Inhaltsverzeichnis

Einführung

Schließlich wird dem Leser vermutlich auffallen, daß auf das Riemann-Integral, diesen ehrwürdigen Gegenstand von Analysisvorlesungen, überhaupt nicht eingegangen wird. Man darf wohl annehmen, daß dieser Begriff, wäre er nicht mit einem so klangvollen Namen verknüpft, schon viel früher übergangen worden wäre; denn bei allem schuldigen Respekt vor dem Genius BERNHARD RIEMANNS ist sich jeder aktive Mathematiker völlig darüber im klaren, daß diese „Theorie" heutzutage in der allgemeinen Maß- und Integrationstheorie bestenfalls die Bedeutung einer halbwegs interessanten Übungsaufgabe besitzt. Nur der sture Konservativismus akademischer Tradition konnte das Riemannsche Integral als vollwertigen Bestandteil der Analysisvorlesung erhalten, lange nachdem es seine historische Bedeutung überlebt hatte.

Das schreibt Jean Dieudonné im 1. Band seiner „Grundzüge der modernen Analysis" (S. 149 der 2. Auflage). Mit dieser Aussage nach 2 oder 3 Semestern Analysis konfrontiert wird sich mancher vielleicht fragen, was an der Riemannschen Integrationstheorie denn so nachteilig ist. Doch zunächst das Positive: Die Einführung des Riemannintegrals scheint absolut natürlich, wirkt anschaulich und ist technisch relativ einfach (deswegen sollte es m.E. auch seinen Platz in der Analysis I behalten). Bekanntlich funktioniert die Sache so:

Sei $f: [a,b] \to \mathbb{R}$ eine beschränkte Funktion. Man zerlegt nun $[a,b]$ in kleine Teilintervalle $a = x_0 < x_1 < \dots < x_n = b$, wählt Zwischenpunkte $x_{i-1} \leq \xi_i \leq x_i$ und betrachtet die Riemannschen Summen

$$(1) \qquad \sum f(\xi_i)\, \ell(J_i),$$

wo $\ell(J_i)$ die Länge des i-ten Intervalls $J_i = [x_{i-1}, x_i]$ ist. f ist R-integrierbar, wenn bei immer feiner werdender Zerlegung die Riemannschen Summen gegen eine Zahl, nämlich $\int_a^b f(x)\,dx$, konvergieren. Insbesondere ist das der Fall, wenn f auf den meisten J_i wenig schwankt, speziell wenn f (gleichmäßig) stetig ist.

Für Funktionen mehrerer Veränderlicher – und als nächstes zeige ich einige Macken des Riemannintegrals auf – ist die Einführung schon lästiger, und die Ergebnisse sind gemessen am Beweisaufwand höchst unbefriedigend (man denke etwa an die Transformationsformel oder die Tatsache, daß nicht einmal stetige Funktionen auf Kompakta R-integrierbar zu sein brauchen). (Deswegen sollte m.E. das Riemannintegral aus der Analysis III auch verschwinden. Das Lebesguesche Integral ist hier eindeutig überlegen; siehe z.B. die Darstellung in Forster, Analysis III.)

Zurück zu den Funktionen auf einem Intervall. Bekanntlich ist nicht jede beschränkte Funktion R-integrierbar, das berühmte Gegenbeispiel ist die Dirichletsche Funktion Δ, die rationalen Zahlen die 1 und irrationalen die 0 zuordnet. Im Lebesgueschen Sinne ist Δ in der Tat integrierbar, und es ist $L-\int_a^b \Delta(x)\,dx = 0$, was plausibel ist, da es ja viel weniger rationale (nämlich nur abzählbar viele) als irrationale Zahlen gibt. An dieser Stelle mag man einwenden, daß es eigentlich recht unerheblich ist, ob man eine vergleichsweise pathologische Funktion wie Δ integrieren kann oder nicht. Die Darstellung

$$\Delta(x) = \lim_{m\to\infty} \lim_{n\to\infty} [\cos(m!\,\pi x)]^{2n} \qquad ^{*)}$$

wirft jedoch ein neues Licht auf diese Frage, weist sie doch Δ als punktweisen Limes von Funktionen aus, die über jeden Zweifel erhaben sind. Hier zeigt sich das eigentliche Manko des R-Integrals: Die punktweise definierte Limesfunktion einer Folge von R-integrierbaren Funktionen braucht nicht

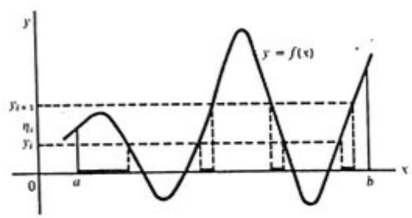
*) Zum Beweis bemerke $\lim_{n\to\infty} [\cos(m!\,\pi x)]^{2n} = \begin{cases} 0 & \text{für } m!x \notin \mathbb{Z} \\ 1 & \text{für } m!x \in \mathbb{Z} \end{cases}$

R-intbar zu sein (selbst wenn sie beschränkt ist). Der Standpunkt, Eigenschaften der Grenzfunktion f einer konvergenten Funktionenfolge (f_n) aus entsprechenden Eigenschaften der f_n herzuleiten, ist jedoch für die moderne Analysis fundamental; man denke z.B. an iterative Verfahren zur Lösung von Differential- oder Integralgleichungen, wo f als $\lim f_n$ definiert ist. Selbst wenn die Grenzfunktion R-integrierbar ist, sind Kriterien, die Grenzübergang und Integration zu vertauschen gestatten, häufig nicht in wünschenswerter Allgemeinheit zu beweisen (insbesondere bei uneigentlichen Integralen).

Lebesgue, portrait by an unknown artist, 1929
By courtesy of the Bibliothèque Nationale, Paris

All diese Schwierigkeiten können bei der Lebesgueschen Integration, die dieser zuerst in seiner Dissertation "Intégrale, longueur, aire" 1902 vorgestellt hatte, überwunden werden. Die Grundidee Lebesgues ist verblüffend einfach; daß sie in der Tat durchführbar ist, ist freilich eine nichttriviale Tatsache.

Die Idee soll am Beispiel einer beschränkten Funktion $f: [a,b] \to \mathbb{R}$ skizziert werden (der Übergang zu unbeschränkten Funktionen stellt sich übrigens als vollkommen unproblematisch heraus). Statt das Urbildintervall in Teilintervalle zu zerlegen, wird nun der Bildbereich zerlegt:



Alsdann sammelt man alle x-Werte, die zu y-Werten in $[y_i, y_{i+1}]$ gehören, ein:

$$E_i = \{ x : y_i \le f(x) \le y_{i+1} \}$$

Es liegt nun nahe, statt der Riemannschen Summen Summen der Form

$$(2) \qquad \sum \eta_i \, \lambda(E_i)$$

zu betrachten, wo $y_i \le \eta_i \le y_{i+1}$ und $\lambda(E_i)$ die „Länge" der Menge E_i ist. Für halbwegs glatte Funktionen wird E_i eine endliche Vereinigung von Intervallen sein (auf der Skizze auf S. 3 sind es deren 4); in diesem Fall ist klar, was unter $\lambda(E_i)$ zu verstehen ist. Im allgemeinen können die Mengen E_i jedoch ziemlich beliebig sein; für $f = \Delta$ kommt etwa $[a,b] \cap \mathbb{Q}$ vor. Das Ziel ist wiederum, die Zerlegungen feiner und feiner zu machen und zu zeigen, daß die Summen vom Typ (2) konvergieren – eben gegen das Lebesguesche Integral von f.

Um mit dieser Methode möglichst viele Funktionen behandeln zu können, ist es also zunächst einmal nötig, möglichst vielen Teilmengen von $\mathbb{R}$ ein Maß, das die Länge von Intervallen verallgemeinert, zuzuordnen (am besten allen, aber das stellt sich leider als unmöglich heraus). Vor dem Problem der Integration steht daher das Maßproblem. In der Tat gelingt es, praktisch jede Teilmenge $E \subset \mathbb{R}$ in diesem Sinn zu messen, so daß die Maßabbildung die Eigenschaften

$$(3) \qquad \lambda([\alpha, \beta]) = \beta - \alpha,$$

$$(4) \qquad \lambda(E \cup F) = \lambda(E) + \lambda(F) \qquad \text{falls } E \text{ und } F \text{ disjunkt,}$$

sogar schärfer

$$(5) \qquad \lambda\Big(\bigcup_{i=1}^{\infty} E_i\Big) = \sum_{i=1}^{\infty} \lambda(E_i) \qquad \text{falls die } E_i \text{ paarweise disjunkt}$$

besitzt. (Die Lebesgue dazu vorgeht, ist auf S. 29f. skizziert.)

Hat man so ein Maß erst einmal zur Hand, kann man für praktisch jede Funktion die approximierende Summe (2) bilden (nämlich sobald nur alle Mengen $E_i = f^{-1}([y_i, y_{i+1}])$ meßbar sind; solche Funktionen heißen meßbar – und praktisch jede Funktion ist meßbar). Es stellt sich dann heraus, daß die Eigenschaft (5) den avisierten Grenzübergang und damit die Definition des L-Integrals gestattet.

Glücklicherweise stimmen für stetige (allgemeiner für R-integrierbare) Funktionen R- und L-Integral überein. Das liegt grob gesagt daran, daß eine Funktion R-integrierbar ist, wenn sie auf vielen Teilintervallen einer Zerlegung von $[a,b]$ wenig schwankt (und umgekehrt), d.h. wenn der Lebesguesche Ansatz des Sammelns aller Argumente mit fast gleichen Funktionswerten zu (endlichen Vereinigungen von) Intervallen führt. Gleichzeitig sieht man die konzeptuelle Schwäche des R-Integrals: Hier bildet man letztendlich auch Mengen der Form $f^{-1}([y_i, y_{i+1}])$, besteht aber darauf, daß es sich mehr oder minder um Intervalle handelt.

Die Klasse der meßbaren Funktionen erweist sich nun als „vollständig", d.h. der punktweise Limes von meßbaren Funktionen ist meßbar. Des weiteren kann man ohne große Mühe weitreichende Konvergenzsätze beweisen und so die Mängel des R-Integrals beheben.

Nachdem das Maß λ konstruiert war, war es für die obige Konstruktion des Integrals vollkommen unerheblich, daß die zu integrierende Funktion auf einem kompakten Intervall definiert war. Jedes Maß – sei es auf einem unbeschränkten Intervall, auf $\mathbb{R}$, auf $\mathbb{R}^n$ oder gar einem abstrakten „Wahrscheinlichkeitsraum" definiert – führt so zu einem Integral.

Im ersten Teil der Vorlesung wird daher das Maßproblem studiert, im möglichst allgemeinen Rahmen Mengen Zahlen zuzuordnen, wobei die σ-Additivität genannte Eigenschaft (5) des Lebesguemaßes von herausragender Bedeutung ist, führt sie doch später direkt zu den Konvergenzsätzen. Der klassische Fall des Lebesguemaßes ist übrigens kaum einfacher als der allgemeinstmögliche, weswegen wir uns gleich diesen vornehmen. Im zweiten Teil geht es dann um die allgemeine Integrationstheorie. Ich hoffe, ihr könnt dabei das Schema der obigen Skizze wiedererkennen; denn es gibt eine Menge technischer Einzelheiten zu berücksichtigen. Der dritte Teil behandelt die Version des Hauptsatzes der Differential- und Integralrechnung, die die Maß- und Integrationstheorie zu bieten hat. Zum Schluß werden die Verbindungen zur klassischen Analysis und einige spezielle Fragen und Anwendungen besprochen.

Der hier vorgestellte Zugang zur Integrationstheorie ist bei weitem nicht der einzig mögliche. Er scheint mir jedoch anschaulicher zu sein als die Methoden, die ohne Maßtheorie direkt zur Integration gelangen. Solche Ansätze findet man z.B. in folgenden Büchern (der Abstraktionsgrad ist hier monoton wachsend):

– Methode von Riesz:
  - Heuser: Analysis II
  - Riesz / Sz. Nagy: Vorlesungen über Funktionalanalysis
  - Weir: Lebesgue Integration and Measure

– Daniell-Integration:
  - Floret: Maß- und Integrationstheorie
  - Weir: General Integration and Measure

– Radonmaße:
  - Pedersen: Analysis Now
  - Dieudonné: Grundlagen der modernen Analysis, Bd. 2
  - Bourbaki: Intégration

I. Grundlagen der Maßtheorie

I.1. Ringe und σ-Algebren

Im folgenden bezeichnet S eine beliebige nicht leere Menge, $\mathcal{P}(S)$ bedeutet die Potenzmenge von S, d.h. die Menge aller Teilmengen von S. (Beachte $\emptyset \in \mathcal{P}(S)$!)

1.1 Definition a) $\mathcal{R} \subset \mathcal{P}(S)$ heißt Ring, falls

(i) $\emptyset \in \mathcal{R}$

(ii) $A, B \in \mathcal{R} \Rightarrow A \setminus B := \{s \in S : s \in A, s \notin B\} \in \mathcal{R}$

(iii) $A_1, \dots, A_n \in \mathcal{R} \Rightarrow \bigcup_{i=1}^{n} A_i \in \mathcal{R}$ ($n \in \mathbb{N}$ beliebig)

b) Gilt sogar

(iii)* $A_1, A_2, \dots \in \mathcal{R} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{R}$,

so heißt $\mathcal{R}$ σ-Ring.

c) Gilt zusätzlich zu (i), (ii), (iii) [bzw. (iii)*]

(iv) $S \in \mathcal{R}$,

so spricht man von einer Algebra [bzw. σ-Algebra].

Der wichtigste dieser Begriffe ist der der σ-Algebra, die übrigen sind mehr technischer Natur. Im Endeffekt werden wir daran interessiert sein, auf einer σ-Algebra definierte „Mengenfunktionen" (die gewissen Mengen Zahlen zuordnen) zu betrachten. Leider ist der „natürliche" Definitionsbereich einer solchen Mengenfunktion i.a. nur ein Ring, und es ist

ein ausgesprochen nichttriviales – für eine befriedigende Theorie von Maß und Integral jedoch unumgängliches – Problem, diesen Definitionsbereich zu erweitern. Dieses Problem wird in I.3.6 gelöst. Die nun folgenden Beispiele sollen der Veranschaulichung der eingeführten Begriffe dienen:

1.2 Beispiele a) $\{\emptyset, S\}$ und $\mathcal{P}(S)$ sind stets σ-Algebren.

b) $\{A \subset S : A$ ist eine endliche Menge$\}$ ist ein Ring.

c) $\{A \subset S : A$ oder $S \setminus A$ ist eine endliche Menge$\}$ ist eine Algebra.

d) $\{A \subset S : A$ ist höchstens abzählbar$\}$ ist ein σ-Ring.

e) $\{A \subset S : A$ oder $S \setminus A$ ist höchstens abzählbar$\}$ ist eine σ-Algebra.

