionales Beispiel, die Cantormenge C. Es gibt eine Folge $(J_n)_n$ von disjunkten offenen Intervallen, derart daß $\sum_{n=1}^{\infty} |J_n| = 1$ und das Einheitsintervall $[0,1]$ die disjunkte Vereinigung von C mit $\bigcup_{n=1}^{\infty} J_n$ ist. Wir wollen aus dieser Beschreibung schließen, daß $vol_1(C) = 0$ ist. Das ist mit den bisherigen Mitteln zwar möglich, aber etwas umständlich. Das folgende anschauliche Argument läuft aber einfach so: Für jedes der offenen Intervalle J_n sollte auch $vol_1(J_n) = |J_n|$ sein, das Volumen der disjunkten Vereinigung dieser Intervalle sollte $\sum_{n=1}^{\infty} vol_1(J_n) = 1$ sein, und C als Differenzmenge von $[0,1]$ und dieser Vereinigung sollte als Volumen die Differenz der betreffenden Volumina, also das Volumen 0 haben. Wir werden im folgenden Paragraphen den Volumenbegriff so erweitern, daß diese Schlußweise unter sehr allgemeinen Bedingungen gerechtfertigt wird.

Kapitel 4

Fortsetzung zu einem Maß

In diesem Paragraphen werden wir das **Lebesgue-Maß** als Fortsetzung des Volumenbegriffs einführen. Gleichzeitig wollen wir die Fortsetzungsprozedur so formulieren, daß sie zur Konstruktion allgemeiner Maße anwendbar bleibt. Zunächst betrachten wir σ-Algebren. Dies sind die natürlichen Definitionsbereiche von Maßen.

4.1 σ-Algebren

Definition 3 *Sei X eine Menge und $\mathcal{F}$ eine Teilmenge der Potenzmenge von X. Die Elemente von $\mathcal{F}$ sind also Teilmengen von X. Dann heißt $\mathcal{F}$ eine* σ-Algebra *über X, falls folgendes gilt:*

a) Es ist $X \in \mathcal{F}$,

b) aus $A \in \mathcal{F}$ folgt, daß auch $A^c \in \mathcal{F}$ ist (A^c ist das Komplement von A in X), und

c) für jede Folge $A_1, A_2, \cdots$ von Elementen von $\mathcal{F}$ ist auch $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$.

Bemerkung: 1. Aus a) und b) folgt, daß auch die leere Menge zu $\mathcal{F}$ gehört, denn es ist $\emptyset = X^c$.
2. Aus b) und c) folgt, daß für jede Folge $A_1, A_2, \cdots$ von Elementen von $\mathcal{F}$ auch $\bigcap_{n=1}^{\infty} A_n \in \mathcal{F}$ ist, denn es gilt zunächst $A_n^c \in \mathcal{F}$ für jedes n wegen b), daher $\bigcup_n A_n^c \in \mathcal{F}$ wegen c) und schließlich wegen b) also auch

$$\bigcap_{n=1}^{\infty} A_n = \left(\bigcup_{n=1}^{\infty} A_n^c \right)^c \in \mathcal{F}.$$

Insbesondere enthält $\mathcal{F}$ mit je zwei Mengen A, B auch $A \cup B$ und $A \cap B$ und damit wegen *b)* auch die Differenzmenge $A \backslash B = A \cap B^c$.
3. Die Potenzmenge $\mathcal{P}(X)$ selbst ist die größte oder umfangreichste σ-Algebra über X, dagegen ist das zweielementige Mengensystem $\{\emptyset, X\}$ die kleinste σ-Algebra über X.

Satz 22 *Ist $(\mathcal{F}_i)_{i\in I}$ eine beliebige Familie von σ-Algebren über X, so ist auch ihr Durchschnitt $\bigcap_{i\in I}\mathcal{F}_i$ eine σ-Algebra über X.*

Beweis. Wir beschränken uns auf den Nachweis der Bedingung c) in der Definition. Seien $A_1, A_2, \cdots$ eine Folge von Elementen von $\bigcap_{i\in I}\mathcal{F}_i$. Dann ist also $A_n \in \mathcal{F}_i$ für alle $n \in \mathbb{N}$ und alle $i \in I$. Für festes $i \in I$ folgt dann auch $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{F}_i$ weil $\mathcal{F}_i$ eine σ-Algebra ist. Dies gilt für jedes i, also ist $\bigcup_{n\in\mathbb{N}} A_n \in \bigcap_{i\in I}\mathcal{F}_i$. ■

Dieser Satz erlaubt die folgende Definition.

Definition 4 *Sei $\mathcal{E}$ irgendein System von Teilmengen von X. Der Durchschnitt aller σ-Algebren, in denen $\mathcal{E}$ enthalten ist, heißt die* von $\mathcal{E}$ erzeugte σ-Algebra *und wird mit $\sigma(\mathcal{E})$ bezeichnet.*

Bemerkungen 1. Diese Begriffsbildung ist also ähnlich wie der von einer Menge E von Vektoren in einem Vektorraum V erzeugte Vektorraum: Er ist der kleinste Vektorraum, der alle Vektoren aus E enthält, oder auch der Durchschnitt aller Unter-Vektorräume von V, die E umfassen. Während man aber die Elemente dieses Vektorraums explizit als die Menge aller Linearkombinationen von Elementen von E darstellen kann, ist eine solche explizite Darstellung der Elemente von $\sigma(\mathcal{E})$ durch die Elemente von $\mathcal{E}$ i.a. nicht möglich (höchstens unter Zuhilfenahme der sogenannten transfiniten Induktion).

2. Aus der Definition folgt: Sind zwei Mengensysteme $\mathcal{E}_1, \mathcal{E}_2$ derart, daß $\mathcal{E}_1 \subset \sigma(\mathcal{E}_2)$ dann ist auch $\sigma(\mathcal{E}_1) \subset \sigma(\mathcal{E}_2)$ denn die σ-Algebra $\sigma(\mathcal{E}_2)$ ist nach Voraussetzung eine der σ-Algebren, deren Durchschnitt $\sigma(\mathcal{E}_1)$ ist. Insbesondere gilt: Je größer ein Mengensystem, desto größer die von ihm erzeugte σ-Algebra.

Beispiel. Sei $\mathcal{Q}$ das System aller kompakten Quader im $\mathbb{R}^m$. Dann ist jede offene Menge $U \subset \mathbb{R}^m$ ein Element der σ-Algebra $\sigma(\mathcal{Q})$.

Beweis. Jeder Punkt $x \in U$ liegt in einem Quader der Form $Q = [r_1, s_1] \times \cdots \times [r_m, s_m]$ mit rationalen Zahlen r_i, s_i, so daß auch noch der ganze Quader Q in U enthalten ist. Die Menge aller in U enthaltenen Quader mit rationalen Seiten ist abzählbar, also gibt es eine Folge $(Q_n)_n$ von kompakten Quadern mit $U = \bigcup_{n=1}^{\infty} Q_n$. Weil $Q_n \in \mathcal{Q} \subset \sigma(\mathcal{Q})$ ist und $\sigma(\mathcal{Q})$ eine σ-Algebra ist, ist auch $U \in \sigma(\mathcal{Q})$. ■

Eine σ-Algebra hat viele verschiedene Erzegendensysteme.

Satz 23 *Die folgenden vier Mengensysteme im $\mathbb{R}^m$, nämlich*
das System $\mathcal{Q}$ aller kompakten Quader,
das System $\mathcal{K}$ aller kompakten Teilmengen
das System $\mathcal{O}$ aller offenen Teilmengen
das System $\mathcal{A}$ aller abgeschlossenen Teilmengen
erzeugen alle die gleiche σ-Algebra, d.h. es ist $\sigma(\mathcal{Q}) = \sigma(\mathcal{K}) = \sigma(\mathcal{O}) = \sigma(\mathcal{A})$.

Beweis. Aus der Kette $\mathcal{Q} \subset \mathcal{K} \subset \mathcal{A}$ folgt auch $\sigma(\mathcal{Q}) \subset \sigma(\mathcal{K}) \subset \sigma(\mathcal{A})$. Weil jede abgeschlossene Menge als Komplement einer offenen Menge in $\sigma(\mathcal{O})$ liegt, ist $\mathcal{A} \subset \sigma(\mathcal{O})$ und nach dem Beipiel oben ist $\mathcal{O} \subset \sigma(\mathcal{Q})$. Aus diesen beiden Inklusionen folgt dann auch $\sigma(\mathcal{A}) \subset \sigma(\mathcal{O}) \subset \sigma(\mathcal{Q})$. ■

Definition 5 *Sei (X, d) ein metrischer Raum. Sei $\mathcal{O}(X)$ das System aller offenen Teilmengen von X. Die von $\mathcal{O}(X)$ erzeugte σ-Algebra heißt* Borel-σ-Algebra *von X und wird mit $\mathcal{B}(X)$ bezeichnet. Ihre Elemente nennt man die* Borel-Mengen *von X.*

Man kann also, ausgehend von offenen und abgeschlossenen Mengen nach Herzenslust abzählbare Vereinigungen, Durchschnitte und Komplemente bilden, ohne je den Bereich der Borel-Mengen zu verlassen.

4.2 Maße

Definition 6 *Sei X eine Menge und $\mathcal{F}$ eine σ-Algebra über X. Dann heißt eine Abbildung $\mu : \mathcal{F} \to [0, \infty]$ ein* **Maß**, *wenn $\mu(\emptyset) = 0$ gilt und μ* σ-additiv *ist, d.h. wenn für jede Folge von disjunkten Mengen $F_n \in \mathcal{F}$ gilt*

$$\mu\left(\bigcup_{n=1}^{\infty} F_n\right) = \sum_{n=1}^{\infty} \mu(F_n).$$

Definition 7 *Ist X eine Menge und $\mathcal{F}$ eine σ-Algebra auf X, so heißt das Paar $(X, \mathcal{F})$ ein* **meßbarer Raum**. *Ist weiter μ ein Maß auf $\mathcal{F}$, so heißt das Tripel $(X, \mathcal{F}, \mu)$ ein* **Maßraum**. *Die Elemente von $\mathcal{F}$ heißen* **meßbare Mengen**.

Wir sammeln zunächst einfache Eigenschaften von Maßen.

Bemerkungen 1. Die Bedingung $\mu(\emptyset) = 0$ folgt automatisch aus der σ-Additivität, wenn es irgendeine Menge $A \in \mathcal{F}$ gibt mit $\mu(A) < \infty$, denn wenn man $F_1 = A$ und $F_n = \emptyset$ für $n \geq 2$ wählt, so folgt $\mu(A) = \mu(A) + \sum_{n=2}^{\infty} \mu(\emptyset)$ und dies ist nur möglich wenn $\mu(\emptyset) = 0$.
2. Ein Maß ist auch endlich additiv: Wenn A und B disjunkt sind. so ist $\mu(A \cup B) = \mu(A) + \mu(B)$, denn man braucht nur $F_1 = A, F_2 = B$ und $F_n = \emptyset$ für $n \geq 3$ zu wählen. Daraus folgt weiterhin, daß ein Maß monoton ist, denn ist $A \subset B$ so ist $\mu(B) = \mu(A \cup (B \setminus A)) = \mu(A) + \mu(B \setminus A) \geq \mu(A)$.
3. Ein Maß ist **subadditiv**: Für je zwei nicht notwendig disjunkte Mengen A, B in $\mathcal{F}$ gilt $\mu(A \cup B) \leq \mu(A) + \mu(B)$; denn wegen der Additivität und der Monotonie gilt

$$\mu(A \cup B) = \mu(A \cup (B \setminus A)) = \mu(A) + \mu(B \setminus A) \leq \mu(A) + \mu(B).$$

Wir schreiben $A_n \uparrow A$ wenn die Mengen-Folge (A_n) monoton wachsend ist mit Vereinigung A; analog wird $A_n \downarrow A$ definiert.

Satz 24 *a) Ein Maß ist nach oben σ-stetig: Aus $A_n \in \mathcal{F}$ mit $A_n \uparrow A$ folgt $\mu(A_n) \to \mu(A)$.*
b) Ein Maß ist nach unten σ stetig: Aus $A_n \in \mathcal{F}$ mit $A_n \downarrow A$ und $\mu(A_1) < \infty$ folgt $\mu(A_n) \to \mu(A)$.
c) Ein Maß ist σ-subadditiv: Für jede Folge von nicht notwendig disjunkten Mengen $A_n \in \mathcal{F}$ gilt

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu(A_n).$$

Beweis. a) Wir betrachten die Mengen $F_1 = A_1$ und $F_n = A_n \backslash A_{n-1}$ für $n \geq 2$. Die F_n sind disjunkt und es gilt $A_n = \bigcup_{k=1}^n F_n$ für jedes $n \in \mathbb{N}$, also wegen der endlichen Additivität $\mu(A_n) \sum_{k=1}^n \mu(F_n)$. Damit ist

$$\begin{aligned} \mu(A) &= \mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \mu\left(\bigcup_{k=1}^{\infty} F_k\right) = \sum_{k=1}^{\infty} \mu(F_k) \\ &= \lim_{n\to\infty} \sum_{k=1}^{n} \mu(F_k) \lim_{n\to\infty} \mu(A_n). \end{aligned}$$

b) Wir setzen $B_n = A_1 \backslash A_n$ und $B = A_1 \backslash A$. Dann gilt $B_n \uparrow B$ und in $A_1 = A_n \cup B_n = A \cup B$ sind die Vereinigungen jeweils disjunkt. Daher gilt $\mu(A) + \mu(B) = \mu(A_n) + \mu(B_n)$ für jedes n. Wegen a) gilt $\mu(B) = \lim_n \mu(B_n)$ also auch $\mu(A) = \lim_n \mu(A_n)$. Beachte, daß wegen der Voraussetzung $\mu(A_1) < \infty$ alle beteiligten Zahlen endlich sind.
c) Wir setzen $C_N = \bigcup_{n=1}^N A_n$. Die Subadditivität liefert $\mu(C_N) \leq \sum_{n=1}^N \mu(A_n)$. Mit Teil a), angewendet auf die C_N, folgt die Behauptung für $N \to \infty$. ■

Definition 8 *Wir nennen eine Menge $N \in \mathcal{F}$ mit $\mu(N) = 0$ eine* **Nullmenge** *oder genauer eine μ-***Nullmenge**.

Aus der σ-Subadditivität folgt, daß die Vereinigung von abzählbar vielen μ-Nullmengen wieder eine μ-Nullmenge ist. Die Nullmengen sind deshalb nützlich, weil man für die meisten Zwecke ignorieren kann, was auf ihnen passiert.

Unser Ziel ist der Nachweis, daß die Volumenfunktion vol_m auf den kompakten Mengen zu einem Maß auf einer die kompakten Mengen enthaltenden σ-Algebra $\mathcal{L}^m$ über $\mathbb{R}^m$ fortgesetzt werden kann. Dann ist notwendig auch $\mathcal{B}(\mathbb{R}^m) \subset \mathcal{L}^m$, diese Fortsetzung enthält also insbesondere die Definition des Volumens beliebiger Borel-Mengen $B \in \mathcal{B}(\mathbb{R}^m)$. Wegen der Additivität in einer Dimension auch das Maß von offenen und halboffenen Intervallen ist gleich ihrer Länge, also ist das am Ende des letzten Paragraphen angegebene Argument für $vol_1(C) = 0$ korrekt.

Der zentrale Fortsetzungssatz für das Volumen lautet

Satz 25 *Für beliebige Teilmengen A von $\mathbb{R}^m$ setze*

$$\lambda_*^m(A) = \sup\{vol_m(K) : K \ kompakt,\ K \subset A\}.$$

Sei ferner

$$\mathcal{L}^m = \{A \subset \mathbb{R}^m : \lambda_*^m(K \cap A) + \lambda_*^m(K \backslash A) = vol_m(K) \ falls\ K\ kompakt\}.$$

Dann ist $\mathcal{L}^m$ eine σ-Algebra über $\mathbb{R}^m$ und die Funktion $\lambda^m = \lambda_*^m{}_{|\mathcal{L}^m}$ ist ein Maß auf $\mathcal{L}^m$, welches die Volumenfunktion vol_m von $\mathcal{K}$ fortsetzt.