Diese Beispiele wirken (zu Recht) etwas künstlich (Beispiel e) hat im folgenden für die Konstruktion einiger Gegenbeispiele eine gewisse Bedeutung); das folgende Beispiel ist für den Aufbau der Lebesgueschen Integrationstheorie fundamental:

f) Sei $S = \mathbb{R}$ und $\mathcal{F}^1$ die Menge aller endlichen Vereinigungen von (links) halboffenen Intervallen:

$$\mathcal{F}^1 = \left\{ \bigcup_{i=1}^{n}]a_i, b_i] : n \in \mathbb{N},\ a_i, b_i \in \mathbb{R},\ a_i \le b_i \ (i = 1, \dots, n) \right\}$$

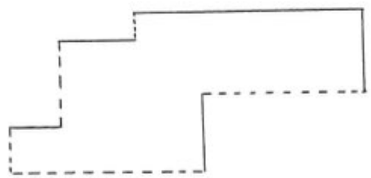
Analog definiert man im Fall $S = \mathbb{R}^k$ das Mengensystem $\mathcal{F}^k$; dabei ist für $a = (\alpha_1, \dots, \alpha_k)$, $b = (\beta_1, \dots, \beta_k)$ mit $\alpha_j \le \beta_j$ (alle j) das halboffene Intervall $]a, b]$ durch

$$]\alpha_1, \beta_1] \times \dots \times]\alpha_k, \beta_k]$$

erklärt.

$\mathcal{F}^k$ ist ein Ring, der der Ring der k-dimensionalen Figuren genannt wird. (Hätte man mit offenen Intervallen operiert, so wäre Bedingung (ii) aus Def. 1.1 verletzt!)

Eine typische zweidimensionale Figur sieht so aus:



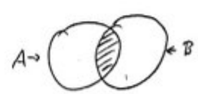
In diesem Beispiel kann $n = 3$ gewählt werden. (Beachte, daß die Zerlegung einer Figur in Intervalle nicht eindeutig ist!)

Eine einfache Eigenschaft von Ringen ist ihre Durchschnittsstabilität:

<u>1.3 Lemma</u> a) Sei $\mathcal{R}$ ein Ring und $A, B \in \mathcal{R}$. Dann ist auch $A \cap B \in \mathcal{R}$.

b) Sei $\mathcal{R}$ ein σ-Ring und $A_1, A_2, \dots \in \mathcal{R}$. Dann ist auch $\bigcap_{i=1}^{\infty} A_i \in \mathcal{R}$.

Beweis: a) $A \cap B = A \setminus (A \setminus B)$:



b) $\bigcap_{i=1}^{\infty} A_i = A_1 \setminus \bigcup_{i=2}^{\infty} (A_1 \setminus A_i)$ (nachrechnen!)

Das Nachrechnen der letzten Identität funktioniert am besten mit dem Komplementbegriff

$$A^c := S \setminus A ,$$

denn damit wird $A \setminus B = A \cap B^c$,

und die de Morganschen Regeln

$$\left(\bigcap_i A_i\right)^c = \bigcup_i A_i^c$$
$$\left(\bigcup_i A_i\right)^c = \bigcap_i A_i^c$$
$$B \cap \left(\bigcup_i A_i\right) = \bigcup_i (B \cap A_i)$$
$$B \cup \left(\bigcap_i A_i\right) = \bigcap_i (B \cup A_i)$$

können angewandt werden.

Wir bemerken noch, daß bei der Definition der Algebra [bzw. σ-Algebra] die Bedingung (ii) in Def. 1.1 durch

$$(\widetilde{ii}) \qquad A \in \mathcal{R} \Rightarrow A^c \in \mathcal{R}$$

ersetzt werden kann. (Wegen (iv) ist $(\widetilde{ii})$ Spezialfall von (ii). Umgekehrt ist $A \setminus B = A \cap B^c \in \mathcal{R}$ für $A, B \in \mathcal{R}$ unter der Bedingung $(\widetilde{ii})$.) *)

Weiterhin sei angemerkt, daß ein Ring seinen Namen zu Recht trägt: Führt man nämlich auf einem Mengensystem $\mathcal{R} \subset \mathcal{P}(S)$ eine „Addition" durch die Vorschrift

$$(A, B) \mapsto A \,\Delta\, B := A \setminus B \cup B \setminus A$$

sowie eine „Multiplikation"

$$(A, B) \mapsto A \cap B$$

ein, so zeigt sich, daß ein Ring im Sinn von Def. 1.1 ebenfalls

) Um zu zeigen, daß ein Mengensystem eine σ-Algebra ist, werden wir demnächst daher (i), $(\widetilde{ii})$, (iii) checken – (iv) folgt dann automatisch.

ein Ring im Sinn der Algebra ist. (Übungsaufgabe!)

$A \Delta B$ heißt übrigens symmetrische Differenz von A und B. (Skizze!)

Ich habe noch kein nichttriviales Beispiel einer σ-Algebra gegeben. In der Tat ist es gar nicht einfach, eine σ-Algebra $\mathfrak{A}$ durch eine Charakterisierung der $\mathfrak{A}$ konstituierenden Teilmengen von S anzugeben.*) Meistens „erzeugt" man sich σ-Algebren im folgenden Sinn: Ist eine Familie $\mathfrak{A}_i$ (wo i eine beliebige Indexmenge I durchläuft) von σ-Algebren ($\subset \mathfrak{P}(S)$) gegeben, so ist ihr Schnitt (das sind also diejenigen Teilmengen von S, die allen $\mathfrak{A}_i$ angehören) ebenfalls eine σ-Algebra, wie man unschwer bestätigt. (Eine entsprechende Aussage gilt für Ringe, σ-Ringe und Algebren.) Ist also $\mathcal{E} \subset \mathfrak{P}(S)$ irgendein Mengensystem, so existiert eine kleinste σ-Algebra $\sigma(\mathcal{E})$, die $\mathcal{E}$ umfaßt, nämlich der Schnitt aller $\mathcal{E}$ umfassenden σ-Algebren. (Es gibt stets garantiert eine solche σ-Algebra: $\mathfrak{P}(S)$!) $\sigma(\mathcal{E})$ heißt die von $\mathcal{E}$ erzeugte σ-Algebra und – umgekehrt – $\mathcal{E}$ Erzeuger von $\sigma(\mathcal{E})$.

Das wichtigste Beispiel liefert

1.4 Definition Die vom Ring der Figuren $\mathcal{F}^k$ erzeugte σ-Algebra heißt die Borel-σ-Algebra, in Zeichen $\mathrm{Bor}(\mathbb{R}^k)$. $E \in \mathrm{Bor}(\mathbb{R}^k)$ heißt Borelmenge oder auch Borel-meßbar.

Die Borel-σ-Algebra kann auch anders erzeugt werden:

*) Bei einer Topologie ist das ganz anders! (Beachte die Analogien zwischen den Definitionen für eine σ-Algebra und eine Topologie [aber auch die Unterschiede]!)

1.5 Satz Setze $\mathcal{E}_0 = \{\,]-\infty, r] : r \in \mathbb{R}\}$,

$\mathcal{E}_1 = \{A \subset \mathbb{R} : A \text{ offen}\}$,

$\mathcal{E}_2 = \{A \subset \mathbb{R} : A \text{ abgeschlossen}\}$,

$\mathcal{E}_3 = \{A \subset \mathbb{R} : A \text{ kompakt}\}$,

$\mathcal{E}_4 = \{\,]-\infty, r[\; : r \in \mathbb{R}\}$,

$\mathcal{E}_5 = \{\,]-\infty, r[\; : r \in \mathbb{Q}\}$,

$\mathcal{E}_6 = \{\,]-\infty, r] : r \in \mathbb{Q}\}$.

Dann ist $\sigma(\mathcal{E}_0) = \dots = \sigma(\mathcal{E}_6) = \mathrm{Bor}(\mathbb{R})$.

(Eine analoge Aussage gilt für $\mathrm{Bor}(\mathbb{R}^k)$.)

Beweis: Wir werden folgende, generell gültigen Schlußweisen benutzen:

- $\mathcal{E} \subset \mathcal{E}' \Rightarrow \sigma(\mathcal{E}) \subset \sigma(\mathcal{E}')$, insbesondere
- $\mathcal{E} \subset \mathfrak{A}$, $\mathfrak{A}$ σ-Algebra $\Rightarrow \sigma(\mathcal{E}) \subset \mathfrak{A}$.

$\sigma(\mathcal{E}_6) = \sigma(\mathcal{E}_0)$: $\subset$: klar, da $\mathcal{E}_6 \subset \mathcal{E}_0$.

$\supset$: zeige $\mathcal{E}_0 \subset \sigma(\mathcal{E}_6)$: wähle dazu zu $r \in \mathbb{R}$ rationale $r_n > r$ mit $r_n \xrightarrow[n\to\infty]{} r$. Dann $]-\infty, r] = \bigcap_n \,]-\infty, r_n] \in \sigma(\mathcal{E}_6)$.

$\sigma(\mathcal{E}_5) = \sigma(\mathcal{E}_6)$: $\supset$: zeige $\mathcal{E}_6 \subset \sigma(\mathcal{E}_5)$: Das geht wie oben: zu $r \in \mathbb{Q}$ wähle $r_n \in \mathbb{Q}$, $r_n > r$, $r_n \xrightarrow[n\to\infty]{} r$. Dann ist

$$]-\infty, r] = \bigcap_n \,]-\infty, r_n[\; \in \sigma(\mathcal{E}_5).$$

$\subset$: zeige $\mathcal{E}_5 \subset \sigma(\mathcal{E}_6)$. zu $r \in \mathbb{Q}$ wähle $r_n \in \mathbb{Q}$, $r_n < r$, $r_n \xrightarrow[n\to\infty]{} r$. Dann $]-\infty, r[\; = \bigcup_n \,]-\infty, r_n]$

$\sigma(\mathcal{E}_4) = \sigma(\mathcal{E}_0)$: Genauso!

$\sigma(\mathcal{E}_1) = \sigma(\mathcal{E}_0)$: $\supset$: Wegen $\mathcal{E}_4 \subset \mathcal{E}_1$ folgt $\sigma(\mathcal{E}_0) \underset{\text{s.o.}}{=} \sigma(\mathcal{E}_4) \subset \sigma(\mathcal{E}_1)$.

$\subset$: Zunächst bemerken wir, daß jede offene Menge A

abzählbare Vereinigung offener Intervalle ist, denn

$$A = \bigcup \,]r_i, s_i[\ ,$$

wo (r_i) bzw (s_i) eine Aufzählung von $\mathbb{Q}$ ist und die Vereinigung sich über diejenigen Intervalle erstreckt, die in A liegen. (Das ist eine abzählbare Vereinigung (warum?).)

Jedes offene Intervall ist in $\sigma(\mathcal{E}_0) \underset{\text{s.o.}}{=} \sigma(\mathcal{E}_4)$, denn

$$]r,s[\; = \;]-\infty, s[\; \cap \;]-\infty, r]^c .$$

Daher $\mathcal{E}_1 \subset \mathcal{E}_0$, folglich $\sigma(\mathcal{E}_1) \subset \sigma(\mathcal{E}_0)$

$\sigma(\mathcal{E}_1) = \sigma(\mathcal{E}_2)$: $\subset$: A offen $\Rightarrow$ A^c abgeschlossen zeigt $\mathcal{E}_1 \subset \sigma(\mathcal{E}_2)$,
denn $A = (A^c)^c$.
$\supset$: analog

$\sigma(\mathcal{E}_2) = \sigma(\mathcal{E}_3)$: $\supset$: wegen A kompakt $\Rightarrow$ A abgeschlossen.
$\subset$: wegen $A = \bigcup_n (A \cap [-n,n])$ und
A abgeschlossen $\Rightarrow$ $A \cap [-n,n]$ kompakt.

$\sigma(\mathcal{E}_0) = \mathrm{Bor}(\mathbb{R})$: $\supset$: wegen $]r,s] =]-\infty,s] \setminus]-\infty,r]$
$\subset$: wegen $]-\infty,r] = \bigcup_{\substack{n\in\mathbb{N} \\ n>|r|}}]-n,r]$

(Der Beweis für $\mathbb{R}^k$ ist (fast) wörtlich derselbe.)

Insbesondere zeigt Satz 1.5, daß $\mathrm{Bor}(\mathbb{R})$ einen abzählbaren Erzeuger besitzt. Man kann daraus mit einem Ordinalzahlargument folgern, daß die Mächtigkeit der Borel-σ-Algebra mit der Mächtigkeit von $\mathbb{R}$ übereinstimmt und daher kleiner ist als die Mächtigkeit der Potenzmenge von $\mathbb{R}$. Lax gesprochen bedeutet das, daß es sehr viel

weniger Borelmengen als Teilmengen überhaupt von $\mathbb{R}$ gibt. Damit hat man noch keine nichtborelsche Menge konstruiert; das ist in der Tat ziemlich schwierig, siehe z.B. Behrends S. 236 ff., insbesondere S. 249. Daß das so ist, liegt in erster Linie daran, daß alle üblicherweise in der Analysis auftretenden Mengen Borel-meßbar sind*); z.B. ist $\mathbb{Q}$ als abzählbare Vereinigung einpunktiger (also abgeschlossener) Mengen borelsch. Hier ein etwas kniffligeres Beispiel:

1.6 Beispiel Jeder Zahl in $]0,1]$ ordne man ihre nicht abbrechende dyadische Entwicklung zu:

$$x = 0.\, d_1 d_2 d_3 \dots ,$$

wo $d_i \in \{0,1\}$ und z.B. statt $0.1000\dots$ $0.01111\dots$ geschrieben werde. Borel nennt eine Zahl *normal*, wenn in ihrer dyadischen Entwicklung Nullen und Einsen „gleich häufig" vorkommen, präziser wenn $\frac{1}{n}\sum_{k=1}^{n} d_k \xrightarrow[n\to\infty]{} \frac{1}{2}$ konvergiert.
Ist die Menge N der normalen Zahlen eine Borelmenge?
Die Antwort steht auf S. 60.

Leider ist eine explizite Beschreibung aller Borelmengen unmöglich; trotzdem ist es möglich, Aussagen über beliebige Borelmengen auch ohne eine solche explizite Darstellung zu beweisen. Hier ein Beispiel (einen weiteren Beweis für 1.7 findet man auf S. 56):

*) Das ist ein immenser Vorzug der Borel-σ-Algebra!

1.7 Satz Ist $A_0 \subset \mathbb{R}^k$ eine Borelmenge und $x_0 \in \mathbb{R}^k$, so ist $x_0 + A_0 := \{x_0 + a : a \in A_0\}$ eine Borelmenge.

Beweis: Setze $\mathfrak{A}_0 := \{A \in Bor(\mathbb{R}^k) : x_0 + A \in Bor(\mathbb{R}^k)\} \subset Bor(\mathbb{R}^k)$

(Es ist also $A_0 \in \mathfrak{A}_0$ zu zeigen.)

Nun ist 1) $\mathbb{F}^k \subset \mathfrak{A}_0$ (denn mit A ist auch $x_0 + A$ eine Figur)

2) $\mathfrak{A}_0$ eine σ-Algebra. (Beweis kanonisch),

daher auch $Bor(\mathbb{R}^k) = \sigma(\mathbb{F}^k) \subset \mathfrak{A}_0$.

Insbesondere $A_0 \in \mathfrak{A}_0$, wie gewünscht.

Die in diesem Beweis verwandte Strategie nennt Ash das "good sets principle"; es funktioniert so:

Problem: Gegeben eine σ-Algebra $\mathfrak{A}$ und eine Eigenschaft (P). Zeige, daß jedes $A_0 \in \mathfrak{A}$ Eigenschaft (P) hat.

Lösungsstrategie: Betrachte die „guten Mengen", also

$$\mathfrak{A}_0 = \{A \in \mathfrak{A} : A \text{ hat } (P)\}$$

und zeige: 1) $\mathfrak{A}_0$ enthält einen Erzeuger von $\mathfrak{A}$

2) $\mathfrak{A}_0$ ist bereits σ-Algebra.

Es folgt dann $\mathfrak{A}_0 = \mathfrak{A}$.

Diesem "good sets principle" werden wir noch häufig begegnen.

Bisweilen werden wir nicht Teilmengen von $\mathbb{R}^k$, sondern nur Teilmengen eines vorgegebenen $E \subset \mathbb{R}^k$ betrachten und dort eine „Spur-σ-Algebra"

im Sinn folgender Definition induzieren:

1.8 Definition Sei S eine Menge, $E \subset S$ und $\mathcal{E} \subset \mathcal{P}(S)$.

$$\mathcal{E} \cap E := \{A \in \mathcal{P}(E) : \text{es ex. } F \in \mathcal{E} \text{ mit } A = F \cap E\}$$

heißt Spur von $\mathcal{E}$ auf E. Insbesondere heißt für $E \subset \mathbb{R}^k$

$$Bor(\mathbb{R}^k) \cap E$$

die Borel-σ-Algebra von E.

Es ist wirklich leicht zu verifizieren, daß die Spur einer σ-Algebra auf S eine σ-Algebra (auf E, freilich nicht auf S!) ist. Allgemeiner hat man das einfache

1.9 Lemma $\sigma(\mathcal{E} \cap E) = \sigma(\mathcal{E}) \cap E$

Beweis zur Übung!

Außerdem ist leicht zu sehen, daß für Borelmengen $E \subset \mathbb{R}^k$

$$Bor(E) = \{A \in Bor(\mathbb{R}^k) : A \subset E\}$$

gilt. (Beweis?) $Bor(E)$ ist jedoch auch für nicht Borel-meßbare E erklärt. (Die obige Definition von $Bor(E)$ ist verträglich mit der in Kap. V studierten der Borel-σ-Algebra eines metrischen (oder gar topologischen) Raums.)

I.2 Inhalte und Maße

Als nächstes sollen Funktionen auf Ringen betrachtet werden, die positive Zahlen oder $+\infty$ als Werte annehmen. Dabei benutzen wir folgende Konventionen über das Rechnen mit dem Symbol ∞:

$$a + \infty = \infty + a = \infty + \infty = \infty \quad \text{für } a \in \mathbb{R}$$

Ferner schreiben wir $\sum_{n=1}^{\infty} a_n = \infty$, falls $a_n \geq 0$, $a_n \in \mathbb{R}$ und die Reihe im eigentlichen Sinn divergent ist oder ein a_n ist.

2.1 Definition Sei R ein Ring und $\mu : R \to [0, \infty]$ mit

$$\mu(\emptyset) = 0.$$

a) μ heißt endlich additiv (oder Inhalt), falls

$$n \in \mathbb{N},\ A_1, \dots, A_n \in R,\ A_i \cap A_j = \emptyset \text{ für } i \neq j$$
$$\Rightarrow \mu\Big(\bigcup_{i=1}^{n} A_i\Big) = \sum_{i=1}^{n} \mu(A_i)$$

b) μ heißt σ-additiv (oder Prämaß), falls

$$A_1, A_2, \dots \in R,\ A_i \cap A_j = \emptyset \text{ für } i \neq j,\ \bigcup_{i=1}^{\infty} A_i \in R$$
$$\Rightarrow \mu\Big(\bigcup_{i=1}^{\infty} A_i\Big) = \sum_{i=1}^{\infty} \mu(A_i)$$

c) Ein Prämaß, das auf einer σ-Algebra definiert ist, heißt Maß.

Einige Bemerkungen: – Die Normierung $\mu(\emptyset) = 0$ dient dazu, das triviale Beispiel $\mu(A) = \infty \ \forall A \in R$ auszuschließen. Sie ergibt sich automatisch

für einen Inhalt, wenn man nur die Existenz eines $A \in R$ mit $\mu(A) < \infty$ voraussetzt ($\mu(A) = \mu(A \cup \emptyset) = \mu(A) + \mu(\emptyset)$).

– Die Bedingung der paarweisen Disjunktheit an die A_i werden wir in Zukunft durch die (etwas laxe) Sprechweise, die A_i bildeten eine disjunkte Folge, ausdrücken.

– Im Falle einer σ-Algebra R ist die Bedingung $\bigcup_{i=1}^{\infty} A_i \in R$ in b) automatisch erfüllt. (Spitzfindiger Zusatz: Dazu reichte bereits, daß R ein σ-Ring ist. Die Integrationstheorie auf σ-Ringen halte ich allerdings für zu unübersichtlich, dazu siehe z.B. Halmos' Buch.)

– Häufig spricht man von einem Maß (Inhalt) auf S statt auf einer σ-Algebra (einem Ring) $\subset \mathcal{P}(S)$, insbesondere, wenn es klar ist, auf welche σ-Algebra (etc.) man sich bezieht.

– Der entscheidende Begriff in Def. 2.1 ist der eines Maßes auf einer σ-Algebra, die übrigen haben eher technische Bedeutung.*)

– Ein Maß mit $\mu(S) = 1$ heißt Wahrscheinlichkeitsmaß (W-Maß). Die Interpretation ist dann, daß $\mu(A)$ die Wahrscheinlichkeit für das Eintreffen des Ereignisses A ist. Diese Interpretation legt nahe, daß der Definitionsbereich von μ eine Algebra und μ additiv sein sollten. Aus innermathematischen Gründen verlangt die moderne Axiomatik der Wahrscheinlichkeitstheorie jeweils die σ-Varianten (die Gründe liegen im wesentlichen in der Beweisbarkeit gewisser Grenzwertsätze). Diese σ-Axiome wurden zuerst von A. Kolmogoroff, Grundbegriffe der Wahrscheinlichkeitsrechnung, Springer 1933 postuliert (siehe insbesondere dort die S. 3 ff.). Übrigens findet sich ein rein endlich additiver

*) Offenbar ist jedes Prämaß ein Inhalt: $A_1 \cup \dots \cup A_n = A_1 \cup \dots \cup A_n \cup \emptyset \cup \emptyset \cup \dots$.

Zugang zu einem Problem der Wahrscheinlichkeitsrechnung in dem Buch von Dubins und Savage mit dem interessanten Titel "How to gamble if you must".

Als nächste sollen einige Beispiele und elementare Eigenschaften notiert werden.

<u>2.2 Beispiele</u> a) $\mu : \mathfrak{P}(S) \to [0,\infty]$ sei durch

$$\mu(\emptyset) = 0, \qquad \mu(E) = 1 \quad \text{für } E \neq \emptyset$$

definiert. Hat S mehr als einen Punkt, so ist μ kein Maß.

b) Hingegen definiert

$$\mu(\emptyset) = 0, \qquad \mu(E) = \infty \quad \text{für } E \neq \emptyset$$

ein Maß.

c) Sei $s_0 \in S$ fest und $\mathfrak{A} \subset \mathfrak{P}(S)$ eine σ-Algebra. Das <u>Dirac-Maß</u> δ_{s_0} ist durch $(A \in \mathfrak{A})$

$$\delta_{s_0}(A) = 1 \quad \text{falls } s_0 \in A$$
$$\delta_{s_0}(A) = 0 \quad \text{sonst}$$

definiert. Es ist wirklich ein Maß.

d) Das <u>zählende Maß</u> auf einer σ-Algebra $\mathfrak{A}$ ist durch $(A \in \mathfrak{A})$

$$\mu(A) = \text{Anzahl der Elemente von } A, \quad \text{falls } A \text{ endlich}$$
$$\mu(A) = \infty \quad \text{sonst}$$

definiert. Das ist auch wirklich ein Maß.

e) Es sei S überabzählbar und $\mathfrak{A}$ die σ-Algebra der abzählbaren und „co-abzählbaren" Mengen aus Bsp. 1.2 e). Für $A \in \mathfrak{A}$ setze

$$\mu_1(A) = \mu_2(A) = 0 \qquad \text{falls } A \text{ abzählbar}$$
$$\mu_1(A) = 1 \qquad \text{sonst}$$
$$\mu_2(A) = \infty$$

μ_1 und μ_2 sind (wohldefiniert!) Maße, die für allerlei Gegenbeispiele gut sind.

f) Die Länge eines halboffenen Intervalls ist (natürlich)

$$\lambda(\,]a,b]\,) = b-a.$$

Ist $A \in \mathcal{F}^1$ als disjunkte Vereinigung $A =]a_1,b_1] \cup \dots \cup]a_r,b_r]$ geschrieben, so ist man versucht, A die „Maßzahl"

$$\lambda(A) = \sum_{i=1}^{r} (b_i - a_i)$$

zuzuschreiben. Man muß sich natürlich davon überzeugen, daß diese Vorschrift wohldefiniert ist und nicht von der speziellen Art der Darstellung von A abhängt:

Das Argument ist schwieriger zu formulieren als zu verstehen und soll daher nur skizziert werden: Sei

$$A = \bigcup_{i=1}^{r} I_i = \bigcup_{j=1}^{s} J_j$$

jeweils disjunkte Vereinigung von halboffenen Intervallen. Es reicht dann,

$$\lambda(I_i) = \sum_{j=1}^{s} (I_i \cap J_j)$$

für alle i zu zeigen, und das stimmt.

Damit ist dann auch klar, daß λ ein Inhalt ist.

g) Allgemeiner wird der k-dimensionale Jordansche Inhalt eines halboffenen Intervalls

$$]a,b] =]\alpha_1,\beta_1] \times \dots \times]\alpha_k,\beta_k]$$

durch $\lambda^k(]a,b]) = (\beta_1-\alpha_1)\cdot\ldots\cdot(\beta_k-\alpha_k)$ erklärt. Einer Figur $A = \bigcup_{i=1}^{r}]a_i,b_i]$ (disjunkte Vereinigung) wird dann $\lambda^k(A) = \sum_{i=1}^{r} \lambda^k(]a_i,b_i])$ zugeordnet. λ^k ist ein Inhalt auf $\mathcal{F}^k$. (Das ist intuitiv klar, ein expliziter Beweis ist jedoch etwas mühsam zu formulieren. Siehe z.B. Behrends, S. 14 ff.)

Der Rest des ersten Kapitels zielt im wesentlichen darauf ab, λ^k zu einem Maß auf $\mathrm{Bor}(\mathbb{R}^k)$ fortzusetzen. Ein notwendiger Zwischenschritt dazu wird in Satz 2.5 gemacht, wo λ^k als σ-additiv auf $\mathcal{F}^k$ nachgewiesen wird.

Zunächst folgen einige allgemeine Eigenschaften.

2.3 Satz μ sei ein Inhalt auf einem Ring $\mathcal{R}$.

a) $A, B \in \mathcal{R}$, $A \subset B \Rightarrow \mu(A) \le \mu(B)$

b) $A_1, A_2, \dots \in \mathcal{R}$, $\bigcup_{i=1}^{\infty} A_i \in \mathcal{R}$, (A_i) disjunkte Folge

$$\Rightarrow \mu\Big(\bigcup_{i=1}^{\infty} A_i\Big) \ge \sum_{i=1}^{\infty} \mu(A_i)$$

c) $A_1, A_2, \dots \in \mathcal{R}$, $\bigcup_{i=1}^{\infty} A_i \in \mathcal{R}$, μ σ-additiv

$$\Rightarrow \mu\Big(\bigcup_{i=1}^{\infty} A_i\Big) \le \sum_{i=1}^{\infty} \mu(A_i)$$

Beweis: a) $B = A \cup B\setminus A$ (disjunkte Vereinigung)

$$\Rightarrow \mu(B) = \mu(A) + \mu(B\setminus A) \ge \mu(A). \qquad (B\setminus A \in \mathcal{R}\ !)$$

b) Sei $n \in \mathbb{N}$ beliebig. Dann $\bigcup_{i=1}^{\infty} A_i \supset \bigcup_{i=1}^{n} A_i$, also nach a)

$$\mu\Big(\bigcup_{i=1}^{\infty} A_i\Big) \ge \mu\Big(\bigcup_{i=1}^{n} A_i\Big) = \sum_{i=1}^{n} \mu(A_i)$$

$$\Rightarrow \mu\Big(\bigcup_{i=1}^{\infty} A_i\Big) \ge \sup_n \sum_{i=1}^{n} \mu(A_i) = \sum_{i=1}^{\infty} \mu(A_i).$$

c) Mit $B_1 = A_1$, $B_2 = A_2\setminus A_1$, $B_3 = A_3 \setminus (A_1 \cup A_2)$ etc. kann man $A := \bigcup_{i=1}^{\infty} A_i$ als disjunkte Vereinigung $A = \bigcup_{i=1}^{\infty} B_i$ schreiben; beachte $B_i \in \mathcal{R}$ sowie $B_i \subset A_i$. Daher

$$\mu(A) = \sum_{i=1}^{\infty} \mu(B_i) \underset{a)}{\le} \sum_{i=1}^{\infty} \mu(A_i).$$

Diese Eigenschaften werden im weiteren ohne viel Aufhebens benutzt.

Mit dem folgenden Kriterium kann häufig einfach die σ-Additivität gezeigt werden.

2.4 Satz μ sei ein Inhalt auf einem Ring $\mathcal{R}$. Betrachte die Eigenschaften

(i) μ ist σ-additiv

$\Downarrow$

(ii) Sind $A, A_1, A_2, \dots \in \mathcal{R}$ mit $A_1 \subset A_2 \subset \dots$, $\bigcup_{i=1}^{\infty} A_i = A$,
$\Uparrow$ so gilt $\mu(A) = \lim_{i\to\infty} \mu(A_i)$.

(iii) Sind $A, A_1, A_2, \dots \in \mathcal{R}$ mit $A_1 \supset A_2 \supset \dots$, $\bigcap_{i=1}^{\infty} A_i = A$,
$\Downarrow$ so gilt $\mu(A) = \lim_{i\to\infty} \mu(A_i)$.

(iv) Sind $A_1, A_2, \dots \in \mathcal{R}$, $A_1 \supset A_2 \supset A_3 \supset \dots$, $\bigcap_{i=1}^{\infty} A_i = \emptyset$,
so gilt $\lim_{i\to\infty} \mu(A_i) = 0$.

Es gelten die Implikationen

$$(i) \Leftrightarrow (ii) \underset{(*)}{\Leftarrow} (iii) \Leftrightarrow (iv).$$

Gilt in (iii) oder (iv) zusätzlich $\mu(A_1) < \infty$, so gilt auch die umgekehrte Implikation in (*).