Definition 9 *Das durch diesen Satz gegebene Maß heißt das m-dimensionale* **Lebesgue-Maß**. *Die σ-Algebra $\mathcal{L}^m$ heißt die σ-Algebra der* **Lebesgue-meßbaren** *Mengen. Die Funktion λ_*^m auf der Potenzmenge von $\mathbb{R}^m$ heißt das vom Lebesgue-Maß induzierte* **innere** *Maß.*

Bevor wir den Beweis im nächsten Abschnitt durchführen, machen wir noch ein paar einfache

Bemerkungen 1. Wegen der Monotonie von vol_m ist $\lambda^m(K) = \lambda_*^m(K) = vol_m(K)$ für alle $K \in \mathcal{K}$.
2. Zu jeder Menge $A \in \mathcal{L}^m$ gibt es nach Definition von λ_*^m bzw. λ^m eine (o.B.d.A. wachsende) Folge (K_n) kompakter Teilmengen von A mit $\lambda^m(K_n) \to \lambda^m(A)$. Ist zusätzlich $\lambda^m(A) < \infty$, so folgt daraus für die Menge $\tilde{A} = \bigcup_{n=1}^{\infty} K_n$, daß $\lambda^m(A \backslash \tilde{A}) = \mu(A) - \lim_n \mu(K_n) = 0$. Also unterscheidet sich A nur auf der λ^m-Nullmenge $N = A \backslash \tilde{A}$ von der Menge $\tilde{A}$.
3. Da man den $\mathbb{R}^m$ durch eine Folge Q_n kompakter Quader überdecken kann, ist jede Menge $A \in \mathcal{L}^m$ mit $\lambda^m(A) = \infty$ Vereinigung der Mengen $A \cap Q_n$, die jeweils endliches Maß haben. Wenn wir die vorige Bemerkung auf diese Mengen anwenden und beachten, daß eine abzählbare Vereinigung von Nullmengen eine Nullmenge ist und daß zweitens kompakte Mengen Borelsch sind, erhalten wir: Jedes $A \in \mathcal{L}^m$ unterscheidet sich nur auf einer λ^m-Nullmenge von nneiner Teilmenge $\tilde{A} \in \mathcal{B}(\mathbb{R}^m)$ von A. Die beiden σ-Algebren $\mathcal{L}^m$ und $\mathcal{B}(\mathbb{R}^m)$ sind also nicht sehr verschieden bezüglich des Lebesgue-Maßes.
4. Offenbar ist jede einpunktige Menge im $\mathbb{R}^m$ eine Nullmenge, denn sie ist enthalten in Quadern von beliebig kleinem Volumen. Also ist auch jede abzählbare Menge wie etwa die Menge $\mathbb{Q}^m$ aller Vektoren mit rationalen Komponenten eine Nullmenge, obwohl $\mathbb{Q}^m$ dicht im $\mathbb{R}^m$ ist.

4.3 Beweis des Fortsetzungssatzes

Dieser Abschnitt wird erst wieder im Anhang benötigt. Er dient nur zum Rechtfertigung der Definition des Lebesgue-Maßes. Das wesentliche Argument hat aber nichts mit der Struktur des $\mathbb{R}^m$ zu tun. Der Beweis des Fortsetzungssatzes beruht auf der σ-Stetigkeit nach unten und der Modularität der Volumenfunktion auf den kompakten Mengen. Eine weitere wichtige Eigenschaft der Volumenfunktion ist, daß die Differenzmenge von zwei kompakten Mengen von innen gut durch kompakte Mengen approximiert werden kann; genauer gilt

Lemma 1 *Seien K, L zwei kompakte Teilmengen des $\mathbb{R}^m$ mit $L \subset K$. Dann ist*

$$\lambda_*^m(K \backslash L) = vol_m(K) - vol_m(L).$$

Beweis. Wir betrachten die Mengen $L_n = \{x \in K : dist(x, L) \leq \frac{1}{n}\}$ und $M_n = \{x \in K : dist(x, L) \geq \frac{1}{n}\}$. Die Folge (L_n) ist monoton fallend mit Durchschnitt L, also gilt wegen der σ-Stetigkeit nach unten $vol_m(L_n) \to vol_m(L)$. Offenbar ist $M_n \subset K \backslash L$ und $K = L_n \cup M_n$ für jedes n. Wegen der Modularität gilt also $vol_m(K) = vol_m(L_n) + vol_m(M_n) - vol(L_n \cap M_n) \leq vol_m(L_n) + vol_m(M_n)$. Für $n \to \infty$ folgt also $\lim_n vol_m(M_n) \geq vol_m(K) - \lim_n vol_m(L_n) = vol_m(K) - vol_m(L)$. Daraus ergibt sich die Behauptung. ∎

Damit wird der im letzten Abschnitt formulierte Fortsetzungssatz zu einem Spezialfall des folgenden Satzes, der in sehr viel allgemeineren Situationen anwendbar ist. Im Fall des Lebesgue-Maßes ist die Menge X der Raum $\mathbb{R}^m$, das

System $\mathcal{K}$ das der kompakten Teilmengen des $\mathbb{R}^m$ und die Funktion m die Volumenfunktion vol_m. Satz 26 stammt im wesentlichen von J. Kysinski. Weitere Kommentare finden sich im Anhang.

Satz 26 *Sei $\mathcal{K}$ ein System von Teilmengen einer Menge X welches stabil unter endlicher Vereinigungs- und abzählbarer Durchschnittsbildung ist. Sei $\mathsf{m} : \mathcal{K} \to \mathbb{R}_+$ eine modulare, monotone, nach unten σ-stetige Funktion mit $\mathsf{m}(\emptyset) = 0$. Für beliebige Mengen $A \subset X$ setze*

$$\mu_*(A) = \sup\{\mathsf{m}(K) : K \subset A,\ K \in \mathcal{K}\}.$$

Es gelte für alle Mengen K, L aus $\mathcal{K}$ mit $L \subset K$ die „Straffheits-Bedingung“

$$\mu_*(K \backslash L) = \mathsf{m}(K) - \mathsf{m}(L).$$

Dann ist das Mengensystem

$$\mathcal{L} = \{A \subset X : \mathsf{m}(K) = \mu_*(K \cap A) + \mu_*(K \backslash A)\ falls\ K \in \mathcal{K}\}$$

eine σ-Algebra über X mit $\mathcal{K} \subset \mathcal{L}$, und die Einschränkung μ von μ_* auf $\mathcal{L}$ ist ein Maß, das m fortsetzt.

Beweis. 1. Die Mengenfunktion μ_* ist superadditiv: Es gilt

$$\mu_*(A \cup B) \geq \mu_*(A) + \mu_*(B)\ wenn\ A, B \subset X\ disjunkt\ sind.$$

Sind nämlich K, L Elemente von $\mathcal{K}$ mit $K \subset A$ and $L \subset B$, so gilt

$$\mu_*(A \cup B) \geq \mathsf{m}(K \cup L) = \mathsf{m}(K) + \mathsf{m}(L)$$

weil die Modularität zusammen mit $\mathsf{m}(\emptyset) = 0$ die Additivität von m impliziert. Wenn man jetzt zum sup bezüglich K, L übergeht ergibt sich die Superadditivität. (Es gibt im Fall des Lebegue-Maßes Mengen $A \subset [0,1]$ derart daß $\mu_*(A) = \mu_*([0,1] \backslash A) = 0$. Also ist die Ungleichung im allgemeinen echt. Die Definition von $\mathcal{L}$ filtert gerade die gutartigen Mengen heraus, bei denen dies Phänomen nicht auftritt.)
2. Aus der Monotonie von m folgt, daß μ_* auf $\mathcal{K}$ mit m übereinstimmt. Ferner ergibt sich aus der Straffheitsbedingung, daß $L \in \mathcal{L}$ gilt für alle $L \in \mathcal{K}$, d.h. also ist die Funktion $\mu = \mu_{*|\mathcal{L}}$ ebenfalls eine Fortsetzung von m. Es bleibt zu zeigen daß $\mathcal{L}$ eine σ-Algebra und μ ein Maß ist.
3. Das Mengensystem $\mathcal{L}$ ist stabil gegenüber Komplementbildung: Wenn man A durch sein Komplement ersetzt, vertauschen sich in der Klammer in der Definition von $\mathcal{L}$ nur die beiden Terme.
4. Wir zeigen, daß $\mathcal{L}$ abgeschlossen gegenüber endlichen Vereinigungen (und damit wegen des letzten Beweisschrittes auch gegenüber endlichen Durchschnitten) ist und daß die Mengenfunktion μ subadditiv ist. Seien hierzu A_1, A_2 zwei Elemente von $\mathcal{L}$ und $K \in \mathcal{K}$. Zu $\varepsilon > 0$ können wir nach Definition von μ_* für $i = 1, 2$ jeweils zwei Mengen L_i, M_i in $\mathcal{K}$ finden mit

$$L_i \subset K \cap A_i\ und\ M_i \subset K \backslash A_i,$$

und

$$\mathsf{m}(L_i) + \mathsf{m}(M_i) \geq \mu_*(K \cap A_i) + \mu_*(K \backslash A_i) - \varepsilon = \mathsf{m}(K) - \varepsilon$$

wobei die Gleichheit wegen $A_i \in \mathcal{L}$ gilt. Wenn wir nun diese Ungleichung über $i = 1, 2$ summieren, erhalten wir mit der Modularität von m unter Vertauschung der Reihenfolge der Terme

$$\begin{aligned} &\mathsf{m}(L_1 \cup L_2) + \mathsf{m}(M_1 \cap M_2) + \mathsf{m}(L_1 \cap L_2) + \mathsf{m}(M_1 \cup M_2) \\ &= \mathsf{m}(L_1) + \mathsf{m}(L_2) + \mathsf{m}(M_1) + \mathsf{m}(M_2) \geq 2\mathsf{m}(K) - 2\varepsilon. \end{aligned}$$

Die beiden Mengen $L_1 \cap L_2$ und $M_1 \cup M_2$ sind in $K \cap (A_1 \cap A_2)$ bzw $K \backslash (A_1 \cap A_2)$ enthalten. Insbesondere sind sie disjunkt und in K enthalten. Also ist die Summe der letzten beiden Terme in der linken Seite dieser Abschätzung höchstens gleich $\mathsf{m}(K)$. Damit ist die Summe der ersten beiden Terme mindestens $\mathsf{m}(K) - 2\varepsilon$. Zusammen mit der Superadditivität erhalten wir wegen $L_1 \cup L_2 \subset K \cap (A_1 \cup A_2)$ und $M_1 \cap M_2 \subset K \backslash (A_1 \cup A_2)$, daß

$$\begin{aligned} \mathsf{m}(K) &\geq \mu_*(K \cap (A_1 \cup A_2)) + \mu_*(K \backslash (A_1 \cup A_2)) \\ &\geq \mathsf{m}(L_1 \cup L_2) + \mathsf{m}(M_1 \cap M_2) \geq \mathsf{m}(K) - 2\varepsilon. \end{aligned}$$

Für jedes $K \in \mathcal{K}$ und jedes ε gibt es L_i, M_i, für die diese Abschätzung gilt, also ergibt sich $A_1 \cup A_2 \in \mathcal{L}$.
Wenn wir in der letzten Abschätzung zusätzllich $K \subset A_1 \cup A_2$ voraussetzen, dann ist $K \backslash (A_1 \cup A_2) = \emptyset = M_1 \cap M_2$, also erhalten wir

$$\mathsf{m}(K) \leq \mathsf{m}(L_1 \cup L_2) + 2\varepsilon \leq \mathsf{m}(L_1) + \mathsf{m}(L_2) + 2\varepsilon \leq \mu(A_1) + \mu(A_2) + 2\varepsilon.$$

Nun gehen wir wieder zum Supremum über alle $K \subset A_1 \cup A_2$ und lassen anschließend ε gegen 0 gehen. Wir erhalten die Subadditivität $\mu(A_1 \cup A_2) \leq \mu(A_1) + \mu(A_2)$. Zusammen mit der Superadditivität ergibt sich auch die endliche Additivität, dh. „=“ gilt für disjunkte Mengen in $\mathcal{L}$. Es folgt wegen $B = A \cup (B \backslash A)$

$$\mu(B \backslash A) = \mu(B) - \mu(A), \ fallsA \subset B, \ A, B \in \mathcal{L}.$$

5. Als nächstes zeigen wir, daß abzählbaren Vereinigungen nicht aus $\mathcal{L}$ herausführen. Sei $(A_i)_{i \in \mathbb{N}}$ eine Folge in $\mathcal{L}$ und $A = \bigcup_{i=1}^{\infty} A_i$. Wir wollen zeigen, daß $A \in \mathcal{L}$ ist. Seien $K \in \mathcal{K}$ und $\varepsilon > 0$ gegeben. Sei $B_n = K \cap \bigcup_{i=1}^{n} A_i$ und $C_n = K \backslash \bigcup_{i=1}^{n} A_i$. Dann gilt $B_n \uparrow K \cap A$ und $C_n \downarrow K \backslash A$. Wähle eine Folge $(\varepsilon_i)_{i \in \mathbb{N}}$ mit $\sum_{i=1}^{\infty} \varepsilon_i < \varepsilon/2$ und wähle für jedes n rekursiv eine Menge $M_n \in \mathcal{K}$ derart, daß $M_0 = K, M_n \subset C_n \cap M_{n-1}$ und $\mu(M_n)(= \mathsf{m}(M_n)) \geq \mu(C_n \cap M_{n-1}) - \varepsilon_n$. Dann ist

$$C_n \backslash M_n \subset \bigcup_{i=1}^{n} (C_i \cap M_{i-1}) \backslash M_i$$

denn wenn x ein Punkt aus der Menge auf der linken Seite ist, gibt es ersten Index i mit $x \notin M_i$. Für diesen Index ist $x \in (C_i \cap M_{i-1}) \backslash M_i$. Aus der Additivität und Subadditivität von μ folgt

$$\mu(C_n) - \mu(M_n) = \mu(C_n \backslash M_n) \leq \sum_{i=1}^{n} \mu((C_i \cap M_{i-1}) \backslash M_i) \leq \sum_{i=1}^{n} \varepsilon_i.$$

Nach Voraussetzung über $\mathcal{K}$ und die σ-Stetigkeit von m ist $M = \bigcap_{n=1}^{\infty} M_n \in \mathcal{K}$ und

$$\begin{aligned} \mathsf{m}(M) &= \lim_{n\to\infty} \mathsf{m}(M_n) \ \geq \ \lim_{n\to\infty} \mu(C_n) - \sum_{i=1}^{\infty} \varepsilon_i \\ &= \lim_{n\to\infty} \big(\mathsf{m}(K) - \mu(B_n)\big) - \varepsilon/2 \\ &\geq \mathsf{m}(K) - \mu(B_n) - \varepsilon \end{aligned}$$

für hinreichend große n. Für diese n gilt $B_n \subset K \cap A$ und $M \subset K \backslash A$ und damit $\mu_*(K \cap A) + \mu_*(K \backslash A) \geq m(M) + m(B_n) \geq \mathsf{m}(K) - \varepsilon$. Dies zeigt $A \in \mathcal{L}$. Also ist $\mathcal{L}$ eine σ-Algebra.
6. Es bleibt die σ-Additivität von μ zu zeigen. Im Fall $K \subset A$ ist in der letzten Rechnung $M = \emptyset$ und daher

$$\mu(K) \leq \lim_n \mu(B_n) \leq \lim_n \mu(\bigcup_{i=1}^{n} A_i).$$

Zusammen mit der Monotonie erhalten wir durch Supremumsbildung über K die Beziehung $\mu(A) = \lim_n \mu(\bigcup_{i=1}^{n} A_i)$. Wenn die A_i zusätzlich disjunkt sind, so ist dieser Grenzwert wegen der Additivität von μ sogar gleich der Reihe $\sum_{i=1}^{\infty} \mu(A_i)$, d.h. μ ist ein Maß. ∎

Kapitel 5

Meßbare Funktionen

Definition 10 *Seien $(X, \mathcal{F})$ und $(Y, \mathcal{G})$ meßbare Räume. Eine Abbildung $f : X \longrightarrow Y$ heißt $\mathcal{F}$-$\mathcal{G}$-**meßbar** (oder kurz: meßbar), wenn*

$$f^{-1}(G) \in \mathcal{F} \quad \text{für alle } G \in \mathcal{G}$$

gilt. Speziell: Eine Abbildung $f : X \longrightarrow \mathbb{R}^m$ heißt $\mathcal{F}$-meßbar (oder meßbar), wenn f eine $\mathcal{F} - \mathcal{B}(\mathbb{R}^m)$-meßbare Abbildung ist.

Für Funktionen mit eindimensionalem Wertebereich tauchen oft die Werte $\pm\infty$ bei Grenzübergängen auf. Deshalb will man sie zulassen.

Notation: Seien $f, g : X \longrightarrow [-\infty, \infty]$ Abbildungen und $a \in \mathbb{R}$. Dann definiere

$$\begin{aligned} \{f \le a\} &:= \{x \in X | f(x) \le a\} = f^{-1}([-\infty; a]) \\ \{f = g\} &:= \{x \in X | f(x) = g(x)\} \end{aligned}$$

und analog $\{f < a\}$, $\{f \ge a\}$, $\{f > a\}$.

Definition 11 *Eine Funktion $f : X \longrightarrow [-\infty, +\infty]$ („numerische Funktion“) heißt $\mathcal{F}$-meßbar, wenn $\{f \le \alpha\} \in \mathcal{F}$ für alle $\alpha \in \mathbb{R}$.*

Bemerkungen: 1. Wir werden im übernächsten Lemma sehen, daß für reellwertige Funktionen kein Widerspruch zwischen diesen beiden Definitionen der Meßbarkeit besteht.

2. Die meßbaren Funktionen sind (mit einfachen später zu diskutierenden Einschränkungen) besonders gut für die Definition des Integrals geeignet.