Beweis: (i) $\Rightarrow$ (ii): $B_1 = A_1$, $B_2 = A_2 \setminus A_1$, $B_3 = A_3 \setminus (A_1 \cup A_2)$ etc.

definiert eine disjunkte Folge mit $A = \bigcup_{i=1}^{\infty} B_i$ und $A_n = \bigcup_{i=1}^{n} B_i$.

Es folgt $\mu(A) = \sum_i^{\infty} \mu(B_i) = \lim_{n\to\infty} \sum_{i=1}^{n} \mu(B_i) = \lim_{n\to\infty} \mu(A_n)$.

(ii) $\Rightarrow$ (i) Sei (B_n) eine disjunkte Folge mit $\bigcup_{i=1}^{\infty} B_i \in R$. Auf $A_i = B_1 \cup \cdots \cup B_i$ ist dann (ii) anwendbar, und es folgt (i) durch Rückwärtslesen des ersten Beweisteils.

(iii) $\Rightarrow$ (iv) ist klar

(iv) $\Rightarrow$ (iii) (A_i) sei wie in (iii), auf $(A_i \setminus A)$ ist dann (iv) anwendbar, daher

$$\mu(A_i) = \mu(A_i \setminus A) + \mu(A) \to \mu(A)$$

(iv) $\Rightarrow$ (i) Sei (A_i) eine disjunkte Folge mit $A := \bigcup_{i=1}^{\infty} A_i \in R$.

Auf B_n mit $B_n = \bigcup_{i=n}^{\infty} A_i$ ist (iv) anwendbar, so daß

$$\mu(A) = \mu(A_1 \cup \cdots \cup A_{n-1}) + \mu(B_n)$$
$$= \sum_{i=1}^{n-1} \mu(A_i) + \mu(B_n) \underset{n\to\infty}{\longrightarrow} \sum_{i=1}^{\infty} \mu(A_i).$$

Daher gilt (i).

(ii) $\Rightarrow$ (iii) für $\mu(A_1) < \infty$ in (iii):

Setze $B_n = A_1 \setminus A_n$, also $B_1 \subset B_2 \subset \cdots$, $\bigcup_{n=1}^{\infty} B_n = A_1 \setminus A$

$$\underset{(ii)}{\Rightarrow} \quad \mu(A_1 \setminus A) = \lim_{n\to\infty} \mu(A_1 \setminus A_n)$$

Da alle $\mu(A_n)$ endlich sind, muß wegen $\mu(A_1 \setminus A_n) = \mu(A_1) - \mu(A_n)$ (warum??) und $\mu(A_1 \setminus A) = \mu(A_1) - \mu(A)$ (iii) gelten.

Im allgemeinen ist die Implikation (ii) $\Rightarrow$ (iii) jedoch falsch:

Betrachte das zählende Maß μ auf $P(\mathbb{N})$ und $A_n = \{n, n+1, n+2, \cdots\}$.

Der in (ii) beschriebene Sachverhalt läßt sich in leicht verständlicher Symbolik auch so ausdrücken:

$$A_n \nearrow A \quad \Rightarrow \quad \mu(A_n) \nearrow \mu(A)$$

($\nearrow$ soll monoton wachsende Konvergenz symbolisieren.) In dieser Schreibweise wird klar, daß σ-Additivität eine Stetigkeitseigenschaft ist.

Besonders wichtig in Satz 2.4 ist die Implikation (iv) $\Rightarrow$ (i).

Zum Abschluß dieses Abschnitts machen wir noch einen entscheidenden Schritt auf dem Weg zum Lebesguemaß.

2.5 Satz λ^k ist σ-additiv auf $\mathcal{F}^k$.

Da der Beweis im wesentlichen ein Kompaktheitsschluß ist, soll an die entscheidende Eigenschaft zuerst erinnert werden:

> Sind $K_1, K_2, \cdots$ abgeschlossene Teilmengen eines kompakten metrischen Raums und gilt für alle n $K_1 \cap \cdots \cap K_n \neq \emptyset$, so ist auch $\bigcap_{i=1}^{\infty} K_i \neq \emptyset$

Das ist die sog. endliche Durchschnittseigenschaft, die durch Komplementbildung aus der üblichen Überdeckungsdefinition der Kompaktheit folgt.

Beweis von 2.5: Wir zeigen (iv) aus Satz 2.4. Sei also (A_n) eine absteigende Folge in $\mathcal{F}^k$ mit $\bigcap_{n=1}^{\infty} A_n = \emptyset$. Ist $\varepsilon > 0$ gegeben, ist $n_0 \in \mathbb{N}$ mit

$$n \geq n_0 \Rightarrow \lambda^k(A_n) < \varepsilon$$

zu produzieren. Dazu wähle für jede $n \in \mathbb{N}$ eine Figur B_n mit $\bar{B}_n \subset A_n$ und $\lambda^k(A_n \setminus B_n) \leq \varepsilon \cdot 2^{-n}$ (indem nämlich A_n „etwas verkleinert wird"; die präzise Formulierung der Konstruktion ist Euch überlassen). Nun ist erst recht $\bigcap_{n=1}^{\infty} \bar{B}_n = \emptyset$, und die $\bar{B}_n$ sind abgeschlossene Teilmengen des Kompaktums $\bar{A}_1$. Wegen der endlichen Durchschnittseigenschaft existiert n_0 mit

$$n \geq n_0 \Rightarrow \bigcap_{i=1}^{n} \bar{B}_i = \emptyset$$

Für $C_n := \bigcap_{i=1}^{n} B_i$ behaupten wir jetzt (Beweis folgt)

$$(*) \qquad \lambda^k(A_n \setminus C_n) \leq \varepsilon(1 - 2^{-n})$$

Da für $n \geq n_0$ $C_n = \emptyset$ ist, folgt daraus

$$\lambda^k(A_n) \leq \varepsilon(1 - 2^{-n}) < \varepsilon$$

für $n \geq n_0$, wie gewünscht. Beweisen wir also (*) durch Induktion. Für $n = 1$ ist (*) nach Konstruktion von B_1 richtig. Gelte nun (*) für n. Dann:

$$\begin{aligned} \lambda^k(A_{n+1} \setminus C_{n+1}) &= \lambda^k(A_{n+1} \setminus (C_n \cap B_{n+1})) \\ &= \lambda^k(A_{n+1} \setminus C_n \cup A_{n+1} \setminus B_{n+1}) \\ &\leq \lambda^k(A_{n+1} \setminus C_n) + \lambda^k(A_{n+1} \setminus B_{n+1}) \end{aligned}$$

(für einen Inhalt gilt stets $\mu(E \cup F) \leq \mu(E) + \mu(F)$)

$$\begin{aligned} &\underset{A_{n+1} \subset A_n}{\leq} \lambda^k(A_n \setminus C_n) + \lambda^k(A_{n+1} \setminus B_{n+1}) \\ &\leq \varepsilon(1 - 2^{-n}) + \varepsilon\, 2^{-(n+1)} \\ &= \varepsilon(1 - 2^{-(n+1)}). \end{aligned}$$

I.3 Die Konstruktion von Maßen nach Carathéodory

In diesem Abschnitt wird eine auf Carathéodory zurückgehende Methode beschrieben, Maße auf σ-Algebren zu erhalten. Insbesondere wird es möglich sein, den Jordanschen Inhalt zu einem Maß auf $\mathrm{Bor}(\mathbb{R}^k)$ auszudehnen. Dieses Maß (das Lebesguemaß! endlich!) wird dann im nächsten Abschnitt im Detail untersucht.

Grundlegend ist der folgende technische Begriff:

3.1 Definition $\alpha : \mathcal{P}(S) \to [0, \infty]$ heißt äußeres Maß, falls

a) $\alpha(\emptyset) = 0$

b) $A \subset B \Rightarrow \alpha(A) \leq \alpha(B)$

c) $A_1, A_2, \ldots \subset S \Rightarrow \alpha\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \alpha(A_i)$

3.2 Beispiele a) $\alpha(\emptyset) = 0$ bzw. $\alpha(A) = 1$ für $A \neq \emptyset$ definiert ein äußeres Maß.

b) $\alpha(A) = 0$ falls A höchstens abzählbar,
$\alpha(A) = 1$ falls A überabzählbar

definiert ebenfalls ein äußeres Maß.

c) Modifiziert man b) zu

$$\alpha(A) = \begin{cases} 0 & \text{falls } A \text{ endlich} \\ 1 & \text{falls } A \text{ unendlich} \end{cases},$$

so entsteht kein äußeres Maß, falls S unendlich ist. (Bedingung c) ist verletzt.)

Die für unsere Zwecke wichtigste Beispielklasse wird durch folgenden Satz geliefert:

3.3 Satz $\mu : \mathcal{R} \to [0,\infty]$ sei ein Inhalt auf einem Ring $\mathcal{R} \subset \mathcal{P}(S)$.
Für $A \subset S$ setze

$$\mu^*(A) = \inf \sum_{i=1}^{\infty} \mu(E_i),$$

wo das Infimum über alle Folgen (E_i) in $\mathcal{R}$ mit $A \subset \bigcup_{i=1}^{\infty} E_i$ zu erstrecken ist, bzw.

$$\mu^*(A) = \infty,$$

wenn es gar keine solche Folge gibt.
Dann ist μ^* ein äußeres Maß.

Beweis: Offenbar bildet μ^* nach $[0,\infty]$ ab, und es gelten a) und b) aus Def. 3.1. Um c) zu zeigen, dürfen wir $\alpha(A_i) < \infty \ \forall i \in \mathbb{N}$ annehmen, da andernfalls die Behauptung trivial ist. Es gibt daher zu $\varepsilon > 0$ E_{ij} mit $A_i \subset \bigcup_{j=1}^{\infty} E_{ij}$ und $\alpha(A_i) \geq \sum_{j=1}^{\infty} \mu(E_{ij}) - \varepsilon 2^{-i}$. Daher $\bigcup_{i=1}^{\infty} A_i \subset \bigcup_{i,j=1}^{\infty} E_{ij}$ und

$$\alpha\left(\bigcup_i A_i\right) \leq \sum_{i,j} \mu(E_{ij}) \leq \sum_i \left(\alpha(A_i) + \varepsilon 2^{-i}\right) \leq \sum_i \alpha(A_i) + \varepsilon.$$

Da $\varepsilon > 0$ beliebig war, zeigt das die Behauptung.

Obwohl ein äußeres Maß i.a. weit davon entfernt ist, additiv (geschweige denn ein Maß) zu sein, führen geeignete Einschränkungen von äußeren Maßen in der Tat zu Maßen. Der folgende Begriff ist nun hilfreich:

3.4 Definition $\alpha : \mathcal{P}(S) \to [0,\infty]$ sei ein äußeres Maß. $A \subset S$ heißt α-meßbar, falls

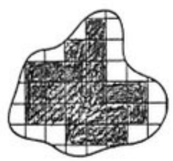
$$(*) \qquad \alpha(Q) = \alpha(Q \cap A) + \alpha(Q \cap A^c) \qquad \forall\ Q \subset S.$$

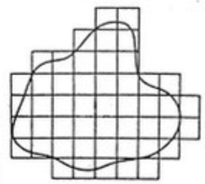
$\mathcal{M}_\alpha$ bezeichnet die Menge aller α-meßbaren Teilmengen von S.

3.5 Satz Für ein äußeres Maß α ist $\mathcal{M}_\alpha$ eine σ-Algebra, und α definiert ein Maß auf $\mathcal{M}_\alpha$.

Bevor Satz 3.5 bewiesen wird, eine Bemerkung zu Def. 3.4, die extrem unanschaulich ist. Hewitt und Stromberg kommentieren sie mit den Worten: "How Carathéodory came to think of this definition seems mysterious, since it is not in the least intuitive." (a.a.O., S.127). Trotzdem möchte ich versuchen, Def. 3.4 zu erklären, indem ich sie Lebesgues Definition einer meßbaren Menge gegenüberstelle.
Betrachten wir etwa eine beschränkte Menge A in $\mathbb{R}^2$, o.E. sei $A \subset Q := [0,1]^2$. Um A einen Inhalt zuzuordnen, kann man so vorgehen (C. Jordan hat diesen Weg vorgeschlagen):

Man überdecke A durch endliche Vereinigungen von Rechtecken (deren Inhalt intuitiv klar ist) und nenne äußeren (Jordanschen) Inhalt $j^*(A)$ das Infimum dieser Rechtecksmaßzahlen. Alsdann schöpfe man A von innen durch endliche Vereinigungen von Rechtecken aus und definiere $j_*(A)$, den inneren Inhalt, als das Supremum der erhaltenen Rechtecksmaßzahlen.





Es liegt nun nahe, A Jordan-meßbar zu nennen, falls $j^*(A) = j_*(A)$ ist. Nun ist zu beachten, daß einer inneren Ausschöpfung von A eine äußere Überdeckung von $Q\setminus A$ entspricht (und umgekehrt) und daß $j_*(A) = 1 - j^*(Q\setminus A)$ ist. A ist daher genau dann Jordan-meßbar, falls

$$j^*(A) + j^*(Q\setminus A) = 1 \quad (= j^*(Q))$$

ist. Lebesgue geht (wie kurz vorher bereits Borel) hier einen scheinbar nur winzigen Schritt weiter, der jedoch umwälzende Konsequenzen nach sich zieht (diese Konsequenzen bilden den Inhalt dieser Vorlesung): Er läßt abzählbare Überdeckungen durch Rechtecke zu und definiert ein entsprechendes äußeres (Lebesguesches) Maß λ^* (genau dasselbe, das Def. 3.3 liefert!), definiert ein inneres Maß λ_* durch $\lambda_*(A) := 1 - \lambda^*(Q\setminus A)$ und nennt meßbar, falls $\lambda_*(A) = \lambda^*(A)$ ist, d.h. falls

$$\lambda^*(A) + \lambda^*(Q\setminus A) = \lambda^*(Q) \quad (=1)$$

ist. Das ist im wesentlichen dasselbe, was in Def. 3.4 gefordert ist, bis auf die Nuance, daß die dortige Bedingung für alle Q gefordert

wird. (Das macht die Angelegenheit technisch einfacher zu handhaben.)

Nun zum Beweis von 3.5.

Klar sind $\emptyset, S \in \mathcal{M}_\alpha$ und $A \in \mathcal{M}_\alpha \Rightarrow A^c \in \mathcal{M}_\alpha$ (Bedingung (ii) von S. 11). Der Rest ist nun recht knifflig.

Wir bemerken zunächst, daß es in (*) von Def. 3.4 reicht „$\geq$" zu zeigen, da die andere Ungleichung als Folge der Bedingungen a) bis c) eines äußeren Maßes stets erfüllt ist.

Als erstes wird jetzt

(i) $A, B \in \mathcal{M}_\alpha \Rightarrow A \cup B \in \mathcal{M}_\alpha$

gezeigt, woraus induktiv folgt, daß $\mathcal{M}_\alpha$ eine Algebra ist.