3. Beim Nachweis der Meßbarkeit ist das folgende Kriterium sehr nützlich.

Lemma 2 *Es sei $\mathcal{E}$ ein System von Teilmengen von Y mit $\mathcal{G} = \sigma(\mathcal{E})$. Dann gilt: Die Abbildung $f : X \longrightarrow Y$ ist $\mathcal{F} - \mathcal{G}$-meßbar genau dann, wenn $f^{-1}(E) \in \mathcal{F}$ ist für alle $E \in \mathcal{E}$.*

Beweis.1. " $\Longrightarrow$ " ist klar, wegen $\mathcal{E} \subseteq \mathcal{G}$.
" $\Longleftarrow$". Sei $\mathcal{A} = \{A \subseteq Y | f^{-1}(A) \in \mathcal{F}\}$. Nach Voraussetzung gilt $\mathcal{E} \subseteq \mathcal{A}$. Weiter gelten die Implikationen

- Es ist $f^{-1}(Y) = X \in \mathcal{F}$. Also ist $Y \in \mathcal{A}$.
- Sei $A \in \mathcal{A}$. Dann ist $f^{-1}(Y \backslash A) = X \backslash f^{-1}(A) \in \mathcal{F}$, also auch $Y \backslash A \in \mathcal{A}$
- Ist $A_n \in \mathcal{A} \quad \forall n \in \mathbb{N}$, so ist $f^{-1}(\bigcup_{n\in\mathbb{N}} A_n) = \bigcup_{n\in\mathbb{N}} f^{-1}(A_n) \in \mathcal{F}$ und damit $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$.

Also ist $\mathcal{A}$ eine σ-Algebra, die $\mathcal{E}$ enthält, dh. es ist auch $\mathcal{G} = \sigma(\mathcal{E}) \subseteq \mathcal{A}$, woraus die Behauptung folgt. ■

Lemma 3 *Sei $(X, \mathcal{F})$ ein meßbarer Raum, $f : X \to [-\infty, \infty]$. Dann sind äquivalent.*
(1) $\{f \leq a\}$ ist meßbar für alle $a \in \mathbb{R}$, d.h. f ist meßbar gemäß Definition 11
(2) $\{f < a\}$ ist meßbar für alle $a \in \mathbb{R}$
(3) $\{f > a\}$ ist meßbar für alle $a \in \mathbb{R}$
(4) $\{f \geq a\}$ ist meßbar für alle $a \in \mathbb{R}$
Ist f endlichwertig so sind diese Bedingungen äquivalent zu
(5) f ist meßbar im Sinn von Definition 10.

Beweis. Wir zeigen nur (5) $\Longleftrightarrow$ (2): Sei f meßbar. Dann ist $f^{-1}(B)$ meßbar für alle $B \in \mathcal{B}(\mathbb{R})$ und daher $\{f < a\} = f^{-1}((-\infty, a])$ meßbar, da $(-\infty, a]$ offen ist.
Umgekehrt gelte (2). Seien $a, b \in \mathbb{R}, \quad a < b$. Dann ist

$$[a, b] = \bigcap_{n\in\mathbb{N}} [a, b + \frac{1}{n}) = \left(\bigcap_{n\in\mathbb{N}} (-\infty, b + \frac{1}{n}) \right) \backslash (-\infty, a)$$

also erzeugt das System $\mathcal{E}$ aller linken offenen Halbgeraden $(-\infty, c)$ auch die von den „Quadern“ $[a, b]$ erzeugte Borel-σ-Algebra. Aus der Bedingung (2) und dem vorigen Lemma ergibt sich, daß f meßbar ist im Sinn von Definition 10. ■

Satz 27 *Seien f, g meßbar. Dann sind die Mengen $\{f \leq g\}, \{f < g\}, \{f = g\}$ in $\mathcal{F}$.*

Beweis. Es ist $\{f < g\} = \bigcup_{r\in\mathbb{Q}} \{f \leq r\} \cap \{g > r\} \in \mathcal{F}$, also auch $\{f \leq g\} = \{g < f\}^c \in \mathcal{F}$ und $\{f = g\} = \{g \leq f\} \cap \{f \leq g\} \in \mathcal{F}$. ■

Bemerkung: 1. Ist $c \in \mathbb{R}$ und $f(x) = c$ für alle $x \in X$, so ist

$$\{f < a\} = \begin{cases} \emptyset \in \mathcal{F} & \text{für} \quad a \leq c \\ X \in \mathcal{F} & \text{für} \quad a > c \end{cases}$$

Also sind die konstanten Funktionen meßbar.
2.Ist $f : \mathbb{R}^m \longrightarrow \mathbb{R}^n$ stetig, so ist $f^{-1}(U)$ offen und damit meßbar für alle offenen Mengen U. Also sind nach Lemma 2 auch alle stetigen Funktionen meßbar.

Lemma 4 *Seien $(X, F), (Y, \mathcal{G}), (Z, \mathcal{H})$ meßbare Räume, $f : Y \to Z, \ g : X \to Y$ meßbare Abbildungen. Dann ist auch $f \circ g : X \to Z$ meßbar.*

Beweis. $H \in \mathcal{H} \Longrightarrow f^{-1}(H) \in \mathcal{G} \Longrightarrow (f \circ g)^{-1}(H) = g^{-1}(f^{-1}(H)) \in \mathcal{F}$. ■

Lemma 5 *Eine Funktion $f = (f_1, \ldots, f_n) : X \to \mathbb{R}^m$ ist $\mathcal{F}$-meßbar genau dann wenn $f_i : X \to \mathbb{R}$ meßbar ist für alle $i \in \{1, \ldots, m\}$.*

Beweis. „$\Longrightarrow$“ Sei f meßbar. Dann ist für jedes $a \in \mathbb{R}$ die Menge $A = (-\infty, a] \times \mathbb{R} \times \ldots \times \mathbb{R} \subseteq \mathbb{R}^m$ abgeschlossen, d.h. $A \in \mathcal{B}(\mathbb{R}^m)$. Also ist $\{f_1 < a\} = f^{-1}(A) \in \mathcal{F}$ für alle $a \in \mathbb{R}$. Damit ist f_1 meßbar. Genauso argumentiert man für die restlichen i.
„$\Longleftarrow$“. Sei f_i meßbar für alle i und $Q = [a_1, b_1] \times \ldots \times [a_n, b_n]$. Dann ist

$$f^{-1}(Q) = \bigcap_{i=1}^{n} f_i^{-1}([a_i, b_i]) \in \mathcal{F}$$

für alle Quader Q. Also ist nach Lemma 2 f meßbar. ∎

Bemerkung: Sind $f_1, f_2 : X \longrightarrow \mathbb{R}$ meßbar, so sind auch $f_1 + f_2$, $f_1 - f_2$ und $f_1 \cdot f_2$ meßbar.

Beweis. Die Abbildungen „+“, „-“, „·“: $\mathbb{R}^2 \longrightarrow \mathbb{R}$ sind stetig, also meßbar und nach dem Lemma ist $f : X \longrightarrow \mathbb{R}^2$ ebenfalls meßbar. Wegen $f_1 + f_2 =$„+“$\circ f$ sind $f_1 + f_2$ etc. ∎

Satz 28 *Sei $(X, \mathcal{F})$ ein Maßraum und $(f_n)_{n \in \mathbb{N}}$ eine Folge meßbarer numerischer Funktionen. a) Dann sind die punktweise definierten Funktionen $\sup_n f_n$, $\inf_n f_n$, $\limsup_{n\to\infty} f_n$, $\liminf_{n\to\infty} f_n$ auch meßbar.*
b) Es ist $F = \{x : \lim f_n(x) \text{ existiert}\} \in \mathcal{F}$ und $\lim_n 1_F f_n$ ist meßbar.

Beweis. a) Für alle $a \in \mathbb{R}$ ist $\{\sup f_n \le a\} = \cap_{n\in\mathbb{N}} \{f_n \le a\} \in \mathcal{F}$. Also ist $\sup f_n$ meßbar. Aus Symmetriegründen ist $\inf f_n$ meßbar. Daher sind auch die Funktionen

$$\limsup_{n\to\infty} f_n = \inf_m (\sup_{n \ge m} f_m)$$

und

$$\liminf_{n\to\infty} f_n = \sup_m (\inf_{n \ge m} f_n)$$

meßbar.
b) Die Menge $F = \{\limsup_n = \liminf_n\}$ ist in $\mathcal{F}$ und $\lim_n 1_F f_n = 1_F \liminf f_n$ ist meßbar. ∎

Definition 12 *Eine Funktion der Form $f(x) = \sum_{r=1}^{n} \alpha_i 1_{A_i}(x)$ mit $\alpha_i \in \mathbb{R}, A_i \in \mathcal{F}$ heißt* **elementare Funktion** *oder* **meßbare Treppenfunktion**.

Die meßbaren Treppenfunktionen bilden offenbar ein Vektorraum.

Lemma 6 *Eine elementare Funktion f hat eine Darstellung $\sum_{j=1}^{m} \tilde{\alpha}_j 1_{\tilde{A}_j}$ mit disjunkten Mengen $\tilde{A}_j \in \mathcal{F}$. Insbesondere ist f meßbar.*

Beweis. Sei $I_1, \ldots, I_m$ eine Nummerierung der endlich vielen Teilmengen I von $\{1, \ldots, n\}$. Sei für jedes $j \in \{1, \cdots, m\}$ die Menge $\tilde{A}_j$ definiert durch

$$\tilde{A}_j = \bigcap_{i \in I_j} A_i \cap \bigcap_{i \notin I_j} A_i^c.$$

Dann ist $\tilde{A}_j \in \mathcal{F}$ für jedes j. Ein Punkt $x \in X$ liegt genau in einem $\tilde{A}_j$, nämlich dann wenn die zu j gehörige Menge I_j gleich der Menge $\{i : x \in A_i\}$ ist. Also bilden die Mengen $\tilde{A}_j$ eine endliche disjunkte Zerlegung von X in meßbare Mengen. Setze $\tilde{\alpha}_j = \sum_{i \in I_j} \alpha_i$. Dann ist $f(x) = \tilde{\alpha}_j$ auf $\tilde{A}_j$, also $f = \sum_I \tilde{\alpha}_j 1_{\tilde{A}_j}$. Dann gilt für jedes $a \in \mathbb{R}$

$$\{f < \alpha\} = \bigcup_{\substack{j=1 \\ \tilde{\alpha}_j < a}}^{m} \tilde{A}_j \in \mathcal{F}$$

also ist f meßbar. ∎

Satz 29 *Zu jeder meßbaren Funktion $f : X \longrightarrow [0, \infty]$ gibt es eine Folge (f_n) von meßbaren Treppenfunktionen mit $f_n \uparrow f$, also $f_1 \leq f_2 \leq \ldots$ und $f = \sup f_n$.*

Beweis. Sei

$$f_n(x) = \sum_{k=0}^{n2^n - 1} k 2^{-n} \mathbf{1}_{\{k2^{-n} \leq f < (k+1)2^{-n}\}} + n 1_{\{f \geq n\}}$$

d.h. falls $f(x) \leq n$, so ist $f_n(x)$ ist das größte Element des Gitters $2^{-n}\mathbb{N}_0$, welches links von $f(x)$ liegt, und es ist $f_n(x) = n$ falls $f(x) \geq n$. Weil diese Gitter sukzessive ineinander enthalten sind, ist $f_n(x) \leq f_{n+1}(x)$. Außerdem ist $f_n \leq f$. Schließlich ist $f(x) - f_n(x) \leq 2^{-n}$, falls $f(x) \leq 2^n \cdot n \cdot 2^{-n} = n$ und $f_n(x) = n \uparrow \infty = f(x)$ falls $f(x) = \infty$. Daher konvergiert (f_n) punktweise gegen f. ∎

Kapitel 6

Das Integral bezüglich eines Maßes

6.1 Das Integral für Treppenfunktionen.

Definition 13 *Wir setzen*

$$\int f d\mu := \int f(x) d\mu(x) := \sum_{i=1}^{n} \alpha_i \mu(A_i)$$

wenn $f = \sum_{i=1}^{n} \alpha_i 1_{A_i}$ *eine meßbare Treppenfunktion ist mit disjunkten Mengen* $A_i \in \mathcal{F}$.

Hierbei verwenden wir die Konvention $0 \cdot \infty = 0$ und ohne Einschränkung der Allgemeinheit können wir $X = \bigcup_i A_i$ voraussetzen, denn der Koeffizient $\alpha_i = 0$ ist zulässig.
Bemerkung: Das Integral ist unabhängig von der Zerlegung: Sei $f = \sum_{j=1}^{m} \beta_j \mathbf{1}_{B_j}$ eine andere Darstellung der Funktion f mit disjunkten $B_j \in \mathcal{F}$. Dann ist $\alpha_i = \beta_j$ für alle Paare (i, j) mit $A_i \cap B_j \neq \emptyset$, denn diese Koeffizienten geben beide den Wert an, den f auf diesem Durchschnitt annimmt. Daher ist wegen der Additivität von μ

$$\sum_{i=1}^{n} \alpha_i \mu(A_i) = \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_i \mu(A_i \cap B_j) = \sum_{j=1}^{m} \sum_{i=1}^{n} \beta_j \mu(A_i \cap B_j) = \sum_{j=1}^{m} \beta_j \mu(B_j).$$

Satz 30 *Das so definierte Integral ist auf den nicht negativen Treppenfunktionen*
a) additiv, dh. $\int f + g \; d\mu = \int f \; d\mu + \int g \; d\mu$
b) positiv homogen, dh. $\int a \; f \; d\mu = a \; \int f \; d\mu$ *für* $a \geq 0$
c) isoton, dh. aus $f \leq g$ *folgt* $\int f \; d\mu \leq \int g \; d\mu$.
d) nach oben σ*-stetig, dh. aus* $f_n \uparrow f$ *folgt* $\int f_n \; d\mu \uparrow \int f \; d\mu$.

Beweis. Die Beweise für a)-c) folgen aus der Tatsache, daß es jeweils für alle beteiligten Funktionen eine gemeinsame Zerlegung in meßbare Mengen gibt, auf denen diese Funktionen konstant sind.

d) Sei $f = \sum_{j=1}^{m} \alpha_j 1_{A_j}$. Wir betrachten die Mengen $B_j^{n,k} = \{x \in A_j : f_n(x) > \alpha_j - \frac{1}{k}\}$. Für festes k gilt $B_j^{n,k} \uparrow A_j$ für $n \to \infty$ denn jedes $x \in A_j$ liegt in $x \in B_j^{n,k}$ für schließlich alle n. Ferner ist $f_n \geq \sum_j (\alpha_j - \frac{1}{k})\, 1_{B_j^{n,k}}$. Also erhalten wir

$$\lim_n \int f_n \, d\mu \geq \lim_n \sum (\alpha_j - \frac{1}{k})\, \mu(B_j^{n,k}) = \sum_{j=1}^{m} (\alpha_j - \frac{1}{k})\, \mu(A_j).$$

Mit $k \to \infty$ erhalten wir $\lim_n \int f_n \, d\mu \geq \sum \alpha_j \, \mu(A_j) = \int f \, dx$. Die umgekehrte Ungleichung folgt aus der Isotonie. ■

6.2 Das Integral für positive meßbare Funktionen

Definition 14 *Für nicht negative meßbare Funktionen definieren wir*

$$\int f \, d\mu = \sup\{\int g \, d\mu : g \leq f, g \textit{ ist Treppenfunktion } \}.$$

Lemma 7 *Sei (f_n) irgendeine Folge meßbarer Treppenfunktionen mit $f_n \uparrow f$. Dann gilt $\int f_n \, d\mu \uparrow \int f \, d\mu$.*

Beweis. Dieser und der folgende Satz sind ganz analog zu entsprechenden Aussagen für fallende Folgen bei der Definition der Volumenfunktion für kompakte Mengen. Offensichtlich folgt aus der Isotonie des Integrals für Treppenfunktionen, daß $\int f \, d\mu \geq \lim \int f_n \, d\mu$. Für die umgekehrte Ungleichung sei g eine Treppenfunktion mit $g \leq f$. Dann gilt $\min(f_n, g) \uparrow \min(f, g) = g$ und daher nach dem vorigen Satz auch $\lim_n \int f_n \, d\mu \geq \int g \, d\mu$. Aus der Definition von $\int f \, d\mu$ ergibt sich die gewünschte Ungleichung. ■

Wieder gilt, daß das Integral für nichtnegative Funktionen additiv, positiv homogen und isoton ist. Hier hat die σ-Stetigkeit nach oben einen eigenen Namen.