Sei $Q \subset S$. Da $B \in \mathcal{M}_\alpha$ ist, gilt

(1) $\alpha(Q \cap A^c) = \alpha(Q \cap A^c \cap B) + \alpha(Q \cap A^c \cap B^c)$.

Ferner ist

$$Q \cap (A \cup B) = (Q \cap A) \cup (Q \cap A^c \cap B);$$

da α äußeres Maß ist, ergibt sich

(2) $\alpha(Q \cap (A \cup B)) \leq \alpha(Q \cap A) + \alpha(Q \cap A^c \cap B)$.

(1) und (2) zusammen zeigen

$$\begin{aligned} &\alpha(Q \cap (A \cup B)) + \alpha(Q \cap (A \cup B)^c) \\ &\leq \alpha(Q \cap A) + \alpha(Q \cap A^c \cap B) + \alpha(Q \cap A^c \cap B^c) \\ &= \alpha(Q \cap A) + \alpha(Q \cap A^c) \\ &= \alpha(Q) \qquad (\text{denn } A \in \mathcal{M}_\alpha) \end{aligned}$$

wie gewünscht.

Nimmt man in der letzten Rechnung A und B disjunkt an und er[setzt] Q durch $Q \cap (A \cup B)$, so ergibt sich

$$\alpha(Q \cap (A \cup B)) = \alpha(Q \cap A) + \alpha(Q \cap B)$$

und daraus induktiv

(ii) $$A_1, \dots, A_n \in \mathcal{M}_\alpha,\ A_i \cap A_j = \emptyset \text{ für } i \neq j,\ Q \subset S \quad \Rightarrow \quad \alpha\Big(Q \cap \bigcup_{i=1}^n A_i\Big) = \sum_{i=1}^n \alpha(Q \cap A_i)$$

Nun können wir die σ-Additivität von α beweisen; genauer: Sind $A_1, A_2, \dots \in \mathcal{M}_\alpha$ paarweise disjunkt, so ist

(iii) $$A := \bigcup_{i=1}^\infty A_i \in \mathcal{M}_\alpha$$

(iv) $$\alpha(A) = \sum_{i=1}^\infty \alpha(A_i)$$

Setze $B_n = \bigcup_{i=1}^n A_i$. $B_n \in \mathcal{M}_\alpha$ nach (i), also gilt für $Q \subset$[S]

$$\begin{aligned} \alpha(Q) &= \alpha(Q \cap B_n) + \alpha(Q \cap B_n^c) \\ &\underset{(ii)}{=} \sum_{i=1}^n \alpha(Q \cap A_i) + \alpha(Q \cap B_n^c) \\ &\underset{\substack{B_n \subset A \\ \alpha \text{ monoton}}}{\geq} \sum_{i=1}^n \alpha(Q \cap A_i) + \alpha(Q \cap A^c) \end{aligned}$$

$$\begin{aligned} \Rightarrow \quad \alpha(Q) &\geq \sum_{i=1}^\infty \alpha(Q \cap A_i) + \alpha(Q \cap A^c) \\ &\geq \alpha(Q \cap A) + \alpha(Q \cap A^c) \end{aligned}$$ (wg. Eigenschaft aus Def.

Das zeigt (iii) nach der Vorbemerkung des Beweises. Folg[lich] gilt in der letzten Ungleichungskette sogar Gleichheit, w[as] für $Q = A$ genau (iv) besagt.

Bleibt zu zeigen:

(v) $\mathcal{M}_\alpha$ ist σ-Algebra (das ist mehr als (iii)!):

Seien $B_1, B_2, \dots \in \mathcal{M}_\alpha$. Setze $A_1 = B_1$, $A_2 = B_2 \setminus B_1$, $A_3 = B_3 \setminus (B_1 \cup B_2)$, etc. Die A_i sind paarweise disjunkt, liegen wegen (i) in $\mathcal{M}_\alpha$, und es ist $\bigcup_{i=1}^\infty A_i = \bigcup_{i=1}^\infty B_i$. (iii) impliziert daher $\bigcup_{i=1}^\infty B_i \in \mathcal{M}_\alpha$.

Ist ein Inhalt μ auf einem Ring R vorgelegt, so ist das gemäß 3.3 zugehörige äußere Maß μ^* auf $\mathcal{M}_{\mu^*}$ σ-additiv. Es fragt sich natürlich, ob $R \subset \mathcal{M}_{\mu^*}$ gilt und, falls ja, was $\mu(A)$ mit $\mu^*(A)$ zu tun hat. Als entscheidende Bedingung stellt sich die σ-Additivität von μ heraus, was im zentralen Satz dieses Abschnitts formuliert wird:

3.6 Theorem (Fortsetzungssatz von Carathéodory)

Sei $R \subset \mathcal{P}(S)$ ein Ring und $\mu : R \to [0, \infty]$ ein Prämaß. Dann kann μ zu einem Maß $\bar{\mu}$ auf die von R erzeugte σ-Algebra fortgesetzt werden. Genauer: Jedes $A \in R$ ist μ^*-meßbar, und es gilt $\mu^*(A) = \mu(A)$. $\bar{\mu}$ kann also als Einschränkung von μ^* auf $\sigma(R)$ gewählt werden; μ kann sogar zu einem Maß auf $\mathcal{M}_{\mu^*}$ fortgesetzt werden.

Beweis: Zeigen wir zuerst $\mu^*(A) = \mu(A)$ für $A \in R$

$\leq$: da $A, \emptyset, \emptyset, \emptyset, \dots$ eine zulässige Überdeckung von A ist.

$\geq$: Im Falle $\mu^*(A) = \infty$ ist nichts zu zeigen. Sei nun

$A \subset \bigcup_{i=1}^{\infty} A_i$ mit $A_i \in R$; dann ist $A = \bigcup_{i=1}^{\infty} (A \cap A_i)$, und 2.3 impliziert ($\mu$ ist σ-additiv!)

$$\mu(A) \le \sum_{i=1}^{\infty} \mu(A \cap A_i) \le \sum_{i=1}^{\infty} \mu(A_i).$$

Es folgt durch Übergang zum Infimum $\mu(A) \le \mu^*(A)$.

und danach $R \subset \mathcal{M}_{\mu^*}$:

Sei $A \in R$ und sei $Q \subset S$. Es ist (vgl. den Beweis zu 3.5) $\mu^*(Q) \geq \mu^*(Q \cap A) + \mu^*(Q \cap A^c)$ zu zeigen. Das ist klar im Fall $\mu^*(Q) = \infty$. Daher können wir die Existenz von $A_1, A_2, \dots \in R$ mit $Q \subset \bigcup_{i=1}^{\infty} A_i$ annehmen. Da $Q \cap A \subset \bigcup_{i=1}^{\infty} (A_i \cap A)$ und $Q \cap A^c \subset \bigcup_{i=1}^{\infty} (A_i \cap A^c) = \bigcup_{i=1}^{\infty} (A_i \setminus A)$ Überdeckungen in R sind, gilt

$$\sum_{i=1}^{\infty} \mu(A_i) = \sum_{i=1}^{\infty} \mu(A_i \cap A) + \sum_{i=1}^{\infty} \mu(A_i \cap A^c) \geq \mu^*(Q \cap A) + \mu^*(Q \cap A^c),$$

was die Behauptung zeigt.

Nun bleibt nur noch, Satz 3.5 anzuwenden!

Im Rest des Abschnitts sollen noch 2 Fragen im Zusammenhang mit 3.6 geklärt werden, nämlich:

- Ist die Fortsetzung $\bar{\mu}$ eindeutig bestimmt?
- Wie hängen $\sigma(R)$ und $\mathcal{M}_{\mu^*}$ zusammen?

Die Diskussion der ersten Frage beginnen wir mit 2 Gegenbeispielen:

- Sei S eine überabzählbare Menge, $R \subset \mathcal{P}(S)$ der Ring der endlichen Teilmengen, $\mu: R \to [0,\infty]$ sei der triviale Inhalt $\mu(A) = 0 \ \forall A \in R$. Da die von R erzeugte σ-Algebra die σ-Algebra der abzählbaren und coabzählbaren Teilmengen (Bsp. 1.2 e)) ist, ist für jedes $r \in [0,\infty]$

$$\mu_r(A) = 0 \quad \text{falls } A \text{ abzählbar}$$
$$\mu_r(A) = r \quad \text{falls } S \setminus A \text{ abzählbar}$$

eine Fortsetzung von μ zu einem Maß auf $\sigma(R)$.

- Betrachte $\mathbb{Q}$ und die σ-Algebra $\mathcal{P}(\mathbb{Q})$. Die Spur (Def. 1.8) von $\mathbb{F}^1$ auf $\mathbb{Q}$ ist ein Erzeuger von $\mathcal{P}(\mathbb{Q})$, auf dem das zählende Maß μ und $2 \cdot \mu$ übereinstimmen.

[Es ist instruktiv, sich zu überlegen, welche Fortsetzung 3.6 liefert!]

Offenbar ist für das Scheitern im 1. Gegenbeispiel verantwortlich, daß die Mengen in R die Menge S „nicht erreichen", im 2. Gegenbeispiel ist es die starke Unendlichkeit von μ auf R. Wir betrachten jetzt Mengenfunktionen, die diese Defekte nicht aufweisen.

<u>3.7 Definition</u> Ein Inhalt μ auf einem Ring R heißt σ-endlich, falls es eine aufsteigende Folge $A_1 \subset A_2 \subset A_3 \subset \cdots$ von Mengen in R mit $\mu(A_n) < \infty \ \forall n \in \mathbb{N}$ und $\bigcup_{n=1}^{\infty} A_n = S$ gibt.

Zum Beispiel ist der Jordansche Inhalt auf $\mathbb{F}^k$ σ-endlich, desgleichen ist es das zählende Maß auf $\mathcal{P}(\mathbb{Q})$, nicht jedoch auf dem Erzeuger $\mathbb{F}^1 \cap \mathbb{Q}$ (!!). Offenbar ist ein endliches Maß auf einer σ-Algebra σ-endlich.

Unser Ziel ist es zu zeigen, daß für σ-endliche μ die Fortsetzung $\tilde{\mu}$ in Satz 3.6 eindeutig bestimmt ist. Zuerst ein allgemeiner Eindeutigkeits

3.8 Satz Sei $\mathcal{E}$ ein $\cap$-stabiler*) Erzeuger der σ-Algebra $\mathfrak{A} \subset \mathfrak{P}(S)$ und μ und ν seien endliche Maße auf $\mathfrak{A}$, die auf $\mathcal{E} \cup \{S\}$ übereinstimmen. Dann ist $\mu = \nu$.

Die Beweisidee ist einfach: betrachte (good sets principle!)

$$\mathcal{D} = \{ A \in \mathfrak{A} : \mu(A) = \nu(A) \}.$$

Nach Voraussetzung enthält $\mathcal{D}$ einen Erzeuger von $\mathfrak{A}$, und man m $\mathcal{D}$ als σ-Algebra entlarven. In der Tat sieht man sofort

(i) $\emptyset \in \mathcal{D}$, $S \in \mathcal{D}$

(ii) $A \in \mathcal{D} \Rightarrow A^c \in \mathcal{D}$

(iii) $A_1, A_2, \dots \in \mathcal{D}$ disjunkte Folge $\Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{D}$

Wenn man noch die Disjunktheit in (iii) loswürde, wäre man fert
Das gelingt mit einer raffinierten Methode, die auf Dynkin zurückg
Diese Methode soll nun beschrieben werden.

3.9 Definition Ein Mengensystem $\mathcal{D}$ mit den obigen Eigenschaften (i), (ii), (iii) heißt Dynkinsystem.

Offenbar ist jede σ-Algebra ein Dynkinsystem, und (wie üblich) existiert zu $\mathcal{E} \subset \mathfrak{P}(S)$ stets ein kleinstes Dynkinsystem $d(\mathcal{E})$, das $\mathcal{E}$ umfaßt.

*) d.h.: $E_1, E_2 \in \mathcal{E} \Rightarrow E_1 \cap E_2 \in \mathcal{E}$.

3.10 Satz Ist $\mathcal{E}$ $\cap$-stabil, so gilt $\sigma(\mathcal{E}) = d(\mathcal{E})$.

Beweis: Wir werden nacheinander zeigen:

a) Ist $\mathcal{D}$ ein Dynkinsystem und gilt $D_1, D_2 \in \mathcal{D}$, $D_1 \subset D_2$, so ist $D_2 \setminus D_1 \in \mathcal{D}$.

b) Ein $\cap$-stabiles Dynkinsystem ist eine σ-Algebra.

c) Mit $\mathcal{E}$ ist auch $d(\mathcal{E})$ $\cap$-stabil.

Daraus folgt die Behauptung des Satzes.

Zu a): $D_2 \setminus D_1 = D_2 \cap D_1^c = (D_2^c \cup D_1)^c = (D_2^c \cup D_1 \cup \emptyset \cup \emptyset \cup \dots)^c \in \mathcal{D}$, denn die letzte Vereinigung ist disjunkt.

Zu b): Seien $A_1, A_2, \dots \in \mathcal{D}$. Wie schon mehrmals zuvor, disjunktifizieren wir die A_i: $B_1 = A_1$, $B_2 = A_2 \setminus A_1$, $B_3 = A_3 \setminus (A_1 \cup A_2)$ etc. Beachte $B_i = A_i \cap A_1^c \cap \dots \cap A_{i-1}^c \in \mathcal{D}$, so daß wirklich $\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{\infty} B_i \in \mathcal{D}$.

Zu c): Sei $D \in d(\mathcal{E})$. Wir haben

$$\mathcal{D}_D := \{ Q \in d(\mathcal{E}) : Q \cap D \in d(\mathcal{E}) \} \supset d(\mathcal{E})$$

zu zeigen (good sets principle!). Zunächst ist $\mathcal{D}_D$ ein Dynkinsystem: (i) und (iii) sind offensichtlich, und für (ii) verwende a):

$$Q \in \mathcal{D}_D \Rightarrow Q^c \cap D = D \setminus Q = D \setminus (Q \cap D) \in d(\mathcal{E}).$$

Es reicht daher, $\mathcal{D}_D \supset \mathcal{E}$ für alle $D \in d(\mathcal{E})$ zu zeigen.

Sammeln wir wieder die guten Mengen ein:

$$\mathcal{D} := \{ D \in d(\mathcal{E}) : \mathcal{D}_D \supset \mathcal{E} \}$$

Wie oben sieht man, daß $\mathcal{D}$ ein Dynkinsystem ist, und $\mathcal{D}$ umfaßt $\mathcal{E}$, denn $\mathcal{E}$ ist $\cap$-stabil. Es folgt $d(\mathcal{E}) \subset \mathcal{D}$, wie gewünscht.

3.10 impliziert sofort 3.8, denn das auf S. 36 vorgeschlagene Mengensystem ist ein Dynkinsystem, das einen $\cap$-stabilen Erzeuger von $\mathfrak{A}$ umfaßt. Daher

$$\mathfrak{A} = \sigma(\mathcal{E}) = d(\mathcal{E}) \underset{3.10}{\subset} \mathcal{D} \subset \mathfrak{A},$$

d.h. $\mathcal{D} = \mathfrak{A}$.

3.8 läßt folgende wichtige Verallgemeinerung zu:

<u>3.11 Satz</u> μ und ν seien Maße auf einer σ-Algebra $\mathfrak{A}$, die auf einem $\cap$-stabilen Erzeuger $\mathcal{E}$ von $\mathfrak{A}$ übereinstimmen. μ und ν seien simultan σ-endlich auf $\mathcal{E}$, d.h. es existieren $E_1 \subset E_2 \subset \dots$ in $\mathcal{E}$ mit $\bigcup_{i=1}^{\infty} E_i = S$ und $\mu(E_i) < \infty$, $\nu(E_i) < \infty$ $\forall i \in \mathbb{N}$.
Dann gilt $\mu = \nu$.

Beweis: Setze $\mu_n(A) = \mu(A \cap E_n)$ und $\nu_n(A) = \nu(A \cap E_n)$.
3.8 impliziert $\mu_n = \nu_n$ $\forall n \in \mathbb{N}$, während 2.4

$$\mu(A) = \lim_{n\to\infty} \mu(A \cap E_n) = \lim_{n\to\infty} \nu(A \cap E_n) = \nu(A)$$

für alle $A \in \mathfrak{A}$ liefert.

<u>3.12 Korollar</u> Ist unter den Voraussetzungen von 3.6 μ auf R σ-endlich, so ist $\mu^*|_{\sigma(R)}$ die einzige σ-additive Fortsetzung von μ auf $\sigma(R)$. ($\bar{\mu}$ ist also eindeutig bestimmt.)

Kommen wir abschließend zur 2. Frage von S. 34. Wir formulieren zunächst ein einfaches Lemma.

<u>3.13 Lemma</u> Sei α ein äußeres Maß und $\alpha(A) = 0$. Dann ist $A \in \mathcal{M}_\alpha$.

Beweis: Da α monoton ist, ist auch ($Q \subset S$ beliebig) $\alpha(Q \cap A) = 0$ sowie

$$\alpha(Q) \geq \alpha(Q \cap A^c) = \alpha(Q \cap A) + \alpha(Q \cap A^c).$$

Nach der Bemerkung von S. 31, Zeile 5 ist A α-meßbar.

<u>3.14 Satz</u> Es sei μ ein σ-endliches Prämaß auf dem Ring R.
Dann sind äquivalent:

(i) $A \in \mathcal{M}_{\mu^*}$.

(ii) Es existiert $N \subset S$ mit $\mu^*(N) = 0$ und $A \cup N \in \sigma(R)$.

Beweis: (ii) $\Rightarrow$ (i) wegen $\sigma(R) \subset \mathcal{M}_{\mu^*}$ und Lemma 3.13.

(i) $\Rightarrow$ (ii): Es sei zuerst $\mu^*(A) < \infty$ angenommen. Zu $n \in \mathbb{N}$ wähle $A_{in} \in R$ mit $A \subset \bigcup_{i=1}^{\infty} A_{in} =: B_n$ und

$$\mu^*(A) + \tfrac{1}{n} \geq \sum_{i=1}^{\infty} \mu(A_{in}) \geq \mu^*(B_n)$$

(bemerke $B_n \in \sigma(R) \subset \mathcal{M}_{\mu^*}$ und daß μ^* dort σ-additiv ist). Setze $C_n = \bigcap_{i=1}^{n} B_i$ und $C = \bigcap_{n=1}^{\infty} C_n \in \sigma(R)$. Da ja

$$A \subset C_n \subset B_n, \quad \text{also auch} \quad A \subset C$$

ist, gilt

$$\mu^*(C) + \tfrac{1}{n} \geq \mu^*(A) + \tfrac{1}{n} \geq \mu^*(C_n).$$

Satz 2.4 liefert $\mu^*(C_n) \xrightarrow[n\to\infty]{} \mu^*(C)$, daher $\mu^*(A) = \mu^*(C)$.
Setze also $N = C \setminus A$.

Nun zum allgemeinen Fall. Wähle $S_n \in R$, $\mu(S_n) < \infty$, $\bigcup_{n=1}^{\infty} S_n = S$

und betrachte $A_n = A \cap S_n$. Wegen $\mu^*(A_n) \le \mu^*(S_n) = \mu(S_n) < \infty$ ist der erste Beweisteil anwendbar: Es existieren N_n mit $\mu^*(N_n) = 0$ und $A_n \cup N_n \in \sigma(\mathcal{R})$. Für $N = \bigcup_{n=1}^{\infty} N_n$ gilt nun $\mu^*(N) = 0$, denn $N_n \in \mathcal{M}_{\mu^*}$ nach 3.13, und μ^* ist σ-additiv auf $\mathcal{M}_{\mu^*}$.

$$A \cup N = \bigcup_n (A \cap S_n) \cup \bigcup_n N_n = \bigcup_n (A_n \cup N_n) \in \sigma(\mathcal{R})$$

zeigt die Behauptung.

3.14 sagt also, daß sich $\mathcal{M}_{\mu^*}$ und $\sigma(\mathcal{R})$ (für σ-endliches μ) nur durch μ^*-Nullmengen unterscheiden. Der Maßraum $(S, \mathcal{M}_{\mu^*}, \mu^*)$ kann im folgenden Sinn als „Vervollständigung" von $(S, \sigma(\mathcal{R}), \mu^*)$ angesehen werden: der erstgenannte Maßraum enthält alle Teilmengen von μ^*-Nullmengen, während dies der zweite i.a. nicht tut.

(Allgemein nennt man einen Maßraum $(S, \mathfrak{A}, \nu)$ vollständig, wenn

$$A \in \mathfrak{A},\ \nu(A) = 0,\ B \subset A \Rightarrow B \in \mathfrak{A}$$

gilt.)

I.4 Das Lebesguemaß

Endlich sind wir soweit, die bisherige Ernte einzufahren und ein nicht triviales Beispiel für ein Maß zu produzieren. Die bisherigen Resultate implizieren (vgl. 2.5, 3.6, 3.12):

4.1 Satz Es gibt genau ein Maß λ^k auf Bor $(\mathbb{R}^k)$, das halboffenen Intervallen ihren Jordanschen Inhalt zuordnet.

λ^k heißt (k-dimensionales) Lebesguesches Maß. Wahlweise kann λ^k auch auf Leb $(\mathbb{R}^k) := \mathcal{M}_{(\lambda^k)^*}$ betrachtet werden; aus 3.14 folgt, daß das Lebesguemaß auch dort noch eindeutig bestimmt ist.

Weiter unten wird übrigens gezeigt, daß

$$\mathrm{Bor}(\mathbb{R}^k) \subsetneq \mathrm{Leb}(\mathbb{R}^k) \subsetneq \mathcal{P}(\mathbb{R}^k)$$

gilt. Eine Menge $A \in \mathrm{Leb}(\mathbb{R}^k)$ soll Lebesgue-meßbar genannt werden.

Ist $A \in \mathrm{Bor}(\mathbb{R}^k)$, so kann λ^k auf naheliegende Weise als Maß auf Bor (A) erklärt werden, da ja für $B \subset A$ $B \in \mathrm{Bor}(A) \Leftrightarrow B \in \mathrm{Bor}(\mathbb{R}^k)$ (vgl. S. 17). Etwas schlampig spricht man dann auch vom Lebesguemaß auf A.

Einige Eigenschaften von λ^k:

4.2 Satz λ^k ist translationsinvariant, d.h. für $A \in \mathrm{Bor}(\mathbb{R}^k)$ ist

$$\lambda^k(x + A) = \lambda^k(A) \quad \forall x \in \mathbb{R}^k.$$

(Dieselbe Aussage gilt auf Leb $(\mathbb{R}^k)$.)

Beweis: Wir wissen bereits $x + A \in \mathrm{Bor}(\mathbb{R}^k)$ (1.7). Definiert man nun μ auf Bor $(\mathbb{R}^k)$ durch $\mu(A) := \lambda^k(x + A)$, so gilt:

- μ ist ein Maß (nachprüfen)
- μ stimmt mit λ^k auf halboffenen Intervallen überein.

Aus der Eindeutigkeitsaussage in 4.1 folgt $\lambda^k = \mu$.

Dieser Satz gestattet folgende Umkehrung:

4.3 Satz Ist μ ein translationsinvariantes Maß auf $\mathrm{Bor}(\mathbb{R}^k)$ mit $c := \mu(]$ ...
so ist $\mu = c \cdot \lambda^k$.

Beweis: Durch fortgesetztes Halbieren oder Verdoppeln folgt aus der Translationsinvarianz, daß μ und $c \cdot \lambda^k$ auf allen Intervallen $]\alpha_1, \beta_1] \times \dots \times]\alpha_k, \beta_k]$, wo α_i, β_i dyadische Brüche sind ($2^n \cdot \alpha_i \in \mathbb{Z}$ für passende n), übereinstimmen. Es sei $\mathcal{E}$ das System aller endlichen Vereinigungen solcher Intervalle. $\mathcal{E}$ erzeugt $\mathrm{Bor}(\mathbb{R}^k)$, und 3.11 ist darauf anwendbar. Es folgt $\mu = c \cdot \lambda^k$.

Man erhält daraus:

4.4 Korollar $T: \mathbb{R}^k \to \mathbb{R}^k$ sei eine lineare Abbildung mit $\det T \neq 0$. Dann gilt für alle $A \in \mathrm{Bor}(\mathbb{R}^k)$
$$\lambda^k(T(A)) = |\det T| \cdot \lambda^k(A).$$

Beweis: Mit einem "good sets" Argument beweist man zunächst $T(A) \in \mathrm{Bor}(\mathbb{R}^k)$ (Übung!) Ferner definiert $\mu(A) = \lambda^k(T(A))$ ein Maß (Beweis?), das wegen
$$\mu(A+x) = \lambda^k(T(A+x)) = \lambda^k(T(A)+Tx) = \lambda^k(T(A)) = \mu(A)$$
translationsinvariant ist. Also ist $\mu = \lambda^k(T(Q)) \cdot \lambda^k$, wo $Q =]0,1]^k$. Bleibt $\lambda^k(T(Q)) = |\det T|$ zu zeigen.

Nun ist jede lineare Abbildung T als Kompositum $T_1 T_2 \cdots T_s$ von linearen Abbildungen folgender einfacher Typen zu schreiben

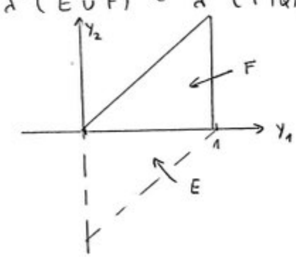
(das wird in der Linearen Algebra gezeigt):

(i) $T_i(x_1, x_2, \dots, x_k) = (x_{\pi(1)}, \dots, x_{\pi(k)})$ für eine Permutation π.

(ii) $T_i(x_1, x_2, \dots, x_k) = (a x_1, x_2, \dots, x_k)$

(iii) $T_i(x_1, x_2, \dots, x_k) = (x_1 + x_2, x_2, \dots, x_k)$

Die gewünschte Formel
$$(*) \qquad \lambda^k(T(Q)) = |\det T|$$
ist nun offensichtlich richtig, falls T vom Typ (i) oder (ii) ist. Ist T vom Typ (iii), so ist offenbar
$$T(Q) = \{(y_1, \dots, y_k) : y_2 < y_1 \le y_2 + 1, \ 0 < y_j \le 1 \ (j \neq 2)\}$$
Sei
$$E = T(Q) \cap \{(y_1, \dots, y_k) : y_1 \le 0\},$$
$$F = T(Q) \cap \{(y_1, \dots, y_k) : y_1 > 0\}.$$
Nun ist $E + (1,0,\dots,0) \cup F = Q$ (disjunkte Vereinigung), daher
$$\det T = 1 = \lambda^k(Q) = \lambda^k(E + (1,0,\dots,0)) + \lambda^k(F) = \lambda^k(E) + \lambda^k(F)$$
$$= \lambda^k(E \cup F) = \lambda^k(T(Q)).$$



Daraus ergibt sich nun
$$\lambda^k(T(Q)) = \lambda^k(T_1(T_2(\cdots T_s(Q)\cdots))) = \lambda^k(T_1(Q)) \cdot (\lambda^k(T_2(\cdots T_s(Q)\cdots)))$$
$$= |\det T_1| \cdot (\lambda^k(T_2(\cdots T_s(Q)\cdots))) = \cdots$$
$$= |\det T_1| \cdot \ldots \cdot |\det T_s| \cdot \lambda^k(Q) = |\det(T_1 \cdots T_s)|$$
$$= |\det T|.$$

4.5 Satz Für $A \in Bor(\mathbb{R}^k)$ gilt

$$\lambda^k(A) = \inf\{\lambda^k(O) : A \subset O,\ O \text{ offen}\}$$
$$= \sup\{\lambda^k(C) : C \subset A,\ C \text{ kompakt}\}$$

(Wegen 3.14 gilt dieser Satz auch für $A \in Leb(\mathbb{R}^k)$.) 4.5 kann a[uf] ähnlichem Wege wie 3.14 gezeigt werden (tu es!), soll hier ab[er] als Spezialfall eines allgemeinen Satzes gewonnen werden.

4.6 Satz $\mu : Bor(\mathbb{R}^k) \to [0,\infty]$ sei ein Maß, das auf kompak[ten] Mengen endlich ist. (Insbesondere ist μ σ-endlich.) Dann gilt

$$\mu(A) = \inf\{\mu(O) : A \subset O,\ O \text{ offen}\}$$
$$= \sup\{\mu(C) : C \subset A,\ C \text{ kompakt}\}$$

für alle $A \in Bor(\mathbb{R}^k)$.

Beweis: Wir nehmen zunächst an, daß μ sogar ein endliches Maß i[st], also $\mu(\mathbb{R}^k) < \infty$ gilt. Für so ein Maß betrachten wir di[e] (ziemlich) guten Mengen

$$\mathcal{D} = \Big\{A \in Bor(\mathbb{R}^k) : \mu(A) = \inf\{\mu(O) : A \subset O,\ O \text{ offen}\} = \sup\{\mu(C) . C \subset A,\ C \text{ abgeschlossen}\}\Big\}$$

und zeigen $\mathcal{D} = Bor(\mathbb{R}^k)$, was schwächer ist als in 4.6 beh[auptet]. Dazu benötigen wir

a) $\mathcal{D}$ ist ein Dynkinsystem (Def. 3.9)

(i) $\emptyset \in \mathcal{D}$, $\mathbb{R}^k \in \mathcal{D}$, denn diese Mengen sind offen [und] abgeschlossen zugleich.

(ii) $A \in \mathcal{D} \Rightarrow A^c \in \mathcal{D}$: das folgt durch Komplementbildung (O offen $\Leftrightarrow$ O^c abgeschlossen) mit Hilfe der Endlichkeit von μ.