Satz 31 (monotone Konvergenz). *Sei (f_n) eine monoton wachsende Folge meßbarer nicht negativer Funktionen mit $f_n \uparrow f$. Dann gilt*

$$\int f \, d\mu = \lim_{n\to\infty} \int f_n \, d\mu$$

Beweis. Für jedes $n \in \mathbb{N}$ existiert eine Folge $(f_n^k)_{k\in\mathbb{N}}$ von Treppenfunktionen mit $f_n^k \uparrow_{k\to\infty} f_n$. Wir betrachten die Funktionen $g_n = \max(f_n^1, \ldots, f_n^n)$. Dann gilt $\sup g_n = \sup f_n = f$, und $g_1 \leq g_2 \leq \ldots$. Also $g_n \uparrow f$. Ferner ist $g_n \leq f_n$. Weil die g_n Treppenfunktionen sind, gilt nach dem vorigen Lemma

$$\int f \, d\mu = \lim_{n\to\infty} \int g_n \, d\mu \leq \lim_{n\to\infty} \int f_n \, d\mu.$$

Die umgekehrte Ungleichung $\int f \, d\mu \geq \lim_{n\to\infty} \int f_n \, d\mu$ folgt wieder aus der Isotonie des Integrals. ■

Satz 32 (Lemma von Fatou). *Sei (f_n) eine Folge nicht-negativer meßbarer Funktionen. Dann gilt*

$$\int \liminf_{n\to\infty} f_n \, d\mu \leq \liminf_{n\to\infty} \int f_n \, d\mu.$$

Beweis. Wir betrachten die Funktion $g_m = \inf_{n\geq m} f_n$. Dann ist $g_m \geq 0$ und $g_m \uparrow \sup g_m = \liminf f_n$. Andererseits ist $g_m \leq f_n$ für alle $n \geq m$, also $\int g_m\, d\mu \leq \inf_{n\geq m} \int f_n\, d\mu$. Zusammen gilt nach dem Satz von der monotonen Konvergenz

$$\begin{aligned}\int \liminf_{n\to\infty} f_n \; d\mu &= \int \sup_m g_m\, d\mu = \sup_m \int g_m\, d\mu \\ &\leq \sup_m \inf_{n\geq m} \int f_n \; d\mu \liminf_{n\to\infty} \int f_n \; d\mu.\end{aligned}$$

■

Bemerkung. Aus dem Satz über monotone Konvergenz und der Additivität folgt, daß für jede nichtnegative meßbare Funktion f die Abbildung $f\mu : \mathcal{F} \to [0,\infty]$, die durch

$$f\mu(A) := \int_A f \; d\mu := \int \mathbf{1}_A f \; d\mu$$

definiert ist, ein Maß ist. In der Tat, für jede Folge (A_n) von paarweise disjunkten Mengen mit Vereinigung A gilt

$$f\mu(A) = \int (\sum_{n=1}^{\infty} \mathbf{1}_{A_n}) f \; d\mu = \sum_{n=1}^{\infty} \int \mathbf{1}_{A_n} f \; d\mu = \sum_{n=1}^{\infty} f\mu(A_n).$$

6.3 Integrierbare Funktionen

Wir lassen jetzt auch Funktionen mit verschiedenem Vorzeichen zu. Jede Funktion $f : X \to [-\infty,\infty]$ besitzt die Zerlegung $f = f^+ - f^-$, wobei $f^+ = \max(f,0)$ und $f^- = \max(-f,0)$. Diese beiden Funktionen sind meßbar, wenn f meßbar ist, denn das Maximum und das Minimum meßbarer Funktionen ist meßbar.

Definition 15 *Eine meßbare Funktion $f : X \to [-\infty,\infty]$ heißt* **μ-integrierbar,** **μ-integrabel***, wenn $\int \; |f| \; d\mu < \infty$ ist. Für eine μ-integrierbare Funktion definiert man das Integral durch*

$$\int \; f \; d\mu := \int f^+ \; d\mu - \int f^- \; d\mu.$$

Der Vektorraum der endlich-wertigen μ-integrierbaren Funktionen wird mit $\mathcal{L}^1(\mu)$ bezeichnet.

Bemerkungen 1. Wegen $\max(f,-f) = \max(f^+,f^-) = |f| = f^+ + f^-$ sind die beiden Integrale auf der rechten Seite endlich, also ist die Differenz sinnvoll. Ferner gilt $\int f^+ \; d\mu - \int f^- \; d\mu \leq \int f^+ \; d\mu + \int f^- \; d\mu$ und daher die nützliche Abschätzung

$$|\int f \; d\mu| \leq \int |f| \; d\mu.$$

2. Die doppelte Verwendung des Symbols $\mathcal{L}$ bei den Lebesgue-meßbaren Mengen und bei den integrierbaren Funktionen beruht darauf, daß man in beiden Fällen sich auf den Namen Lebesgue beziehen will. Im Fall $(X,\mathcal{F},\mu) = (\mathbb{R}^m,\mathcal{L}^m,\lambda^m)$

spricht man von **Lebesgue-integrierbaren** Funktionen und schreibt auch einfach $\mathcal{L}^1(\mathbb{R}^m)$ und $\int f(x)\ dx$ statt $\mathcal{L}^1(\lambda^m)$ und $\int f(x)\ d\lambda^m(x)$. Im Spezialfall $m = 1$ erweitert dieser neue Integralbegriff insbesondere das Integral für Regelfunktionen auf kompakten Intervallen. Allerdings sind die im ersten Semester eingeführten uneigentlichen Integrale nicht immer auch Lebesgue-Integrale, z.B. ist das Integral in der Formel $\int_0^\infty \frac{\sin x}{x}\ dx = \pi/2$ kein Lebesgue-Integral, denn der Integrand ist wegen $\int_0^\infty |\frac{\sin x}{x}|\ dx = \infty$ nicht λ^1-integrierbar.
3. Wenn f eine integrierbare Funktion ist, dann gilt für jedes $a > 0$ die sogenannte Markov-sche Ungleichung

$$\mu\{|f| \geq a\} \leq \frac{1}{a} \int |f|\ d\mu < \infty.$$

In der Tat ist $|f| \geq a\mathbf{1}_{\{|f|\geq a\}}$ und daher wegen der Isotonie des Integrals $\int |f|\ d\mu \geq \int a\mathbf{1}_{\{|f|\geq a\}}\ d\mu = a\mu\{|f| \geq a\}$. Insbesondere, wenn man a gegen $+\infty$ gehen läßt, erhält man $\mu\{|f| = \infty\} = 0$.
4. Man sagt, eine Eigenschaft $\mathcal{P}$ von Punkten von X ist **μ-fast überall** erfüllt, wenn die Menge der Punkte, die diese Eigenschaft nicht haben, in einer μ-Nullmenge enthalten ist, z.B. haben wir eben bewiesen, daß eine μ-integrierbare Funktion f μ-fast überall endlich ist.
5. Analog folgt aus der Markovschen Ungleichung auch, daß jede meßbare Funktion f mit $\int |f|\ d\mu = 0$ sogar μ-fast überall verschwindet, denn es gilt $\{|f| > 1/n\} \uparrow \{|f| > 0\}$ und daher

$$\mu(\{|f| > 0\}) = \lim_n \mu(\{|f| > 1/n\}) \leq \lim_n n \int |f| d\mu = 0.$$

6. Zwei meßbare Funktionen f, g die μ-fast überall gleich sind, haben das gleiche Integral: Sei nämlich $N = \{f \neq g\}$. Dann gilt $\int f\ d\mu = \int \mathbf{1}_{N^c} f\ d\mu = \int \mathbf{1}_{N^c} g\ d\mu = \int g\ d\mu$, wie man leicht sieht, indem man alle approximierenden Treppenfunktionen ebenfalls mit $\mathbf{1}_{N^c}$ multipliziert, was deren Integral nicht ändert, weil $\mu(A) = \mu(A \cap N^c) = \mu(A \backslash N)$ gilt für alle A in der σ-Algebra $\mathcal{F}$. Aus diesem Grund identifiziert man häufig zwei meßbare Funktionen, die μ-f.ü. gleich sind. Man nennt zwei solche Funktionen auch **äquivalent** modulo μ. Insbesondere ist jede integrierbare Funktion μ äquivalent zu der endlichwertigen Funktion $f\mathbf{1}_{\{|f|<\infty\}} \in \mathcal{L}^1(\mu)$.

Wieder gilt

Satz 33 *Auf dem Raum $\mathcal{L}^1(\mu)$ ist das Integral linear und isoton.*

Beweis. 1. Es ist $\int af\ d\mu = a \int f\ d\mu$ für alle $a \in \mathbb{R}$. Für $a \geq 0$ folgt dies unmittelbar aus der entsprechenden Eigenschaft im Fall $f \geq 0$. Für $a < 0$ beachte $(af)^+ = |a| f^-$ und $(af)^- = |a| f^+$, also ist

$$\begin{aligned} \int af\ d\mu &= \int (af)^+\ d\mu - \int (af)^-\ d\mu = \int |a| f^-\ d\mu - \int |a| f^+\ d\mu \\ &= -|a| (\int f^+\ d\mu - \int f^-\ d\mu) = a \int f\ d\mu. \end{aligned}$$

2. Sind $f, g \in \mathcal{L}^1(\mu)$, so ist $(f+g)^+ - (f+g)^- = f + g = f^+ + g^+ - f^- - g^-$ oder $(f+g)^+ + f^- + g^- = (f+g)^- + f^+ + g^+$. Aus der Additivität des Integrals

für nichtnegative Funktionen und anschließendes Subtrahieren erhält man

$$\int f + g \; d\mu = \int f \; d\mu + \int g \; d\mu.$$

3. Ist $f \leq g$, so ist $\int g \; d\mu = \int f \; d\mu + \int g - f \; d\mu \geq \int f \; d\mu$ wegen $g - f \geq 0$. ■

Bemerkung. 1. Jede meßbare Funktion f zu der eine integrierbare Funktion h existiert mit $|f| \leq h$ μ- f.ü. ist integrierbar: Denn sei $N = \{|f| > h\}$. Dann ist offenbar f μ-f.ü. gleich der Funktion $\tilde{f} f \mathbf{1}_{N^c}$ und es gilt $|\tilde{f}| \leq h$ und daher $\int |f| \; d\mu = \int |\tilde{f}| \; d\mu \leq \int h \; d\mu < \infty$.
2. Wir betrachten über der Menge $X = \mathbb{N}$ die Potenzmenge als σ-Algebra. Dann sind alle Funktionen mit Definitionsbereich $\mathbb{N}$ meßbar. Das Zählmaß μ ordnet jeder Menge A die Anzahl ihrer Elemente zu. Dann ist leicht zu sehen, daß für jedes $f : \mathbb{N} \to [0, \infty]$ gilt $\int f \; d\mu = \sum_{n=1}^{\infty} f(n)$. Eine Funktion $f : \mathbb{N} \to \mathbb{R}$ ist genau dann μ-integrierbar, wenn die zugehörige Reihe absolut konvergiert. Jeder allgemeine Satz über integrierbare Funktionen impliziert damit einen Satz über absolut konvergente Reihen. Im Fall des nächsten Satzes haben wir den entsprechenden Satz über Reihen schon bei der Diskussion der Exponentialfunktion verwendet.

Der folgende Satz ist vielleicht das nützlichste einzelne Ergebnis der (Lebesgueschen) Integrationstheorie.

Satz 34 Satz von der majorisierten Konvergenz. *Sei (f_n) eine Folge integrierbarer Funktionen so daß eine integrierbare Funktion h existiert mit $|f_n| \leq h$ μ-f.ü.. Sei f eine meßbare Funktion, so daß $f(x) = \lim_{n\to\infty} f_n(x)$ μ-f.ü.. Dann ist auch f integrierbar und es gelten*

$$\int f \; d\mu = \lim_{n\to\infty} \int f_n \; d\mu$$

und

$$\lim_{n\to\infty} \int |f - f_n| \; d\mu = 0.$$

Beweis. Die abzählbare Vereinigung N der Nullmengen $\{f_n \to f\}^c$, $\{|f_n| > h\}$ und $\{|h| = \infty\}$ ist wieder eine Nullmenge. Ohne die Integrale oder die Integrierbarkeit der beteiligten Funktionen zu beeinflussen können wir alle Funktionen auf dieser Nullmenge auf den Wert 0 setzen. Es gilt auch $|f| = \lim |f_n| \leq h$, also ist f integrierbar. Wir betrachten zunächst die Hilfsfunktion $g_n = 2h - (f - f_n)$. Dann gilt $\lim_n g_n = 2h$ und $g_n \geq 0$, also nach dem Lemma von Fatou

$$\begin{aligned} \int 2h \; d\mu &= \int \lim_n g_n d\mu \leq \liminf_n \int g_n \; d\mu \\ &= \int 2h \; d\mu - \limsup_n \Big(\int f \; d\mu - \int f_n \; d\mu \Big). \end{aligned}$$

Also ist $\limsup_n \big(\int f \; d\mu - \int f_n \; d\mu \big) \leq 0$. Analog können wir das Lemma von Fatou auf die ebenfalls nichtnegativen Hilfsfunktionen $2h + f - f_n$ anwenden und erhalten auf die gleiche Weise die Abschätzung $\liminf_n \big(\int f \; d\mu - \int f_n \; d\mu \big) \geq 0$. Zusammen ergibt sich die erste der beiden im Satz behaupteten Limes-Aussagen.

Die zweite folgt, wenn wir die erste auf die Funktionenfolge $(|f - f_n|)$ anwenden, die majorisiert wird durch $2h$ und überall gegen 0 konvergiert. ■

Als Folgerung ergibt sich zum Beispiel die folgende Version des Satzes über das Vertauschen von Differentiation und Integration.

Satz 35 *Sei $f : [a,b] \times X \to \mathbb{R}$ eine Funktion, die bei festem $t \in [a,b]$ μ-integrierbar in x und bei festem $x \in X$ stetig differenzierbar in t ist mit Ableitung $f'(t,x)$. Wenn es eine μ-integrierbare Funktion h gibt, derart daß $|f'(t,x)| \leq h(x)$ für alle $x \in X$ und alle $t \in [a,b]$ gilt, dann ist die Funktion $F : t \mapsto \int f(t,x)d\mu(x)$ stetig differenzierbar mit Ableitung $F'(t) = \int f'(t,x)d\mu(x)$.*

Beweis. Sei (t_n) eine gegen einen Punkt t konvergente Folge in $[a,b]$. Betrachte die μ-integrierbare Funktion $g_n(x)\frac{f(t_n,x)-f(t,x)}{t_n-t}$. Nach dem Mittelwertsatz gibt es für jedes x einen Punkt $t_n^*(x) \in [a,b]$ mit $g_n(x)f'(t_n^*(x),x)$. Nach Voraussetzung ist also $|g_n(x) \leq h(x)$. Ferner gilt $g_n(x) \to f'(t,x)$ für alle $x \in X$. Also gilt nach dem Satz über die majorisierte Konvergenz

$$\frac{F(t_n) - F(t)}{t_n - t} = \int g_n(x)\ d\mu \to \int f'(t,x)d\mu.$$

■

Kapitel 7

Der Satz von Fubini

Wir wollen in diesem Abschnitt die Technik der mehrfachen Integrale auf das Lebegue-Integral ausdehnen. Beim Beweis brauchen wir noch einen Hilfsbegriff, der auch sonst oft nützlich ist, wenn man eine Aussage für einige Elemente einer σ-Algebra kennt und auf die ganze σ-Algebra fortsetzen will.