(iii) Seien $A_1, A_2, \dots \in \mathcal{D}$ paarweise disjunkt, $A := \bigcup_{i=1}^{\infty} A_i$, $\varepsilon > 0$.
Wähle offene $O_i \supset A_i$ mit $\mu(O_i \setminus A_i) \le 2^{-i} \cdot \varepsilon$. Dann ist $O := \bigcup_{i=1}^{\infty} O_i$ offen, $O \supset A$, und

$$\mu(O \setminus A) = \mu(O) - \mu(A)$$
$$\le \sum \mu(O_i) - \sum \mu(A_i) \qquad (2.3c)$$
$$= \sum \mu(O_i \setminus A_i)$$
$$\le \varepsilon$$

Andersherum wähle N mit $\mu(\bigcup_{i>N} A_i) = \sum_{i=N+1}^{\infty} \mu(A_i) < \frac{\varepsilon}{2}$ (möglich, da μ endlich), und zu A_i finde abgeschlossene $C_i \subset A_i$ mit $\mu(A_i \setminus C_i) \le 2^{-i} \cdot \varepsilon/2$.
$C := \bigcup_{i=1}^{N} C_i$ ist dann abgeschlossen, $C \subset A$, und

$$\mu(A \setminus C) = \sum_{i=1}^{N} \mu(A_i \setminus C_i) + \sum_{i>N} \mu(A_i)$$
$$\le \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

Damit ist $A \in \mathcal{D}$ gezeigt. //

b) Jede abgeschlossene Menge liegt in $\mathcal{D}$.

Sei C abgeschlossen. $d(x, C) = \inf_{y \in C} |x - y|$ definiert eine stetige Funktion, da nach der umgekehrten Dreiecksungleichung $|d(x_1, C) - d(x_2, C)| \le |x_1 - x_2|$ gilt. (Schreibe selbst den Beweis dafür hin!) Folglich ist

$$O_n := \{x . \ d(x, C) < \tfrac{1}{n}\}.$$

offen, $O_1 \supset O_2 \supset \cdots$, und $C = \bigcap_{n=1}^{\infty} O_n$, da C [illegible]
schlossen ist. (Beweis?) 2.4 impliziert $\mu(C) = \lim \mu(O_n)$,
also $\mu(C) = \inf\{\mu(O) : O \supset C,\ O \text{ offen}\}$.
Die andere Approximation ist trivial, da C selbst abgeschlossen [illegible]

c) $\mathcal{D} = \mathrm{Bor}(\mathbb{R}^k)$

3.10 impliziert $\mathrm{Bor}(\mathbb{R}^k) = d(\mathcal{C})$ (= das von den abgeschlossenen Mengen erzeugte Dynkinsystem), denn $\mathrm{Bor}(\mathbb{R}^k) = \sigma(\mathcal{C})$ nach 1.5. Wegen a) und b) ist aber

$$d(\mathcal{C}) \subset \mathcal{D} \subset \mathrm{Bor}(\mathbb{R}^k).$$

Der Beweis des Satzes für endliche μ ist daher mit

d) Für jedes abgeschlossene $C \subset \mathbb{R}^k$ gilt

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

erbracht.

Wähle eine aufsteigende Folge von Kompakta K_n mit $\mathbb{R}^k = \bigcup$ [illegible] (z.B. K_n = Kugel um 0 mit Radius n). Dann ist $C \cap K_n$ eine kompakte Menge, und 2.4 liefert $\mu(C) = \lim \mu$ [illegible] $\left(C = \bigcup_{n=1}^{\infty} (C \cap K_n)\ !\right)$.

Der allgemeine Fall soll jetzt auf den Fall endlicher Maße zurückgeführt [illegible] Dazu betrachte K_n wie im Beweisteil d) und definiere endliche Maße μ_n durch $\mu_n(A) = \mu(A \cap K_n)$. Es gilt dann (2.4)

$$\lim \mu_n(A) = \mu(A)$$

für alle $A \in \mathrm{Bor}(\mathbb{R}^k)$.

Die Approximation von unten ist dann leicht: Wähle kompakte $C_n \subset A$ mit $\mu_n(A \setminus C_n) \to 0$. Dann ist

$$\mu(A) = \lim \mu_n(A) = \lim \mu_n(C_n) = \lim \mu(C_n \cap K_n),$$

und $C_n \cap K_n$ ist eine kompakte Teilmenge von A.
Für die Approximation von oben dürfen wir $\mu(A) < \infty$ annehmen (sonst ist man mit $O = \mathbb{R}^k$ schon fertig). Folgende Hilfsbehauptung ist nun am Platze:

e) Zu jedem kompakten $K \subset \mathbb{R}^k$ existiert offenes $U \supset K$ mit $\mu(U) < \infty$.

Jedes $x \in K$ hat eine Umgebung U_x, deren Abschluß kompakt ist (z.B. eine Kugel). Alle U_x $(x \in K)$ zusammen überdecken K, daher (Kompaktheit) reichen bereits endlich viele davon:

$$K \subset \bigcup_{i=1}^{n} U_{x_i} =: U,$$

und

$$\mu(U) \le \sum_{i=1}^{n} \mu(U_{x_i}) \le \sum_{i=1}^{n} \mu(\bar{U}_{x_i}) < \infty$$

nach Voraussetzung an μ (Kompakta haben endliches Maß).

Zu den obigen K_n wählen wir nun U_n gemäß e) und setzen $(B \in \mathrm{Bor}(\mathbb{R}^k))$ $\nu_n(B) = \mu(B \cap U_n)$. Zu $\varepsilon > 0$ findet man nach dem ersten Beweisteil offene $O_n \supset A$ mit $\nu_n(O_n \setminus A) \le \varepsilon \cdot 2^{-n}$. Für $O = \bigcup_{n=1}^{\infty} (O_n \cap U_n)$ gilt dann: O ist offen, $A \subset O$, denn

$$A = \bigcup_{n=1}^{\infty} (A \cap U_n) \subset \bigcup_{n=1}^{\infty} (O_n \cap U_n),$$

sowie

$$\begin{aligned}\mu(O \setminus A) &= \mu\Big(\bigcup_{n=1}^{\infty} O_n \cap U_n \cap A^c\Big) \\ &\le \sum_{n=1}^{\infty} \mu(O_n \cap U_n \cap A^c) \\ &= \sum_{n=1}^{\infty} \nu_n(O_n \setminus A) \\ &\le \varepsilon.\end{aligned}$$

Die in 4.6 ausgedrückte äußerst nützliche Approximationseigenschaft heißt Regularität oder Straffheit von μ. Sie wird uns später noch im allgemeinen Rahmen wiederbegegnen.

Ganz ohne Endlichkeitsvoraussetzung ist Satz 4.4 übrigens falsch; als Gegenbeispiel betrachte $\mu(A) =$ Anzahl der Punkte in $A \cap \mathbb{Q}$ auf Bo[…]

Zurück zum Lebesguemaß. Als nächstes werden mit Hilfe des Auswahlaxioms nicht-Lebesguemeßbare Mengen nachgewiesen.

4.7 Satz Es gibt eine nicht-Lebesguemeßbare Teilmenge von $[0,1]$.

Beweis: Auf $[0,1]$ führe die Äquivalenzrelation

$$x \sim y \iff x - y \in \mathbb{Q}$$

ein. (Das ist wirklich eine Äquivalenzrelation.) Nach dem Auswahlaxiom gibt eine Teilmenge $A \subset [0,1]$, die aus jeder Äquivalenzklasse genau einen Vertreter enthält. Ist $\{r_1, r_2, r_3, \dots\}$ eine Abzählung von $\mathbb{Q} \cap [-1,1]$, so gilt offenbar

$$[0,1] \subset \bigcup_{n=1}^{\infty} (r_n + A) \subset [-1,2],$$

und die Vereinigung ist disjunkt. Wäre $A \in \mathrm{Leb}(\mathbb{R})$, so folgte aus der Translationsinvarianz (4.2) $3 \geq \lambda\left(\bigcup_n (r_n + A)\right) = \sum_n \lambda(r_n + A) = \sum_n \lambda(A)$, woraus $\lambda(A) = 0$ folgt. Das widerspricht jedoch

$$1 \leq \lambda\left(\bigcup_n (r_n + A)\right) = 0.$$

Mit gleichem Beweis zeigt man, daß jede meßbare Menge in $\mathbb{R}^k$ positiven Maßes eine nicht-meßbare Teilmenge hat. (Gemeint ist jeweils „Lebesgue-meßbar".)

Die Methode des Beweises von 4.7 zeigt sogar: Es gibt kein translationsinvariantes Maß, das auf ganz $\mathcal{P}(\mathbb{R}^k)$ definiert ist und Intervallen ihren üblichen Inhalt zuordnet. Verzichtet man jedoch auf die σ-Additivität zugunsten der bloßen Additivität, so ist dieses Inhaltsproblem lösbar.*) Verlangt man hingegen, daß kongruente Mengen denselben Inhalt zugeordnet bekommen, so existieren Lösungen dieses Inhaltsproblems in $\mathcal{P}(\mathbb{R})$ und $\mathcal{P}(\mathbb{R}^2)$ (Banach), nicht jedoch für $\mathcal{P}(\mathbb{R}^k)$ mit $k \geq 3$ (Hausdorff)!**)

Als nächstes begründen wir $\mathrm{Bor}(\mathbb{R}) \neq \mathrm{Leb}(\mathbb{R})$. Dazu produzieren wir eine überabzählbare Menge C vom Maß 0. Da dann jede Teilmenge von C Lebesgue-meßbar ist (3.13), gibt es mehr Lebesgue- als Borelmengen, vgl. S. 14 f.. (Auch dieses Argument ist nicht konstruktiv!)

Die gewünschte Menge ist die berühmte Cantormenge. Sie wird so konstruiert: Aus $[0,1]$ entferne das offene mittlere Drittel $]\frac{1}{3}, \frac{2}{3}[=: O_1$. Aus den beiden Restintervallen entferne wiederum die offenen mittleren Drittel $O_2 :=]\frac{1}{9}, \frac{2}{9}[$, $O_3 :=]\frac{7}{9}, \frac{8}{9}[$. Aus den noch verbleibenden 4 Restintervallen werden wieder die offenen mittleren Drittel $O_4, \dots, O_7$ entfernt etc. Was übrig bleibt, ist die Cantormenge C:

$$C = [0,1] \setminus \bigcup_{i=1}^{\infty} O_i$$

Bleibt überhaupt etwas übrig? Scharfes Hinsehen zeigt, daß C genau aus den Zahlen besteht, die in der Entwicklung im Dreiersystem ohne die Ziffer 1 geschrieben werden können (etwa $\frac{1}{3} = 0.02222\dots$). Mit anderen Worten, die Abbildung

$$\varphi : \{0,2\}^{\mathbb{N}} \to C, \quad (a_n)_{n \in \mathbb{N}} \mapsto \sum_{n=1}^{\infty} a_n \cdot 3^{-n}$$

ist bijektiv. C ist also überabzählbar.

C ist auch kompakt, da $\bigcup_{i=1}^{\infty} O_i$ offen ist, und enthält kein

*) vgl. Hewitt/Ross, Abstract Harmonic Analysis, vol I, §17

**) vgl. Floret, S. 119 ff.

Intervall (dieses müßte „auf jeder Stufe" in einem der Restintervalle liegen, deren Längen jedoch eine Nullfolge bilden). Man sagt, C sei *nirgends dicht*. Letztendlich gilt $\lambda(C) = 0$, da

$$\lambda(O_1) = \tfrac{1}{3}, \quad \lambda(O_2) = \lambda(O_3) = \tfrac{1}{9}, \quad \lambda(O_4) = \dots = \lambda(O_7) = \tfrac{1}{27}, \dots$$

also

$$\begin{aligned} \lambda(C) &= 1 - \sum_{i=1}^{\infty} \lambda(O_i) \\ &= 1 - \left(\tfrac{1}{3} + \tfrac{2}{9} + \tfrac{4}{27} + \dots\right) \\ &= 1 - \tfrac{1}{3}\left(1 + \tfrac{2}{3} + \tfrac{4}{9} + \dots\right) \\ &= 1 - \tfrac{1}{3} \cdot \frac{1}{1-\frac{2}{3}} = 0. \end{aligned}$$

Zusammengefaßt:

4.8 Beispiel. Es gibt eine überabzählbare, kompakte, nirgends dichte Teilmenge von $[0,1]$ vom Maß 0.

Die Cantormenge ist noch für mancherlei andere Überraschung gut!

Abschließend soll das Beispielreservoir für Maße weiter angefüllt werden. Das Lebesguemaß ist das Maß auf $\mathrm{Bor}(\mathbb{R})$, das von der üblichen Längenmessung abgeleitet wird. Was passiert nun bei einer gewichteten Längenmessung? Präziser sei $F: \mathbb{R} \to \mathbb{R}$ eine monoton wachsende Funktion, die rechtsseitig stetig ist, d.h.

$$\lim_{\substack{x \to x_0 \\ x > x_0}} F(x) = F(x_0) \qquad \forall\, x_0 \in \mathbb{R}.$$

Einem halboffenen Intervall werde die gewichtete Länge

$$(*) \qquad \mu_F(\,]a,b]\,) = F(b) - F(a)$$

zugeordnet.

4.9 Satz Es gibt genau ein Maß μ_F auf $\mathrm{Bor}(\mathbb{R})$, für das $(*)$ gilt. Es heißt das von F erzeugte *Lebesgue-Stieltjes-Maß*.

Beweis: Der Beweis ist ähnlich wie für λ:

1. Schritt: μ_F erzeugt einen Inhalt auf $\mathcal{F}^1$.

2. Schritt: Dieser Inhalt ist sogar ein (σ-endliches!) Prämaß.

[Der gleiche Beweis wie für 2.5 geht durch. Damit der „Verkleinerungstrick" dort klappt, ist die *rechtsseitige* Stetigkeit entscheidend!

3. Schritt: Wende den Fortsetzungssatz von Carathéodory an (3.6 & 3.12)!

Umgekehrt ist ein Maß auf $\mathrm{Bor}(\mathbb{R})$, das auf kompakten Mengen endlich ist, ein Lebesgue-Stieltjes-Maß: So einem μ ordne

$$\begin{aligned} F(x) &= \mu(\,]0,x]\,) && x \geq 0 \\ &= -\mu(\,]x,0]\,) && x < 0 \end{aligned}$$

zu.[*] (Die rechtsseitige Stetigkeit folgt aus 2.4.)

Es gilt: $\quad \mu_F(\{x_0\}) = 0 \iff F$ stetig bei x_0,

allgemeiner $\quad \mu_F(\{x_0\}) = F(x_0) - \lim\limits_{\substack{x \to x_0 \\ x < x_0}} F(x)$

(Beweis zur Übung; der Limes existiert wegen der Monotonie von F.)

Bei einem Lebesgue-Stieltjes-Maß ist also i.a. zwischen

$$\mu(\,]a,b[\,), \quad \mu(\,[a,b[\,), \quad \mu(\,[a,b]\,) \quad \text{und} \quad \mu(\,]a,b]\,)$$

zu unterscheiden!

---

[*] Die maßerzeugende Funktion ist bis auf eine Konstante eindeutig bestimmt (Beweis?).