Definition 16 *Ein System $\mathcal{D}$ von Teilmengen einer Menge X, das die drei Eigenschaften*

a) $X \in \mathcal{D}$.

b) Aus $A, B \in \mathcal{D}$ und $A \subset B$ folgt $B \backslash A \in \mathcal{D}$.

c) Ist (A_n) eine Folge von **paarweise disjunkten** *Elementen von $\mathcal{D}$, dann ist auch die Vereinigung $\bigcup_{n=1}^{\infty} A_n$ in $\mathcal{D}$.*

hat, heißt **Dynkin-System** *über X.*

Bemerkung. Dynkin-Systeme sind also so ähnlich gebaut wie σ-Algebren. Insbsesondere ist wieder der Durchschnitt von beliebig vielen Dynkin-Systemen ein Dynkin-System. Der wesentliche Unterschied ist aber, daß in c) die Mengen als paarweise disjunkt angenommen sind. Das führt dazu, daß diese Eigenschaften a)-c) wesentlich leichter nachzuweisen sind, als die einer σ-Algebra. Andererseits gilt der folgende

Satz 36 *Sei $\mathcal{E}$ ein durchschnittsstabiles System von Teilmengen der Menge X, d.h. aus $A, B \in \mathcal{E}$ folge $A \cap B \in \mathcal{E}$. Wenn ein Dynkin-System $\mathcal{D}$ das System $\mathcal{E}$ umfaßt, dann umfaßt $\mathcal{D}$ auch die von $\mathcal{E}$ erzeugte σ-Algebra $\mathcal{F} = \sigma(\mathcal{E})$.*

Beweis. Sei $\mathcal{D}^*$ der Durchschnitt aller Dynkin-Systeme die $\mathcal{E}$ enthalten, d.h. $\mathcal{D}^*$ ist das von $\mathcal{E}$ erzeugte Dynkin-System. Wir wollen $\sigma(\mathcal{E}) \subset \mathcal{D}^*$ zeigen. Dazu gehen wir schrittweise vor.
1. Aus den Eigenschaften a) und b) folgt, daß $\mathcal{D}^*$ stabil ist gegenüber Komplementbildung.
2. Wir zeigen: Für jede Menge $E \in \mathcal{E}$ und jedes $A \in \mathcal{D}^*$ ist auch $A \cap E \in \mathcal{D}^*$. Denn die Menge $\mathcal{D}_E$ aller A mit $A \cap E \in \mathcal{D}^*$ ist erstens selber ein Dynkin-System, weil $\mathcal{D}^*$ eines ist, und enthält zweitens alle Elemente von $\mathcal{E}$, weil $\mathcal{E}$ durchschnittsstabil ist und $\mathcal{E} \subset \mathcal{D}^*$ gilt. Also ist nach Definition von $\mathcal{D}^*$ auch

$\mathcal{D}^* \subset \mathcal{D}_E$ und daraus folgt die Behauptung.
3. Wir zeigen: $\mathcal{D}^*$ ist durchschnittsstabil. Seien $A, B \in \mathcal{D}^*$. Die Menge $\mathcal{D}_A$ aller B mit $A \cap B \in \mathcal{D}^*$ ist ein Dynkinsystem, welches nach dem vorigen Beweisschritt das System $\mathcal{E}$ enthält. Also gilt nach Definition von $\mathcal{D}^*$ wieder $\mathcal{D}^* \subset \mathcal{D}_A$, woraus die Behauptung folgt.
Wegen 1. und 3. ist das Dynkin-System $\mathcal{D}^*$ auch stabil gegenüber endlichen Vereinigungen. Zusammen mit den Eigenschaften b) und c) folgt daraus, daß $\mathcal{D}^*$ sogar abgeschlossen gegenüber abzählbaren Vereinigungen ist, denn jede Vereinigung einer monoton wachsenden Folge von Mengen läßt sich mit Hilfe der Differenz-Operation $\setminus$ in eine disjunkte Vereinigung verwandeln. Wegen a) und 1. ist $\mathcal{D}^*$ damit eine σ-Algebra, die $\mathcal{E}$ umfaßt. Also gilt $\mathcal{F} \subset \mathcal{D}^*$. ∎

Ferner brauchen wir noch einen technischen Begriff

Definition 17 *Ein Maßraum $(X, \mathcal{F}, \mu)$ heißt* **σ-endlich**, *wenn es eine Folge (Q_n) von Elementen von $\mathcal{F}$ gibt mit $\mu(Q_n) < \infty$ so daß $X = \bigcup_{n=1}^{\infty} Q_n$.*

Bemerkung. Dies ist eine milde Endlichkeitsbedingung. Das Lebesgue-Maß λ^m ist offenbar σ-endlich. Aber das Zählmaß auf der Potenzmenge von $\mathbb{R}$ ist nicht σ-endlich.

Der Satz über Mehrfachintegrale lautet wie folgt.

Satz 37 (Fubini) *Seien $(X, \mathcal{F}, \mu)$ und $(Y, \mathcal{G}, \nu)$ zwei σ-endliche Maßräume. Sei $\mathcal{F} \otimes \mathcal{G}$ die kleinste σ-Algebra über der Produktmenge $X \times Y$, die alle Rechteckmengen $F \times G$ mit $F \in \mathcal{F}, G \in \mathcal{G}$ enthält. Dann gelten*
a) Es gibt es genau ein Maß $\mu \otimes \nu$ auf der σ-Algebra $\mathcal{F} \otimes \mathcal{G}$, derart daß

$$\mu \otimes \nu(F \times G) = \mu(F)\nu(G)$$

gilt für alle $F \in \mathcal{F}, G \in \mathcal{G}$.
b) Sei $f : X \times Y \to [0, \infty]$ eine $\mathcal{F} \otimes \mathcal{G}$-meßbare Funktion. Für jedes $x \in X$ sind die Schnitt-Funktionen $f(x, \cdot) : y \mapsto f(x, y)$ $\mathcal{G}$-meßbar und die Funktion $x \mapsto \int_Y f(x, y)\ d\nu(y)$ ist $\mathcal{F}$-meßbar. Analog sind die Funktionen $f(\cdot, y)$ $\mathcal{F}$-meßbar und die Funktion $y \mapsto \int_X f(x, y)\ d\mu(x)$ ist $\mathcal{G}$-meßbar. Ferner gilt

$$\int_{X \times Y} f(x, y)\ d\mu{\otimes}\nu(x, y) = \int_X \int_Y f(x, y)\ d\nu(y)\ d\mu(x) = \int_Y \int_X f(x, y)\ d\mu(x)\ d\nu(y).$$

c) Eine $\mathcal{F} \otimes \mathcal{G}$-meßbare Funktion $f : X \times Y \to \mathbb{R}$ ist genau dann $\mu \otimes \nu$-integrierbar, wenn $\int_X \int_Y |f(x, y)| d\nu(y)\ d\mu(x) < \infty$ ist. In diesem Fall gibt es eine $\mathcal{F} \otimes \mathcal{G}$-meßbare Funktion $\tilde{f}$, die $\mu \otimes \nu$-fast überall mit f übereinstimmt, derart daß für $\tilde{f}$ die Formel aus b) sinnvoll und richtig ist.

Bevor wir den Satz beweisen, notieren wir noch zwei wichtige Folgerungen:

Folgerung 3 (verallgemeinertes Cavalierisches Prinzip) *Für jede Menge $A \in \mathcal{F} \otimes \mathcal{G}$ sei $A_x = \{y \in Y : (x, y) \in A\}$ und $A^y = \{x \in X : (x, y) \in A\}$. Dann gilt*

$$\mu \otimes \nu(A) = \int \nu(A_x)\ d\mu(x) = \int \mu(A^y)\ d\nu(y).$$

Folgerung 4 *Sei $N \in \mathcal{F} \otimes \mathcal{G}$. Die folgenden Aussagen sind äquivalent:*
a) N ist eine $\mu \otimes \nu$-Nullmenge,
b) Für μ-fast alle $x \in X$ gilt $\nu(N_x) = 0$.
c) Für ν-fast alle $y \in Y$ gilt $\mu(N^y) = 0$.

Beweis. 1. Wir nehmen zunächst zur Vereinfachung an, daß $\mu(X) < \infty$ und $\nu(Y) < \infty$. Wir betrachten nun das System $\mathcal{D}$ aller Teilmengen A von $X \times Y$, derart daß die Menge A_x in $\mathcal{G}$ ist für alle $x \in X$ und außerdem die Abbildung $x \mapsto \nu(A_x)$ $\mathcal{F}$-meßbar ist. Weil die Operation $A \mapsto A_x$ vertauschbar ist mit abzählbaren Vereinigungen und der Differenz $\setminus$, und außerdem ν σ-additiv ist und nach unserer Zusatzvoraussetzung nur endliche Werte annimmt, läßt sich leicht verifizieren, daß $\mathcal{D}$ ein Dynkin-System ist. Andererseits enthält $\mathcal{D}$ offenbar das System $\mathcal{E}$ aller Rechteckmengen $F \times G$ mit $F \in \mathcal{F}, G \in \mathcal{G}$. Der Durchschnitt zweier solcher Rechteckmengen ist wieder eine Rechteckmenge, d.h. $\mathcal{E}$ ist stabil gegenüber Durchschnittsbildung. Also ist der Satz zu Beginn des Paragraphen anwendbar und zeigt $\mathcal{F} \otimes \mathcal{G} \subset \mathcal{D}$. Ferner ist auch die Abbildung $\mu \otimes \nu : A \mapsto \int \nu(A_x)\, d\mu(x)$ wegen der σ-Additivität von ν und dem Satz von der monontonen Konvergenz ein Maß; und es gilt offensichtlich $\mu \otimes \nu(F \times G) = \mu(F)\nu(G)$.

2. Analog erhalten wir ein Maß ρ auf $\mathcal{F} \otimes \mathcal{G}$ mit $\rho(A) = \int_Y \mu(A^y)\, d\nu$. Dies Maß ρ erfüllt ebenfalls $\rho(F \times G) = \mu(F)\nu(G)$ für alle $F \times G \in \mathcal{E}$ und das System aller Mengen A für die $\mu \otimes \nu(A) = \rho(A)$ ist, ist wieder ein Dynkin-System. Also gilt $\rho = \mu \otimes \nu$. Dies beweist Teil a) der Behauptung und die Aussage der ersten Folgerung. Die zweite Folgerung ergibt sich aus der ersten, wenn man noch verwendet, daß eine nichtnegative Funktion (hier die Funktion $x \mapsto \nu(A_x)$), deren μ-Integral verschwindet, selber μ-fast überall verschwindet.

3. Sei nun f nicht negativ und $\mathcal{F} \otimes \mathcal{G}$-meßbar. Dann folgt die Aussage von Teil b) für f aus dem eben bewiesenen auf Grund der Additivität des Integrals, falls f eine Treppenfunktion ist. Außerdem enthält die Klasse aller Funktionen, für die die Aussage von b) richtig ist, auf Grund einer mehrfachen Verwendung des Satzes von der monotonen Konvergenz mit jeder monoton wachsenden Folge von Funktionen auch deren Grenzwert. Weil jede nichtnegative meßbare Funktion als monotoner Grenzwert einer Folge von Treppenfunktionen dargestellt werden kann, folgt die Aussage b).

4. Sei nun f eine $\mathcal{F} \otimes \mathcal{G}$-meßbare Funktion mit beliebigem Vorzeichen. Daß f genau dann $\mu \otimes \nu$ integrierbar ist, wenn das in c) angegebene Doppelintegral über $|f|$ endlich ist, ergibt sich, wenn man Teil b) auf die Funktion $|f|$ anwendet. Wenn f $\mu \otimes \nu$-integrierbar ist, dann ist also insbesondere die Funktion $x \mapsto \int |f(x,y)| d\nu(y)$ μ-integrierbar, also gibt es eine μ-Nullmenge $N_1 \in \mathcal{F}$, so daß dies Integral für alle $x \notin N_1$ endlich ist. Für diese x ist dann auch die Funktion $f(x, \cdot)$ ν-integrierbar. Analog gibt es eine ν-Nullmenge $N_2 \in \mathcal{G}$, so daß für alle $y \notin N_2$ das Integral $\int |f(x,y)|\, d\mu(x)$ endlich ist und das Integral $\int_X f(x,y)\, d\mu(x)$ sinnvoll ist. Sei schließlich $N = (N_1 \times Y) \cup (X \times N_2)$ und $\tilde{f} = \mathbf{1}_{N^c} f$. Dann ist $f = \tilde{f}$ außerhalb der $\mu \otimes \nu$-Nullmenge N und für $\tilde{f}$ sind alle beteiligten Integrale definiert und die Formel in b) ergibt sich, indem man die Zerlegung $\tilde{f} = \tilde{f}^+ - \tilde{f}^-$ benutzt.

5. Schließlich lösen wir uns von der eingangs gemachten Endlichkeitsvoraussetzung. Seien (Q_n) und (R_n) zwei Folgen in $\mathcal{F}$ resp. $\mathcal{G}$, so daß $\mu(Q_n) < \infty$, $\nu(R_n) < \infty$ und $Q_n \uparrow X$, $R_n \uparrow Y$ für $n \in \infty$. Sei $\mu_n(A) = \mu(A \cap Q_n)$ und analog $\nu(B) = \nu(B \cap R_n)$. Dann gilt $\mu_n(X) < \infty$ und $\nu_n(Y) < \infty$. Also können wir das bisherige anwenden und erhalten für jedes n ein eindeutiges Maß $\mu_n \otimes \nu_n$, für

welches die Aussagen a)-c) sinngemäß gelten. Diese Maße sind monoton wachsend in n und die durch $\mu\otimes\nu(A) = \lim \mu_n\otimes\nu_n(A)$ definierte Mengenfunktion ist wieder nach dem Satz über monotone Konvergenz ein Maß. Man verifiziert leicht mit monotoner Konvergenz und Approximation durch Treppenfunktionen, daß

$$\int f \; d\mu\otimes\nu = \lim_n \int_{Q_n\times R_n} f \; d\mu\otimes\nu = \lim_n \int f \; d\mu_n\otimes\nu_n$$

gilt für alle $\mathcal{F}\otimes\mathcal{G}$-meßbaren $f \geq 0$. Daraus folgen dann die Behauptungen auch in diesem Fall. ∎

Für den Spezialfall des Lebesgue-Maßes sind noch eine Reihe von Zusatzbemerkungen angebracht.

Bemerkungen 1. Wenn wir $\mathcal{F} = \mathcal{B}(\mathbb{R}^m)$ und $\mathcal{G} = \mathcal{B}(\mathbb{R}^k)$ wählen, so ist $\mathcal{B}(\mathbb{R}^{m+k}) = \mathcal{B}(\mathbb{R}^m)\otimes\mathcal{B}(\mathbb{R}^k)$, denn beide σ-Algebren werden von den Quadern der Form $Q\times R$ erzeugt. Außerdem ist $\lambda^m_{|\mathcal{B}(\mathbb{R}^m)}\otimes\lambda^k_{|\mathcal{B}(\mathbb{R}^k)}(B) = \lambda^{m+k}(B)$ für alle $B\in\mathcal{B}(\mathbb{R}^{m+k})$, wie man zunächst durch Nachprüfen auf dem System dieser Quader, dann mit dem Satz über Dynkin-Systeme auf meßbare Teilmengen von beschränkten Quadern und schließlich durch Anwendung des Satzes über monotone Konvergenz für allgemeine B verifiziert.

2. Insbesondere erhalten wir eine neue Rechtfertigung der Definition des Integrals für stetige Funktionen durch sukzessive eindimensionale Integration. Strenggenommen scheinen wir allerdings hier die bisher noch nicht bewiesene Tatsache zu verwenden, daß unsere ursprüngliche Definition des Integrals für stetige Funktionen mit kompaktem Träger verträglich ist mit der neuen Integraldefinition gemäß des letzten Paragraphen. Aber im eindimensionalen ist diese Tatsache implizit enthalten in der Definition des Integrals stetiger Funktionen auf Intervallen via gleichmäßiger Approximation durch Treppenfunktionen, und in der mehrdimensionalen Situation gibt der Satz von Fubini gerade den noch ausstehenden Nachweis.

3. Dagegen ist $\mathcal{L}^{m+k}$ echt größer als $\mathcal{L}^m\otimes\mathcal{L}^k$. Dies liegt an der Tatsache, daß jede Teilmenge einer λ^{m+k}-Nullmenge, etwa im Fall $m = k = 1$ jede Teilmenge der Diagonale $\{(x,x)\}\subset\mathbb{R}^2$, wieder ein Element von $\mathcal{L}^{m+k}$, bzw. im Beispiel von $\mathcal{L}^2$ liegt. Es ist nicht schwer zu sehen, daß eine solche Menge i.a. aber nicht in der Produkt-σ-Algebra zu liegen braucht. Trotzdem unterscheidet sich jedes Element von $\mathcal{L}^{m+k}$ aber nur auf einer Lebesgue-Nullmenge von einer Borel-Menge, die nach der ersten Bemerkung sogar in der kleineren σ-Algebra $\mathcal{B}(\mathbb{R}^m)\otimes\mathcal{B}(\mathbb{R}^k)$, also erst recht in $\mathcal{L}^m\otimes\mathcal{L}^k$ liegt.

Kapitel 8

Integraltransformation

Dieser Abschnitt ist dem Beweis des folgenden Satzes gewidmet.

Satz 38 (Integraltransformationsformel) *Seien $U, V \subset \mathbb{R}^m$ zwei offene Mengen und $g : U \to V$ eine bijektive stetig differenzierbare Funktion. Dann gilt für jede $\mathcal{L}^m$-meßbare Funktion $f : V \to \mathbb{R}$ die Formel*

$$\int_V f(y)\ dy = \int_U f(g(x))|\det Dg(x)|\ dx$$

in dem Sinn, daß jedes der beiden Integrale genau dann definiert ist, wenn das andere definiert ist.

Wir beginnen mit einigen Vorbereitungen.

Lemma 8 *Für jede Kugel $K = K(x, r)$ in einem metrischen Raum bezeichne mit K^3 die Kugel $K(x, 3r)$. Sei D die Vereinigung von endlich vielen Kugeln. Dann kann man aus diesen Kugeln eine Teilklasse $K_1, \cdots, K_n$ so auswählen, daß die Kugeln $K_1, \cdots, K_n$ disjunkt sind und $D \subset \bigcup_{i=1}^n K_i^3$ ist.*

Beweis. 1. Seien $K(x, r_1)$ und $K(y, r_2)$ zwei Kugeln mit $r_1 \geq r_2$, die nicht disjunkt sind. Dann gilt $K(x_2, r_2) \subset K(x_1, 3r_1)$, denn ist y im Durchschnitt, dann ist $d(x_i, z) < r_i$ für $i = 1, 2$ und nach der Dreiecksungleichung gilt $d(x_1, z) \leq d(x_1, y) + d(y, x_2) + d(x_2, z) \leq 2r_2 + r_1 \leq 3r_1$ für alle $z \in K(x_2, r_2)$.
2. Wir ordnen die ursprünglichen Kugeln an nach absteigendem Radius. Wir wählen die erste Kugel als K_1. Danach gehen wir die Kugeln der Reihe nach durch und wählen eine Kugel aus, falls sie disjunkt sind zu den schon ausgewählten, und sonst nicht. Dann sind die ausgewählten Kugeln offenbar disjunkt, und jede nicht ausgewählte Kugel K hat nichtleeren Durchschnitt mit einer schon ausgewählten $\tilde{K}$, wobei $\tilde{r} \geq r$. Nach Beweisschritt 1. ist daher $K \subset \tilde{K}^3$. Damit ist auch die zweite Eigenschaft bewiesen. ■

Satz 39 (Vitalischer Überdeckungssatz) *Sei $A \in \mathcal{L}^m$. Sei Φ eine Familie von Kugeln, derart daß für jeden Punkt $x \in A$ und jedes $\varepsilon > 0$ eine Kugel $K \in \Phi$ mit Radius $< \varepsilon$ existiert so daß $x \in K$. Dann gibt es eine disjunkte Folge (K_i) von Kugeln aus Φ, die A fast überdecken, d.h. es ist*

$$\lambda^m \left(A \backslash \bigcup_{i=1}^{\infty} K_i \right) = 0.$$

Beweis. Die Idee des Beweises ist zunächst, mit dem vorangegangenen Lemma zu zeigen, daß man jedenfalls einen gewissen Prozentsatz einer beliebigen Menge durch endlich viele Kugeln heraustrennen kann und dann diesen Prozesß zu iterieren.
1. Wir zeigen zunächst: Ist U offen mit $\lambda^m(U) < \infty$, so gibt es endlich viele disjunkte in U enthaltene Kugeln $K_1, \cdots, K_n$ aus Φ, derart daß

$$\lambda^m\left((A \cap U) \setminus \bigcup_{i=1}^n K_i\right) \leq (1 - 3^{-(m+1)})\lambda^m(A \cap U).$$

Schreibe $B = A \cap U$. Wir setzen $\eta = 3^{-(m+1)}\lambda^m(B)$ und wählen eine kompakte Menge $C \subset B$ mit $\lambda^m(C) \geq \lambda^m(B) - \eta$. Sei $\delta > 0$ so daß $C^\delta \subset U$ und $\lambda^m(C^\delta) \leq \lambda^m(C) + \eta$, wobei $C^\delta = \{y : \mathrm{dist}(y, C) < \delta\}$. Weil C kompakt und in A enthalten ist, können wir C durch endlich viele Kugeln aus Φ, die in C^δ enthalten sind, überdecken. Auf diese wenden wir das Lemma an und erhalten disjunkte Kugeln $K_1, \cdots, K_n$ in Φ derart, daß $C \subset \bigcup_{i=1}^n K_i^3$ ist. Dann gilt

$$\lambda^m(\bigcup_{i=1}^n K_i) = \sum_{i=1}^n \lambda^m(K_i) = 3^{-m} \sum_{i=1}^n \lambda^m(K_i^3) \geq 3^{-m}\lambda^m(C),$$

also erhalten wir aus $K_i \subset C^\delta$ und $\lambda^m(C) \geq \lambda^m(B) - \eta$

$$\begin{aligned}
\lambda^m(B \setminus \bigcup_{i=1}^n K_i) &\leq \lambda^m(C^\delta \setminus \bigcup_{i=1}^n K_i) + \eta \leq \lambda^m(C) - \lambda^m(\bigcup_{i=1}^n K_i) + 2\eta \\
&\leq (1 - 3^{-m})\lambda^m(C) + 2\eta \leq (1 - 3^{-(m+1)})\lambda^m(B).
\end{aligned}$$

2. Wir können, ohne die beteiligten Volumina zu ändern, annehmen, daß die Kugeln abgeschlossen sind. Sei U weiter beschränkt. Wähle $A_0 = A \cap U$ und sukzessive disjunkte Kugeln $K_1, K_2, \cdots$ aus Φ und Indizes $n_1 < n_2 < \cdots$ so daß die Menge $A_k = A \cap U \setminus \bigcup_{i=1}^{n_k} K_i$ die Abschätzung $\lambda^m(A_k) < (1 - 3^{-(m+1)})\lambda^m(A_{k-1})$ erfüllt. Dies geht indem man den ersten Beweisschritt auf die Menge A_{k-1} anwendet. Also ergibt sich, daß $\lambda^m(A_k)$ für wachsende k gegen 0 konvergiert. Die Menge $A \cap U \setminus \bigcup_{i=1}^\infty K_i$ ist dann eine Nullmenge.
3. Sei (U_n) eine Abzählung des disjunkten Systems aller offenen Quader der Seitenlänge 1 mit Ecken in $\mathbb{Z}^m$. Für jedes n kann man das schon bewiesene auf die Mengen $A \cap U_n$ anwenden und erhält die Behauptung, wenn man alle beteiligten Kugeln zusammen nimmt, denn die in verschiedenen U_n liegenden Kugeln treffen sich nicht und die Vereinigung der abzählbar vielen Hyperebenen, die von den U_n nicht überdeckt werden, sind zusammen eine Nullmenge. ■

Eine andere Anwendung des „Kugellemma" ist die folgende Aussage

Lemma 9 *Sei $U \subset \mathbb{R}^m$ offen und $g : U \to \mathbb{R}^m$ lokal Lipschitz-stetig, d.h. um jeden Punkt $x \in U$ gebe es eine Kugel K, so daß $\|g(x) - g(y)\| \leq L\|x - y\|$ für eine Lipschitz-Konstante $L < \infty$ und alle $x, y \in K$. Sei $N \subset U$ eine λ^m-Nullmenge. Dann ist auch $g(N)$ eine λ^m-Nullmenge.*

Beweis. Sei $C \subset U$ kompakt. Wir wollen $\lambda^m(g(C \cap N)) = 0$ zeigen. Wähle endlich viele offene Kugeln $B_1, \cdots, B_k$, die C überdecken, mit zugehörigen

Lipschitz-Konstanten L_j. Sei $L = \max\{L_1, \cdots, L_k\}$. Sei $\varepsilon > 0$ gegeben. Wegen der äußeren Regularität des Lebesgue-Maßes gibt es eine offene Menge V mit $N \cap C \subset V$ und $\lambda^m(V) \leq \frac{\varepsilon}{3^m L^m}$. Wir können auch $V \subset \bigcup_{j=1}^k B_j$ annehmen. Es genügt zu zeigen, daß $\lambda^m(g(V)) \leq \varepsilon$. Hierzu sei D eine kompakte Teilmenge von V. Überdecke D durch Kugeln K, die erstens in V enthalten sind und so daß jedes K^3 in einem der B_j enthalten ist. Weil D kompakt ist, genügen endlich viele solche Kugeln. Wegen der Lipschitz-Stetigkeit ist das Bild $g(K^3)$ in einer Kugel von L-fachem Radius enthalten, also ist $\lambda^m(g(K^3)) \leq L^m \lambda^m(K^3) = 3^m L^m \lambda^m(K)$. Seien nun $K_1, \cdots, K_n$ wie in dem Kugellemma. Dann gilt $D \subset \bigcup_{i=1}^n K_i^3$, $\bigcup_{i=1}^n K_i \subset V$ und daher

$$\begin{aligned} \lambda^m(g(D)) &\leq \lambda^m\left(\bigcup_{i=1}^n g(K_i^3)\right) \leq \sum_{i=1}^n \lambda^m(g(K_i^3)) \\ &\leq 3^m L^m \sum_{i=1}^n \lambda^m(K_i) \leq 3^m L^m \lambda^m(V) < \varepsilon. \end{aligned}$$

Wenn D_k eine wachsende Folge von kompakten Teilmengen von V ist, die V ausschöpfen, folgt $g(D_k) \uparrow g(V)$ und daher $\lambda^m(g(V)) = \lim \lambda^m g(D_k) \leq \varepsilon$ und schließlich $\lambda^m(C \cap N) = 0$.
Ein analoge Ausschöpfung der offen Menge U durch eine wachsende Folge von kompakten Mengen C_k liefert $\lambda^m(g(N)) = \lim \lambda^m(g(C_k \cap N)) = 0$. ■

Lemma 10 *Sei U offen und $g\colon U \to \mathbb{R}^m$ differenzierbar an der Stelle $x \in U$. Dann gilt*

$$\limsup_{\delta \to 0} \frac{\lambda^m\big(g(K(x,\delta))\big)}{\lambda^m\big(K(x,\delta)\big)} \leq |\det D(g(x))|.$$

Beweis. Wir bezeichnen mit $K(r)$ die Kugel um den Ursprung mit Radius r. Es gilt $g(x+h) = g(x) + Dg(x)h + r(h)$, wobei $\|r(h)\| \leq \eta(\|h\|)\|h\|$ mit $\lim_{\rho \to 0} \eta(\rho) = 0$. Sei jetzt $\rho_n \downarrow 0$. Dann ist also das Bild der Kugel $K(x, \rho_n)$ unter g enthalten in der Menge

$$g(x) + Dg(x)\left(K(\rho_n)\right) + K(\eta(\rho_n)\rho_n) = g(x) + \rho_n \left(Dg(x)K(1) + K(\eta(\rho_n))\right).$$

Nun ist $Dg(x)K(1) + K(\eta(\rho_n)) \downarrow Dg(x)K(1)$. Aus der σ-Stetigkeit nach unten folgt $\lambda^m\left(Dg(x)K(1) + K(\eta(\rho_n))\right) \downarrow \lambda^m(Dg(x)K(1)) = |\det Dg(x)|\tau_m$ und daher

$$\limsup \frac{\lambda^m(g(K(x,\rho_n))}{\lambda^m(K(x,\rho_n)} \leq \frac{\rho_n^m \lambda^m\left(Dg(x)K(1) + \mathcal{K}(\eta(\rho_n))\right)}{\rho_n^m \tau_m} = |\det Dg(x)|.$$

■

Der wichtigste Schritt im Beweis ist das folgende Lemma. Beachte, daß hier g nicht injektiv sein muß.

Lemma 11 *Sei U eine offene Teilmenge des $\mathbb{R}^m$. Sei $g : U \to \mathbb{R}^m$ eine stetig differenzierbare Funktion. Dann ist $g(U) \in \mathcal{B}(\mathbb{R}^m)$ und es gilt*

$$\lambda^m(g(U)) \leq \int_U |\det Dg(x)|\, dx.$$

Beweis. Wir nehmen zunächst an, daß U beschränkt ist. Aus dem Mittelwertsatz folgt, daß g lokal Lipschitz-stetig ist. Sei $\varepsilon > 0$. Sei Φ das System aller in U enthaltenen Kugeln $K = K(x, r)$, derart daß

$$\frac{\lambda^m(g(K(x,r)))}{\lambda^m(K(x.r))} \leq |\det Dg(x)|(1+\varepsilon)$$

und außerdem

$$|\det Dg(y)| \geq |\det Dg(x)| - \frac{\varepsilon}{\lambda^m(U)}$$

gilt für alle $y \in K(x,r)$. Dann erfüllen Φ und U nach Lemma 10 und der stetigen Differenzierbarkeit von g die Voraussetzung im Vitalischen Überdeckungssatz. Sei $(K_i) = K(x_i, r_i)$ eine Folge von disjunkten offenen Kugeln aus Φ, die U fast überdecken, d.h. daß $N = U \setminus \bigcup_{i=1}^\infty K_i$ eine λ^m-Nullmenge ist. Weil die K_i offen sind, ist $N \cap L$ kompakt für jede kompakte Teilmenge von U. Weil U Vereinigung von abzählbar vielen kompakten Mengen ist, gilt das gleiche auch für $g(U)$ und für N. Beide Mengen sind also Borelsch. Dann gilt $\lambda^m(g(N)) = 0$ nach dem vorigen Lemma und daher

$$\begin{aligned}
\lambda^m(g(U)) &= \lambda^m\left(g(N) \cup \bigcup_{i=1}^\infty g(K_i)\right) \leq \lambda^m(g(N)) + \sum_{i=1}^\infty \lambda^m(g(K_i)) \\
&\leq \sum_{i=1}^\infty |\det Dg(x_i)|(1+\varepsilon)\lambda^m(K_i) \\
&\leq \sum_{i=1}^\infty \int \mathbf{1}_{K_i}(y)(1+\varepsilon)\Big(|\det Dg(y)| + \frac{\varepsilon}{\lambda^m(U)}\Big)\, dy \\
&= (1+\varepsilon)\int_{\bigcup K_i} |\det Dg(y)| + \frac{\varepsilon}{\lambda^m(U)}\, dy \\
&= (1+\varepsilon)\int_U |\det Dg(y)|\, dy + \varepsilon.
\end{aligned}$$

Daraus folgt die behauptete Ungleichung, denn $\varepsilon > 0$ war beliebig.

Falls U unbeschränkt ist, stelle U dar als aufsteigende Vereinigung beschränkter offener Mengen und verwende monotone Konvergenz. ∎

Bemerkung. In der Situation des vorigen Lemmas heißt $A = \{x \in U : |\det Dg(x)| = 0\}$ die Menge der **kritischen Punkte** von g und ihr Bild unter g die Menge der **kritischen Werte**. Aus dem Lemma folgt, daß die Menge der kritischen Werte immer eine Nullmenge ist. Zum Beweis kann man wieder annehmen daß U beschränkt ist. Die Menge $A^\delta = \{x \in U \mid \det Dg(x)| < \delta\}$ ist offen und für ihre Bildmenge gilt dann

$$\lambda^m(g(A)) \leq \lambda^m(g(A^\delta)) \leq \int_U \delta dx = \delta\lambda^m(U),$$

also $\lambda^m g(A) = 0$.

Lemma 12 *Ist im vorigen Lemma g sogar injektiv, so gilt Gleichheit.*

Beweis. Da die Menge der kritischen Werte eine Nullmenge ist, können wir ohne Einschränkung annehmen, daß $Dg(x)$ vollen Rang hat für alle $x \in U$. Also ist die Umkehrabbildung g^{-1} auch stetig differenzierbar.
Wir verschärfen zunächst das lokale Lemma: Es gilt für jeden Punkt $x \in U$ sogar

$$\lim_{\delta \to 0} \frac{\lambda^m(g(K(x,\delta))}{\lambda^m(K(x,\delta)} = |\det Dg(x)|.$$

Denn sei $\eta > 0$ gegeben und $V_\delta = g(K(x,\delta))$. Nach dem Satz über inverse Funktionen ist V_δ offen und nach dem letzten Lemma, angewendet auf g^{-1} : $V_\delta \to K(x,\delta)$ folgt

$$\lambda^m(K(x,\delta)) \leq \int_{V_\delta} |\det Dg^{-1}(y)| \; dy \leq \frac{(1+\eta)}{|\det Dg(x)|} \lambda^m(V_\delta)$$

für hinreichend kleines δ. Daraus folgt die Behauptung.
2. Nun können wir den gleichen Beweis wie für das vorige Lemma wiederholen, wobei wir in der Definition der Menge Φ die Abschätzungen nach oben durch analoge Abschätzungen nach unten ersetzen. ∎

Schließlich können wir den **Beweis des Hauptsatzes** abschließen. Wieder nehmen wir zunächst U und V als beschränkt an. Wir betrachten das System $\mathcal{D}$ aller Elemente D von $\mathcal{L}^m$, für die die Gleichheit

$$\lambda^m(g(D \cap U)) = \int_{D \cap U} |\det Dg(x)| \; dx$$

gilt. Aus dem Satz von der monotonen Konvergenz folgt leicht, daß $\mathcal{D}$ ein Dynkin-System ist. außerdem enthält $\mathcal{D}$ alle offenen Mengen. Daher gilt diese Gleichheit auch für alle Borel-Mengen und, da jede Menge aus $\mathcal{L}^m$ sich nur auf einer Nullmenge von einer Borel-Menge unterscheidet, auch für alle $D \in \mathcal{L}^m$. Sei schließlich $f = \sum_{i=1}^n b_i \mathbf{1}_{B_i}$ eine nicht negative $\mathcal{L}$-meßbare Treppenfunktion. Dann gilt wegen $\mathbf{1}_{B_i} \circ g = \mathbf{1}_{g^{-1}(B_i)}$ auch

$$\begin{aligned}
\int_V f(y) \; dy &= \sum_{i=1}^n b_i \lambda^m(B_i \cap V) \\
&= \int_U \sum_{i=1}^n b_i \mathbf{1}_{B_i} \circ g(x) \left|\det Dg(x)\right| \; dx \\
&= \int_X f(g(x)) \left|\det Dg(x)\right| \; dx.
\end{aligned}$$

Für allgemeines $f \geq 0$ wähle eine Folge (f_n) von Treppenfunktion mit $f_n \uparrow f$. Dann gilt auch $f_n \circ g \uparrow f \circ g$ und die Gleichheit der Integrale folgt mit monotoner Konvergenz. Für integrierbares f mit allgemeinem Vorzeichen folgt nun die Aussage unmittelbar aus der Definition der Integrierbarkeit.

Beispiel. Wir betrachten zweidimensionale Polarkoordinaten. Um die Voraussetzung des Satzes genau zu erfüllen, entfernen wir aus der Ebene die positive reelle Achse. Sei also $U = (0,\infty) \times (0, 2\pi)$ und $g : U \longrightarrow V = \mathbb{R}^2 \backslash \mathbb{R}_+$ definiert durch $g(r,\varphi) = (r\cos\varphi,\; r\sin\varphi)$. Dann ist g bijektiv und stetig differenzierbar und es ist

$$\begin{aligned} |\det Dg(r,\varphi)| &= |\det \begin{pmatrix} \frac{\partial g_1}{\partial r} & \frac{\partial g_1}{\partial \varphi} \\ \frac{\partial g_2}{\partial r} & \frac{\partial g_2}{\partial \varphi} \end{pmatrix}| \\ &= |\det \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix}| = r. \end{aligned}$$

Weil $\mathbb{R}^2 - V$ eine λ^2-Nullmenge ist, gilt

$$\int_{\mathbb{R}^2} f(y)dy = \int_V f(y)dy = \int_0^\infty \int_0^{2\pi} rf(r\cos\varphi,\ r\sin\varphi)dy\ dr$$

für jedes $f \in \mathcal{L}^1(\mathbb{R}^2)$.

Wir verwenden diese Formel zur Berechnung des eindimensionalen Integrals $I = \int_{-\infty}^{+\infty} e^{-\frac{s^2}{2}} ds$. Mit $f(y_1, y_2) = e^{\frac{-y_1^2+y_2^2}{2}}$ gilt $f(g(r,\varphi)) = e^{\frac{-r^2}{2}}$ und damit dem Satz von Fubini

$$\begin{aligned} I^2 &= \int e^{\frac{-y_1^2}{2}} dy_1 \int e^{\frac{-y_2^2}{2}} dy_2 = \int_{\mathbb{R}^2} f(y)dy \\ &= \int_0^\infty \int_0^{2\pi} r\ e^{-r^2} d\varphi dr = 2\pi \int_0^\infty r\ e^{\frac{-r^2}{2}} dr = 2\pi. \end{aligned}$$

Also ist $I = \sqrt{2\pi}$.

Kapitel 9

Anhang

Der Kern des Integrationsteils dieses Skripts besteht in der Einführung des Lebesgue-Maßes und des Lebesgue-Integrals mit den entsprechenden Konvergenzsätzen, dem Satz von Fubini und der Integraltransformationsformel. Wie eingangs erwähnt, sind in anderen Einführungstexten die Beweis-Methoden etwas anders. Unser Zugang erweist sich bei weiterer Ausdehnung der Theorie als besonders geschmeidig. Dies soll in diesem Anhang an einigen weiteren Sätzen illustriert werden, die in der Vorlesung nicht mehr besprochen wurden.

9.1 Der Fortsetzungssatz von Carathéodory

Das klassische Hilfsmittel bei der Konstruktion von Maßen ist der Maßfortsetzungssatzes von Carathéodory. Wir wollen hier zeigen, daß er relativ einfach aus dem Fortsetzungssatz 26 folgt. Wir übernehmen die Formulierung von Satz 40 aus dem Lehrbuch von H. Bauer.

Definition 18 *Ein System $\mathcal{R}$ von Teilmengen einer Menge X heißt* **Ring**, *wenn es folgende Eigenschaften besitzt.*
a) $\emptyset \in \mathcal{R}$,
b) $A \cup B \in \mathcal{R}$ *für alle* $A, B \in \mathcal{R}$,
c) $A \backslash B \in \mathcal{R}$ *für alle* $A, B \in \mathcal{R}$.

Die Grundmenge X braucht also nicht zu $\mathcal{R}$ zu gehören. Wenn $X \in \mathcal{R}$ ist, ist der Ring eine Algebra. Das bekannteste Beispiel eines Ringes ist das System der endlichen Vereinigungen aller halboffenen Quader der Form $(a_1, b_1] \times \cdots \times (a_m, b_m]$ im $\mathbb{R}^m$. Die Schwäche dieses Begriffs ist, daß ohne eine solche Rechteckstruktur die Ring-Eigenschaften selten nachweisbar sind. Der Satz 26 von Kysinski ist im wesentlichen eine Variante des Satzes von Carathéodory für etwas allgemeinere Mengensysteme.

Definition 19 *Sei $\mathcal{R}$ ein Ring. Eine Funktion $\mathsf{m} : \mathcal{R} \to [0, \infty]$ heißt ein* **Prämaß**, *wenn $\mathsf{m}(\emptyset) = 0$ und m σ-*additiv *ist, dh. wenn für jede Folge $(A_n)_n$ von disjunkten Elementen von $\mathcal{R}$, deren Vereinigung ebenfalls zu $\mathcal{R}$ gehört, gilt*

$$\mathsf{m}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathsf{m}(A_n).$$

Dagegen heißt m *ein* Inhalt, *wenn* m *nur* endlich additiv *ist.*

Bemerkung Ein Prämaß m unterscheidet sich von einem Maß also nur durch die geringeren Anforderungen an den Definitionsbereich. Insbesondere ist ein Prämaß auf einem Ring $\mathcal{R}$ auch monoton und nach unten σ-stetig, dh. für jede Folge (A_n) von Mengen aus $\mathcal{R}$ mit $A_n \downarrow A, A \in \mathcal{R}$ und $\mathsf{m}(A_1) < \infty$ gilt $\mathsf{m}(A_n) \downarrow \mathsf{m}(A)$. Der Beweis ist der gleiche wie für die entsprechende Aussage für Maße in Abschnitt 4.2.

Der Satz von Carathéodory lautet nun

Satz 40 *Jedes Prämaß* m *auf einem Ring* $\mathcal{R}$ *in* X *kann auf mindestens eine Weise zu einem Maß* μ *auf die von* $\mathcal{R}$ *erzeugte* σ*-Algebra* $\sigma(\mathcal{R})$ *fortgesetzt werden.*

Beweis. Wir betrachten das Mengensytem $\mathcal{K}$ aller abzählbaren Durchschnitte von Elementen von $\mathcal{R}$. Auf $\mathcal{K}$ definiere $\underline{\mathsf{m}}$ durch

$$\underline{\mathsf{m}}(K) = \inf\{\mathsf{m}(A) : K \subset A, A \in \mathcal{R}\}. \tag{9.1}$$

Dann ist $\mathcal{K}$ stabil gegen abzählbare Durchschnitte und $\underline{\mathsf{m}}$ setzt wegen der Monotonie m fort. Wir zeigen, daß das Paar $\mathcal{K}, \underline{\mathsf{m}}$ die Voraussetzung des Fortsetzungssatzes aus Abschnitt 4.3 sinngemäß erfüllt.
Erstens ist zu zeigen, daß auch $\underline{\mathsf{m}}$ nach unten σ-stetig ist. Dies sieht man durch eine Wiederholung des Arguments, das sowohl im Beweis der entsprechenden Aussage für die Volumenfunktion vol_m in 3.1 als auch in ähnlicher Form im Beweis des Satzes über monotone Konvergenz in 6.2 verwendet wurde: Zunächst gilt ähnlich wie dort, daß $\underline{\mathsf{m}}(K) = \lim_n \mathsf{m}(A_n)$ gilt für jede Folge (A_n) in $\mathcal{R}$ mit $A_n \downarrow K$. Sei nun $K_n \in \mathcal{K}$ für jedes n gegeben mit $K_n \downarrow K$. Wähle für jedes n eine fallende Folge $(A^n_m)_m$ in $\mathcal{R}$ mit $A^n_m \downarrow K_n$. Setze $B_n = \bigcap_{m=1}^n A^m_n \in \mathcal{R}$. Dann gilt auch $B_n \downarrow K$ und $\underline{\mathsf{m}}(K) = \lim_n \mathsf{m}(B_n) \geq \lim_n \underline{\mathsf{m}}(K_n)$. Die umgekehrte Ungleichung folgt wieder aus der Monotonie.
Zweitens ist zu zeigen, daß für alle $L, K \in \mathcal{K}$ mit $L \subset K$ gilt

$$\underline{\mathsf{m}}_*(K \backslash L) = \underline{\mathsf{m}}(L) - \underline{\mathsf{m}}(K).$$

Hierzu seien (K_n) und (L_n) zwei Folgen in $\mathcal{R}$ gegeben mit $K_n \downarrow K$ und $L_n \downarrow L$. Wir können ohne Einschränkung der Allgemeinheit annehmen, daß $L_n \subset K_n$ indem wir gegebenfalls L_n durch $L_n \cap K_n$ ersetzen. Für jedes $m \in \mathbb{N}$ betrachte die Menge

$$M_m = \left(\bigcap_{n=m}^{\infty} K_n \right) \backslash L_m = \bigcap_{n=m}^{\infty} (K_n \backslash L_m).$$

Wegen der zweiten Darstellung ist $M_m \in \mathcal{R}$ und es gilt $\underline{\mathsf{m}}(M_m) = \lim \mathsf{m}(K_n) - \mathsf{m}(L_m)$. Wegen der ersten Darstellung ist $M_m \subset K \backslash L$ und es folgt

$$\underline{\mathsf{m}}_*(K \backslash L) \geq \lim_m \underline{\mathsf{m}}(M_m) = \lim_n \mathsf{m}(K_n) - \lim \mathsf{m}(L_m) = \underline{\mathsf{m}}(K) - \underline{\mathsf{m}}(L).$$

Wieder liefert die Monotonie die umgekehrte Ungleichung. Damit ist der Fortsetzungssatz aus 4.3 anwendbar und liefert die Behauptung. ■

Die Eindeutigkeit der Maßfortsetzung gilt unter etwas einfacheren Bedingungen an das Mengensystem, aber unter der leichten Zusatzvoraussetzung der σ-Endlichkeit. Implizit haben wir das entscheidende Argument schon in der Eindeutigkeitsaussage des Satzes von Fubini benutzt. Genauer gilt:

Satz 41 (Eindeutigkeitssatz) *Sei $(X, \mathcal{F})$ ein Meßraum und seien μ, ν zwei Maße auf $\mathcal{F}$, die auf einem durchschnittsstabilen Erzeugendensystem $\mathcal{E}$ der σ-Algebra $\mathcal{F}$ übereinstimmen. Wenn $\mathcal{E}$ eine Folge (A_n) enthält mit $A_n \uparrow X$ und $\mu(A_n) = \nu(A_n) < \infty$, dann stimmen μ und ν auf ganz $\mathcal{F}$ überein.*

Beweis. Wir betrachten zunächst den Fall $\mu(X) = \nu(X) < \infty$. Dann ist sehr leicht zu verifizieren, daß das System $\mathcal{D} = \{A \in \mathcal{F} : \mu(A) = \nu(A)\}$ ein Dynkin-System ist. Also ist nach Satz 36 $\mathcal{D} = \sigma(\mathcal{E}) = \mathcal{F}$, woraus die Behauptung folgt. Im allgemeinen Fall können wir ähnlich wie im Beweis des Satzes von Fubini die Maße μ_n, ν_n einführen durch $\mu_n(B) = \mu(A_n \cap B)$ und $\nu_n(B) = \nu(A_n \cap B)$. Wegen

$$\mu_n(X) = \mu(A_n) = \nu(A_n) = \nu_n(X) < \infty$$

gilt $\mu_n = \nu_n$ für jedes n und damit auch $\mu = \nu$. ■

9.2 Der Rieszsche Darstellungssatz

Der Rieszsche Darstellungssatz ist die klassische Wurzel des „funktionalanalytischen Zugangs zur Integrationstheorie“. Wir formulieren ihn für lokalkompakte metrische Räume, dies ist die am meisten benutzte Form. Er stellt eine Bijektion her zwischen den lokal-endlichen kompakt-regulären Borel-Maßen und positiven linearen Funktionalen auf dem Raum der stetigen Funktionen mit kompaktem Träger. Es muß betont werden, daß dieser in etwas modifizierter Form ebenfalls richtig ist für lokalkompakte Räume, deren Topologie nicht von einer Metrik stammt und sogar in allgemeineren Räumen. Darüber hinaus gibt es ähnliche, aber vollständig Topologie-freie Varianten (Satz von Daniell-Stone). Der folgende Satz ist der Prototyp dieser Aussagen.

Ein metrischer Raum X heißt **lokalkompakt**, wenn jede Umgebung eines Punkts $x \in X$ eine kompakte Umgebung besitzt. Er heißt im Unendlichen abzählbar, wenn er Vereinigung einer Folge kompakter Mengen ist. Ein Maß μ auf $\mathcal{B}(X)$ heißt **Radon-Maß**, wenn $\mu(K) < \infty$ ist für alle kompakten Mengen K und wenn μ **von innen regulär** ist, dh. für alle $B \in \mathcal{B}(X)$ gilt:

$$\mu(B) = \sup\{\mu(K) : K \subset B, K \text{ kompakt}\}.$$

Satz 42 *Sei X ein lokalkompakter metrischer Raum. Dann gibt es zu jedem linearen isotonen Funktional $I : \mathcal{C}_c(X) \to \mathbb{R}$ ein eindeutiges Radon-Maß μ auf $\mathcal{B}(X)$, so daß für alle $f \in \mathcal{C}_c(X)$ gilt*

$$I(f) = \int f \, d\mu.$$

Wir können über weite Strecken unser Vorgehen bei der Konstruktion des Lebesgue-Maßes übernehmen. Nur der Beginn und das Ende des Beweises sind etwas anders:

Lemma 13 *Ist eine Folge (f_n) in $\mathcal{C}_c(X)$ gleichmäßig konvergent gegen f und sind die Mengen supp(f_n) in einer gemeinsamen kompakten Menge M enthalten, so ist $I(f) = \lim I(f_n)$.*

Beweis. Aus der Definition der Lokalkompaktheit von X zusammen mit der Überdeckungseigenschaft kompakter Mengen folgt, daß es zu jeder kompakten Menge M ein $\varepsilon > 0$ existiert, derart daß $\{x \in X : \mathrm{dist}(x, M) \leq \varepsilon\}$ wieder kompakt ist. Genau wie in 15 folgt daraus: Die Indikatorfunktion jeder kompaktem Menge ist monoton fallender punktweiser Grenzwert einer Folge in $\mathcal{C}_c(X)$.
Insbesondere gibt eine Funktion $g \in \mathcal{C}_c(X)$ mit $0 \leq \mathbf{1}_M \leq g$. Dann ist wegen der gleichmäßigen Konvergenz $|f - f_n| \leq \varepsilon_n g$ für eine Nullfolge (ε_n). Aus der Isotonie von I folgt

$$
\begin{aligned}
|I(f) - I(f_n)| &= \max\{I(f) - I(f_n), I(f_n) - I(f)\} = \max\{I(f - f_n), I(f_n)\} \\
&\leq I(|f - f_n|) \leq \varepsilon_n I(g) \longrightarrow_{n\to\infty} 0.
\end{aligned}
$$

■

Beweis des Satzes: Wir bezeichnen wieder mit $\mathcal{K}$ das System aller kompakten Teilmengen und definieren für $K \in \mathcal{K}$

$$\mathsf{m}(K) = \inf\{I(f) : \mathbf{1}_K \leq f, f \in CcX\}.$$

Dies ist also eine direkte Verallgemeinerung der Definition der Volumenfunktion vol_m, wobei nur das Integral durch I ersetzt ist. Dann erfüllt auch die Funktion m alle Voraussetzungen des Fortsetzungssatzes in Abschnitt 26. Die Beweise der zugehörigen Teilaussagen lassen sich wörtlich von den entsprechenden Beweisen für die Volumenfunktion Übernehmen. Also ist nach diesem Satz die Einschränkung der durch

$$\mu_*(A) = \sup\{\mathsf{m}(K) : K \subset A,\ K \in \mathcal{K}\}$$

definierten Mengenfunktion auf die σ-Algebra

$$\mathcal{L} = \{A \subset X : \mathsf{m}(K) = \mu_*(K \cap A) + \mu_*(K \backslash A)\ falls\ K \in \mathcal{K}\}$$

ein Maß, das m fortsetzt.
Das Maß μ ist von innen regulär nach Defintion von μ_* und jede kompakte Menge erhält wegen $\mu(K)\mathsf{m}(K)$ endliches Maß. Also ist μ ein Radon-Maß.
Es bleibt zu zeigen, daß das mit diesem Maß nach Kapitel 6 definierte Integral $\int f\ d\mu$ für alle $f \in \mathcal{C}_c(X)$ mit $I(f)$ übereinstimmt. Hier müssen wir etwas anders argumentieren als beim Lebesgue-Maß. Sei also $f \in \mathcal{C}_c(X)$ gegeben. Sei zunächst $f \geq 0$ und wähle $n \in \mathbb{N}$ so groß, daß $f \leq n$. Wir betrachten die Mengen $A_0 = \mathrm{supp} f$ und $A_k = \{k2^{-n} \leq f\}$ für $k \geq 1$. Sei

$$f_n = \sum_{k=0}^{n2^n-1} k2^{-n}\mathbf{1}_{A_k \backslash A_{k+1}} = 2^{-n}\sum_{k=1}^{n2^n-1} \mathbf{1}_{A_k}.$$

Es ist $f_n(x) \leq f(x) < f_n(x) + 2^{-n}\mathbf{1}_{A_0}(x)$ für alle $x \in X$. Da die Mengen A_k kompakt sind, gibt es für jedes $k \geq 0$ eine Folge $(g_m^k)_m$ in $\mathcal{C}_c(X)$ mit $g_m^k \downarrow \mathbf{1}_{A_k}$. Sei $g_m = 2^{-n}\sum_{k=1}^{n2^n-1} g_m^k$. Dann gilt $g_m \downarrow f_n \leq f$, also konvergiert nach dem

Satz von Dini $\max(g_m, f)$ gleichmäßig gegen f. Also erhalten wir nach dem Lemma

$$\begin{aligned} I(f) &= \lim_m I(\max(g_m, f)) \geq \lim I(g_m) = \lim_m 2^n \sum_{k=1}^{n2^n-1} I(g_m^k) \\ &= 2^{-n} \sum_{k=1}^{n2^n-1} \mu(A_k) = \sum_{k=0}^{n2^n-1} k2^{-n}\mu(A_k \backslash A_{k+1}) = \int f_n \, d\mu \end{aligned}$$

und daher für $n \to \infty$ auch $I(f) \geq \int f \, d\mu$. Andererseits ist $f \leq g_m + 2^{-n} g_m^0$ und damit

$$I(f) \leq \lim_m I(g_m) + 2^{-n} I(g_m^0) = \int f_n \, d\mu + 2^{-n}\mu(A_0)$$

woraus für $n \to \infty$ die umgekehrte Ungleichung $I(f) \leq \int f \, d\mu$ folgt.
Die Eindeutigkeit ergibt sich wie folgt: Aus dem Satz von der majorisierten Konvergenz folgt für jedes Maß μ mit der Eigenschaft des Satzes und jede kompakte Menge K, daß die Zahl $\mu(K)$ eindeutig durch das Funktional I bestimmt ist wegen

$$\mu(K) = \int \mathbf{1}_K \, d\mu = \lim_n \int f_n \, d\mu = \lim_n I(f_n),$$

wobei (f_n) irgendiene Folge in CcX ist mit $f_n \downarrow \mathbf{1}_K$. Wegen der inneren Regularität eines Radon-Maßes ist dann $\mu(B)$ auch eindeutig bestimmt durch I für alle Borel-Mengen B. ■

Beispiel. Sei X irgendeine Menge und $d(x, y) = 1$ für $x \neq y$ und $d(x, x) = 0$. Diese Metrik induziert die **diskrete Topologie**, in der jede Teilmenge von X offen und damit Borelsch ist. Die kompakten Mengen sind dann die endlichen Mengen und das Zählmaß μ, das jeder unendlichen Mengen den wert ∞ und jeder endlichen Menge die Anzahl ihrer Elemente zuordnet, ist das Radon-Maß, welches von dem Funktional $I : f \mapsto \sum_{x \in X} f(x)$ induziert wird.

Folgerung 5 *Sei zusätzlich X Vereinigung einer Folge von kompakten Mengen. Dann ist jedes Borel-Maß μ mit $\mu(K) < \infty$ für alle kompakten Mengen automatisch von innen regulär, also ein Radon-Maß. Ferner ist jedes Radon-Maß auch von außen regulär, d.h. für jedes $B \in \mathcal{B}(X)$ gilt*

$$\mu(B) = \inf\{\mu(U) : U \; offen, B \subset U\}.$$

Beweis. 1. Für den Beweis der ersten Aussage definiere das Funktional I durch $I(f) := \int_X f \, d\mu$, $f \in CcX$. Sei $\overline{\mu}$ das Radon-Maß, welches der Satz liefert. Die beiden Maße μ und $\overline{\mu}$ stimmen auf den kompakten Mengen nach dem Argument am Ende des obigen Beweises überein. Sei (K_n) die Folge der kompakten Mengen in unserer Voraussetzung. Ohne Einschränkung kann sie als aufsteigend gewählt werden. Für jede abgeschlossene Menge A ist $A = \bigcup_n A \cap K_n$ in der von den kompakten Mengen erzeugten σ-Algebra, also erzeugen die kompakten Mengen die Borel-σ-Algebra $\mathcal{B}(X)$. Das System aller kompakten Mengen kann also die Rolle des Erzeugenden-Systems $\mathcal{E}$ im Eindeutigkeitssatz übernehmen und wir erhalten damit $\mu = \overline{\mu}$, also ist μ ein Radon-Maß.

2. Wegen der Lokalkompaktheit können wir zusätzlich noch annehmen, daß für jedes n die Menge K_n im offenen Kern K_{n+1}° von K_{n+1} enthalten ist. Seien jetzt $B \in \mathcal{B}(X)$ und $\varepsilon > 0$ gegeben. Für jedes n setze $B_n = B \cap K_{n+1}^\circ$ und wähle eine kompakte Teilmenge L_n der Menge $K_{n+1}^\circ \backslash B_n$ mit

$$\mu\left((K_{n+1}^\circ \backslash B_n)\backslash L_n\right) < \varepsilon 2^{-n}.$$

Dann ist B_n in der offenen Menge $U_n = K_{n+1}^\circ \backslash L_n$ enthalten und das Maß der Menge

$$U_n \backslash B_n = (K_{n+1}^\circ \backslash B_n)\backslash L_n$$

ist $< \varepsilon 2^{-n}$. Die offene Menge $U = \bigcup_n U_n$ erfüllt dann wegen $B = \bigcup B_n$ auch $\mu(U\backslash B) \leq \sum \mu(U_n \backslash B_n) < \varepsilon \sum_n 2^{-n} = \varepsilon$. ■

9.3 Zum Vitalischen Überdeckungssatz

Der Vitalische Überdeckungssatz, den wir in Kapitel 8 zum Beweis der Integraltransformationsformel benutzt haben, zeigt seine eigentlichen Stärke erst, wenn man sich von der Voraussetzung der stetigen Differenzierbarkeit lösen will.

Der folgende Satz (er wird auch Lebesguescher Differentiationssatz genannt) ist für stetige Funktionen sehr einsichtig: Der Wert einer stetigen Funktion an einer beliebigen Stelle ergibt sich approximativ durch Mittelung der Werte in kleinen Kugeln um diese Stelle. Einfache Sprungfunktionen zeigen, daß diese Aussage i.a. falsch ist. Erstaunlich ist, daß die Ausnahmepunkte, bei denen diese Aussage nicht stimmt, selbst für beliebige integrierbare Funktionen nur eine Nullmenge bilden.

Satz 43 *Sei $f \in \mathcal{L}^1(\lambda^m)$. Dann gilt für λ^m-fast alle $x \in \mathbb{R}^m$: Für jede Folge $(K_n(x))$ von Kugeln, die x enhalten und deren Radius gegen 0 konvergiert, ist*

$$\lim_{n\to\infty} \frac{1}{\lambda^m(K_n(x))} \int_{K_n(x)} f(y)\, dy = f(x).$$

Beweis. Offenbar genügt es den Fall $f \geq 0$ zu betrachten. Wir zeigen zunächst, daß die Funktion $\overline{f}$ mit

$$\overline{f}(x) = \lim_{\delta\to 0} \sup\{\mathbf{1}_K(x) \frac{\int_K f(y)\, dy}{\lambda^m(K)} : K \textit{ Kugel vom Radius} < \delta\}$$

meßbar ist. Aus $x_n \to x$ und $\varepsilon_n \to \varepsilon$ folgt $\mathbf{1}_{K(x_n,\varepsilon_n)}(y) \to \mathbf{1}_{K(x,\varepsilon)}(y)$ für alle y außerhalb des Randes von $K(x,\varepsilon)$. Dieser Rand ist eine Nullmenge, also folgt aus dem Satz von der majorisierten Konvergenz, angewendet auf die Funktionen $\mathbf{1}_{K(x_n,\varepsilon_n)} f$, daß das Integral $\int_{K(x,\varepsilon)} f(y)\, dy$ stetig von x und ε abhängt. Daraus ergibt sich, daß in der Definition von $\overline{f}$ es reicht, das Supremum für festes δ nicht über alle Kugeln vom Radius $< \delta$ zu bilden, sondern nur über die abzählbare Teilklasse derjenigen Kugeln, deren Mittelpunkt außerdem rationale Komponenten hat. Also ist dies Supremum meßbar in x. Zweitens fällt dies Supremum monoton mit δ, daher genügt es, den Grenzwert längs einer beliebigen Nullfolge von δ-s zu berechnen. Daraus folgt die Meßbarkeit von $\overline{f}$.
2. Sei jetzt $r > 0$ gegeben. Wir betrachten die meßbare Menge $A_r = \{x \in \mathbb{R}^m$:

$f(x) < r < \overline{f}(x)\}$ und wollen zeigen, daß A_r eine Nullmenge ist. Wir beweisen dies indirekt und nehmen an, es sei $\lambda^m(A_r) > 0$. Dann ist $\int_{A_r} f(x)\ dx < \int_{A_r} r\ dx = r\lambda^m(A_r)$. Wir können eine offene Menge U finden, die A_r enthält und für die auch noch gilt

$$\int_U f(x)\ dx < r\lambda^m(A_r).$$

Der einfachste Weg, dies zu sehen, besteht darin, die äußere Regularität aus der Folgerung 5 auf das durch $\mu(B) = \int_B f(x)\ dx$ definierte Maß μ anzuwenden.
Wir betrachten nun die Menge Φ aller in U enthaltenen Kugeln K, für die gilt

$$\frac{\int_K f(y)\ dy}{\lambda^m(K)} > r.$$

Nach Definition von A_r ist jeder Punkt $x \in A_r$ in beliebig kleinen solchen Kugeln. Nach dem Vitalischen Überdeckungssatz gibt es also eine disjunkte Folge (K^l) in Φ, deren Vereinigung A_r bis auf eine Nullmenge überdecken. Daher gilt

$$\begin{aligned} r\lambda^m(A_r) &> \int_U f(y)\ dy \geq \int_{\bigcup_l K^l} f(y)\ dy \\ &= \sum_l \int_{K_l} f(y)\ dy \geq r \sum_l \lambda^m(K_l) \geq r\lambda^m(A_r). \end{aligned}$$

Dies ist ein Widerspruch, die Menge A_r muß also eine Nullmenge sein.
Es folgt, daß auch die Menge $\{f < \overline{f}\} = \bigcup_{r\in\mathbb{Q}_+} A_r$ eine Nullmenge ist. Daraus ergibt sich, daß für λ^m-fast alle Punkte x und jede Folge $K_n(x)$ wie im Satz der lim sup der Folge auf der linken Seite der Behauptung $\leq f(x)$ ist. Analog wird die umgekehrte Ungleichung für den lim inf bewiesen. Zusammen folgt die Behauptung des Satzes. ∎

Dies hat eine Anwendung auf sogenannte Dichte-Punkte. Ein Punkt $x \in \mathbb{R}^m$ heißt **Dichtepunkt** der Menge $A \in \mathcal{L}^m$, wenn gilt

$$\lim_{\delta\to 0} \frac{\lambda^m\left(A \cap \mathcal{K}(x,\delta)\right)}{\lambda^m(K(x,\delta))} = 0.$$

Folgerung 6 *Fast alle Punkte einer meßbaren Menge $A \in \mathcal{L}^m$ sind Dichtepunkte von A.*

Beweis. Es genügt, den Fall $\lambda^m(A) < \infty$ zu betrachten. Dann ist die Indikatorfunktion $\mathbf{1}_A$ integrierbar. Mit $f = \mathbf{1}_A$ folgt die Behauptung aus dem vorigen Satz. ∎

Eine andere Anwendung ist auf eine Variante des Fundamentalsatzes der Differential und Integralrechnung:

Satz 44 *Sei $f \in \mathcal{L}^1(\mathbb{R})$ und $F(x) = \int_{-\infty}^x f(s)\ ds$. Dann ist F λ^1-fast überall differenzierbar mit $F'(x) = f(x)$.*

Beweis. Sei $x \in \mathbb{R}$ ein Punkt, für den die Aussage aus dem Satz gilt. Wir behaupten, daß F an der Stelle x differenzierbar ist. Sei hierfür (t_n) eine Nullfolge von positiven Zahlen. Wir wenden die Formel des Satzes auf die Kugeln $K_n(x) = [x, x + t_n]$ an und erhalten

$$\frac{F(x+t_n) - F(x)}{t_n} = \frac{\int_x^{x+t_n} f(x)\ dx}{\lambda^1(K_n(x))} \longrightarrow_{n\to\infty} f(x).$$

Also ist die rechtssetige Ableitung von F gleich $f(x)$. Das gleiche Argument für $K_n(x) = [x - t_n, x]$ liefert die entsprechende Beziehung für die linksseitige Ableitung. ■

Im Kapitel 8 hatten wir vorausgesetzt, daß die Variablen-Transformation durch eine stetig differenzierbare Funktion g gegeben ist. Wir wollen nun diese Voraussetzung abschwächen, weil in einer Reihe von Anwendungen die Funktion nur fast überall differenzierbar ist und die Ableitung nicht notwendig stetig ist. Im Beweis hatten wir verwendet, daß ein stetig differenzierbares g auch (jedenfalls lokal) Lipschitz-stetig ist. Bemerkenswerter Weise gilt hiervon auch eine abgeschwächte Umkehrung:

Satz 45 (Rademacher) *Sei $U \subset \mathbb{R}^m$ offen und $g : U \to \mathbb{R}^n$ Lipschitz-stetig. Dann ist g an λ^m-fast allen Punkten $x \in U$ (total) differenzierbar.*

Wir wollen diesen Satz hier nicht beweisen. Wir verwenden ihn für folgende Variante des Hauptsatzes aus Kapitel 8. Wenn man den Satz von Rademacher nicht verwenden will, muß man eben die Differenzierbarkeit fast überall zusätzlich fordern. Das Ergebnis ist dann immer noch deutlich allgemeiner als Satz 38.

Satz 46 *Seien U, V zwei offene Mengen im $\mathbb{R}^m$. Sei $g : U \to V$ bijektiv und in beiden Richtungen Lipschitz-stetig. Dann gilt auch die Integralstransformationsformel aus Satz 38.*

Beweis. (Skizze) Wir können wie üblich U als beschränkt voraussetzen. Sei $U_0 = \{x \in U : Dg(x)\ existiert\}$. Nach dem Satz von Rademacher ist $\lambda^m(U \backslash U_0) = 0$, insbesondere ist auch $U_0 \in \mathcal{L}^m$. Da die Funktion g Lipschitz-stetig ist, ist die Operatornorm $\|Dg(x)\|$ gleichmäßig auf der Menge U_0 beschränkt. Daraus folgt, daß die Einträge der die linearen Abbildungen $Dg(x)$ darstellenden Funktionalmatrizen und schließlich auch die Determinanten $\det Dg(x)$ beschränkt sind. Also ist die Funktion $\mathbf{1}_{U_0} |\det Dg|$ integrierbar, wir können also den vorigen Satz anwenden und erhalten für fast alle $x \in U$

$$\frac{\int_{K(x,\delta)} |detDg(y)|\ dy}{\tau_m \delta^m} \geq |\det Dg(x)| - \frac{\varepsilon}{\lambda^m(U)}$$

für alle hinreichend kleinen $\delta > 0$. Daher kann man die Bedingung

$$|\det Dg(y)| \ \geq \ |\det Dg(x)| - \frac{\varepsilon}{\lambda^m(U)}$$

im Beweis von Lemma 11 durch diese Ungleichung ersetzen ohne den Rest des Beweises dieses Lemmas zu ändern. Diese Bedingung war aber die einzige Stelle in diesem Lemma, an der die stetige Differenzierbarkeit verwendet wurde, also bleibt es für Lipschitz-Funktionen, die fast überall differenzierbar sind, gültig. Der Beweis des Lemmas 12 bleibt ebenfalls mit diesen Modifikationen richtig, wobei man in dem lokalen Schritt statt des alten Arguments wieder den Lebesgueschen Differentiationssatz einsetzt und beachtet, daß auch die Umkehrabbildung als Lipschitz-stetig vorausgesetzt wird und daher tatsächlich der Beweis der unteren Abschätzung wie im Lemma 11 auf g^{-1} angewendet werden kann.

■

Literaturverzeichnis

[Bau92] H. Bauer. *Maß- und Integrationstheorie, 2nd ed.* de Gruyter, Berlin - New York, 1992.

[Els99] J. Elstrodt. *Maß- und Integrationstheorie.* Springer-Verlag, New York etc., 1999.

[Fora] O. Forster. *Analysis 1.* Vieweg Studium 24, Wiesbaden.

[Forb] O. Forster. *Analysis 2.* Vieweg Studium 36, Wiesbaden.

[Forc] O. Forster. *Analysis 3.* Vieweg Studium 52, Wiesbaden.

[Rud70] W. Rudin. *Reelle und Komplexe Analysis.* Oldenbourg, München, 1970.