4.10 Beispiele a) $F(x) = x$ liefert das Lebesguemaß.

b) $F: \mathbb{R} \to \mathbb{R}$ sei monoton wachsend und stetig differenzierbar.

Nach dem Hauptsatz der Diff.-Int.-Rechnung gilt $(f := F')$

$$\mu_F(]a,b]) = \int_a^b f(x)\,dx \qquad \text{(Riemann-Integral).}$$

c) Ein für die Wahrscheinlichkeitstheorie wichtiger Spezialfall ist

$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$; μ_F heißt (standardisierte) Normalverteilung.

d) $F(x) = 0$ für $x < x_0$, $F(x) = 1$ für $x \geq x_0$

liefert das Dirac-Maß δ_{x_0} (Bsp. 2.2.c)) auf $\mathcal{B}or(\mathbb{R})$.

e) Sei μ das endliche Maß $\mu = \sum_{n=1}^{\infty} 2^{-n} \delta_{r_n}$, wo (r_n) ein Aufzählung von $\mathbb{Q}$ ist. Aus der obigen Bemerkung ergibt sich für die μ erzeugende Funktion F:

$$F \text{ stetig bei } x_0 \iff x_0 \text{ irrational} \qquad (!)$$

f) Es sei an die Konstruktion der Cantormenge erinnert (Bsp. 4.8).

Mit den obigen Bezeichnungen sei für $x \in \bigcup_{i=1}^{\infty} O_i$

$$\tilde{F}(x) = (2j+1) \cdot 2^{-(n+1)} \quad \text{falls } x \in O_{2^n+j} \quad (n \geq 0, j = 0, \ldots$$

(also $\tilde{F}|_{O_1} = \frac{1}{2}$, $\tilde{F}|_{O_2} = \frac{1}{4}$, $\tilde{F}|_{O_3} = \frac{3}{4}$, $\tilde{F}|_{O_4} = \frac{1}{8}$ etc.)

$F: \mathbb{R} \to \mathbb{R}$ wird durch

$$F(x) = \begin{cases} 0 & x \leq 0 \\ \sup\limits_{y \leq x} \tilde{F}(y) & x > 0 \end{cases}$$

definiert:

Es ist nicht schwer zu sehen, daß F monoton wachsend und stetig ist. F und das zugehörige LS-Maß haben einige spektakuläre Eigenschaften. Hier eine davon: F bildet die λ-Nullmenge C auf $[0,1]$ ab! Das stetige Bild einer λ-Nullmenge braucht also keine λ-Nullmenge zu sein!

Lebesgue-Stieltjes-Maße im $\mathbb{R}^k$ und ihre maßerzeugenden Funktionen werden z.B. bei Ash oder Billingsley besprochen.

II. Integration

II.1 Meßbare Funktionen

Zuerst eine Sprechweise: Ist S eine nichtleere Menge und $\mathfrak{A} \subset \mathfrak{P}(S)$ eine σ-Algebra, so heißt das Paar $(S, \mathfrak{A})$ ein <u>meßbarer Raum</u>. (Es ist hier noch nicht von einem Maß auf $\mathfrak{A}$ die Rede.)

<u>1.1 Definition</u> $(S_1, \mathfrak{A}_1)$ und $(S_2, \mathfrak{A}_2)$ seien meßbare Räume, $T : S_1 \to S_2$ eine Abbildung. T heißt meßbar (präziser: $\mathfrak{A}_1$-$\mathfrak{A}_2$-meßbar), falls gilt

$$T^{-1}(B) \ \left(= \{ s \in S_1 : T(s) \in B \}\right) \in \mathfrak{A}_1 \qquad \forall B \in \mathfrak{A}_2 .$$

Besonders wichtig ist der Fall $(S_2, \mathfrak{A}_2) = (\mathbb{R}, \mathrm{Bor}(\mathbb{R}))$. In diesem Fall spricht man von Borel-meßbaren Funktionen.

Beachte die Ähnlichkeit von Def. 1.1 mit der Definition der Stetigkeit von Abbildungen zwischen metrischen Räumen!

Das folgende Lemma wird es ermöglichen, eine große Anzahl von Beispielen anzugeben.

<u>1.2 Lemma</u> $(S_i, \mathfrak{A}_i)$ seien meßbare Räume $(i = 1,2,3)$.

a) $T : S_1 \to S_2$ ist genau dann meßbar, wenn
$$T^{-1}(B) \in \mathfrak{A}_1 \qquad \forall B \in \mathcal{E},$$
wo $\mathcal{E}$ ein Erzeuger der σ-Algebra $\mathfrak{A}_2$ ist (d.h. $\mathfrak{A}_2 = \sigma(\mathcal{E})$).

b) Sind $T_1 : S_1 \to S_2$ und $T_2 : S_2 \to S_3$ meßbar, so auch
$T_2 \circ T_1 : S_1 \to S_3$.

c) $T : S_1 \to \mathbb{R}^k$ ist genau dann Borel-meßbar, wenn alle $\pi_j \circ T : S_1 \to \mathbb{R}$ meßbar sind, wo $\pi_j : \mathbb{R}^k \to \mathbb{R}$ die j-te Projektion $(x_1, \dots, x_k) \mapsto x_j$ ist.

d) Stetige Abbildungen $T : \mathbb{R}^m \to \mathbb{R}^k$ sind Borel-Borel-meßbar.

e) $f : S_1 \to \mathbb{R}$ ist genau dann Borel-meßbar, wenn gilt
$$\{ s : f(s) \le r \} \in \mathfrak{A}_1 \qquad \forall r \in \mathbb{R} .$$

Beweis: a) Natürlich geht's mit dem "good sets principle": Setzt man $\mathcal{B} = \{ B \in \mathfrak{A}_2 : T^{-1}(B) \in \mathfrak{A}_1 \}$, so ist $\mathcal{B}$ eine σ-Algebra [fast trivial nachzuprüfen; der Witz ist, daß sich T^{-1} durch die mengentheoretischen Operationen $\cap, \cup, {}^c$ durchziehen läßt], die nach Voraussetzung einen Erzeuger von $\mathfrak{A}_2$ enthält. Konsequenz: $\mathcal{B} = \mathfrak{A}_2$.

b) $B \in \mathfrak{A}_3 \Rightarrow T_2^{-1}(B) \in \mathfrak{A}_2 \Rightarrow (T_2 \circ T_1)^{-1}(B) = T_1^{-1}(T_2^{-1}(B)) \in \mathfrak{A}_1$.

Zuerst nun zu d): Ist T stetig, so ist für offenes $B \subset \mathbb{R}^k$ $T^{-1}(B)$ offen, insbesondere eine Borelmenge. Die Behauptung folgt nun aus a), da $\mathrm{Bor}(\mathbb{R}^k)$ von den offenen Mengen erzeugt wird (I.1.5).

c) Da die π_j stetig sind, ergibt sich die Hinlänglichkeit aus b) und d).
— Seien nun alle $\pi_j \circ T$ meßbar. Für ein Intervall
$$B =]\alpha_1, \beta_1] \times \dots \times]\alpha_k, \beta_k] = \bigcap_{j=1}^{k} \pi_j^{-1}(]\alpha_j, \beta_j])$$
gilt dann
$$T^{-1}(B) = \bigcap_{j=1}^{k} (\pi_j \circ T)^{-1}(]\alpha_j, \beta_j]) \in \mathfrak{A}_1 ,$$

und nach Teil a) ist T meßbar, da ja die Intervalle $\mathrm{Bor}(\mathbb{R}^k)$ erzeugen.
e). schließlich ist ein Spezialfall von a); I.1.5 erlaubt, weitere Spezialfälle zu formulieren (die wir im Folgenden auch stillschweigend verwenden werden).

Formulieren wir einige einfache Beispiele:

a) Konstante Funktionen sind stets meßbar.

b) Die Indikatorfunktion 1_A, definiert durch $1_A(s) = 1$ falls $s \in A$, $= 0$ sonst, ist genau dann meßbar, wenn $A \in \mathfrak{A}$ ist.*) Offenbar ist das der einfachste Typ einer meßbaren Funktion. Interessanterweise läßt sich jede meßbare Funktion aus ihnen zusammensetzen (vgl. 1.8).

c) Sei $x_0 \in \mathbb{R}^k$ fest und $T : \mathbb{R}^k \to \mathbb{R}^k$ die Translationsabbildung $T(y) = y - x_0$. Da offenbar $T^{-1}(A) = x_0 + A$ ist, erhalten wir aus 1.2. d) einen neuen Beweis für Satz I.1.7.
[So neu ist der Beweis freilich nicht: eigentlich ist es haargenau der Beweis aus Kap. I, nur mit den neuen Begriffen formuliert.]

Im weiteren soll der wichtigen Konvention gefolgt werden, daß $\mathbb{R}^k$ – sofern nichts anderes angedeutet wird – stets mit der Borel-σ-Algebra versehen wird. Der Meßbarkeitsbegriff bezieht sich also immer auf $(\mathbb{R}^k, \mathrm{Bor}(\mathbb{R}^k))$.

*) statt „$A \in \mathfrak{A}$" ist daher auch die Ausdrucksweise „A ist meßbar" geläufig.

Die Zusammensetzung meßbarer Funktionen liefert wieder meßbare Funktionen:

1.3 Satz $(S, \mathfrak{A})$ sei ein meßbarer Raum, $f, g : S \to \mathbb{R}$ seien meßbar. $f+g$, $f-g$, $f \cdot g$, $\max(f,g)$, $\min(f,g)$, $f^+ (= \max(f,0))$, $f^- (= \max(-f,0))$, $|f|$, αf $(\alpha \in \mathbb{R})$ und $\frac{1}{f}$ (sofern stets $f(s) \neq 0$ ist) sind dann ebenfalls meßbar.
Insbesondere bilden die meßbaren Funktionen einen Vektorraum.

Beweis: Zu $+$: Die Abbildung $\varphi : \mathbb{R}^2 \to \mathbb{R}$, $(x,y) \mapsto x+y$ ist stetig, und nach 1.2 c) ist $F : S \to \mathbb{R}^2$, $F(s) = (f(s), g(s))$ meßbar, daher ist (1.2.d+b) das Kompositum $\varphi \circ F = f+g$ ebenfalls meßbar.
Genauso funktionieren die Beweise für $-, \cdot, \max, \min$. ($\max(x,y) = \frac{x+y}{2} + \frac{|x-y|}{2}$ ist stetig auf $\mathbb{R}^2$!)
f^+, f^- und αf sind Spezialfälle davon ($g = 0$ bzw. $g = \alpha$), und $|f| = f^+ + f^-$ (oder $|f| = |\cdot| \circ f$, wenn das lieber ist), womit fast alle Behauptungen gezeigt sind. $\frac{1}{f}$ bleibt zur Übung.

Wegen 1.2 e) tauchen Mengen der Form $\{s : f(s) \leq r\}$ bzw. allgemeiner $\{s : f(s) \in B\}$ häufig bei der Diskussion der Meßbarkeit auf. Für solche Mengen werden wir in Zukunft abkürzend

$$\{f \leq r\} \quad \text{bzw.} \quad \{f \in B\}$$

schreiben. Entsprechend ist $\{f = g\}$ $(= \{s : f(s) = g(s)\})$ etc. zu verstehen.

Insbesondere folgt aus Satz 1.3, daß für meßbare Funktionen

$\{f = g\}$, $\{f \leq g\}$ etc. meßbar sind (d.h. in $\mathfrak{A}$ liegen). Beweis?

Wenden wir uns nun Folgen meßbarer Funktionen zu. Folgende Erweiterung des Begriffs der Borel-Meßbarkeit erweist sich dabei als praktisch: Man läßt nun $[-\infty, +\infty]$ wertige Funktionen zu und versieht $[-\infty, +\infty]$ mit der Borel-σ-Algebra aller Teilmengen $A \cup E$, wo $A \subset \mathbb{R}$ und borelsch sowie $E \subset \{-\infty, +\infty\}$ ist. Es ist leicht zu sehen, daß das wirklich eine σ-Algebra ist, die z.B. von den Intervallen $[-\infty, r]$ $(r \in \mathbb{R})$ erzeugt wird. Eine Funktion $f: S \to \overline{\mathbb{R}}$ ist daher genau dann meßbar, wenn

$$\{-\infty \leq f \leq r\} \in \mathfrak{A} \qquad \forall\, r \in \mathbb{R}$$

ist. (Beweis?)

Der Grund für die Einführung $\overline{\mathbb{R}}$-wertiger Funktionen liegt hauptsächlich darin, daß nun die Funktion

$$(\sup_n f_n)(s) = \sup_n f_n(s)$$

stets definiert ist.

Außer den bereits getroffenen Vereinbarungen (S. 18) über die Addition von ∞ benötigen wir noch

$$a - \infty = -\infty \qquad \text{für } a \in \mathbb{R} \text{ oder } a = -\infty$$

$$a \cdot \infty = \infty \qquad \text{für } 0 < a \leq \infty$$

$$a \cdot \infty = -\infty \qquad \text{für } 0 > a \geq -\infty$$

sowie

$$0 \cdot \infty = 0.$$

Der Ausdruck $\infty - \infty$ ist verboten, daher ist für $\overline{\mathbb{R}}$-wertige Funktionen $f - g$ nicht unbedingt überall definiert.

Mit diesen Vereinbarungen gilt Satz 1.3 entsprechend.

1.4 Satz (f_n) sei eine Folge auf einem meßbaren Raum definierter Funktionen nach $\overline{\mathbb{R}}$. Dann sind die punktweise definierten Funktionen

$$\sup f_n, \quad \inf f_n, \quad \limsup f_n, \quad \liminf f_n$$

ebenfalls meßbar. Falls $f := \lim_{n \to \infty} f_n$ punktweise existiert, ist f meßbar.

Beweis: Zu sup: $\{-\infty \leq \sup f_n \leq r\} = \bigcap_{n=1}^{\infty} \{-\infty \leq f_n \leq r\}$.

inf behandelt man analog. Daraus folgt die Behauptung über lim sup, da ja $\limsup f_n = \inf_k \sup_{n \geq k} f_n$ ist.

lim inf geht wieder analog, und im Fall seiner Existenz ist natürlich $\lim = \limsup = \liminf$.

Im übrigen ist $\{\lim f_n \text{ existiert}\}$ stets meßbar; das folgt durch Anwendung von

$$g, h: S \to \overline{\mathbb{R}} \text{ meßbar} \overset{(*)}{\Rightarrow} \{g = h\} \in \mathfrak{A}$$

auf $g = \liminf f_n$, $h = \limsup f_n$.

Der Beweis von (*) folgt nicht unmittelbar aus 1.3, da evtl. $g - h$ nicht definiert ist. Die nachstehende Beweisvariante ist sehr instruktiv: Offenbar genügt es, $\{g < h\} \in \mathfrak{A}$ zu zeigen, und $\{g < h\}$ kann als abzählbare (!!) Vereinigung

$$\{g < h\} = \bigcup_{r \in \mathbb{Q}} (\{-\infty \leq g < r\} \cap \{r < h \leq \infty\})$$

geschrieben werden.

Damit kann ein Versprechen aus dem I. Kapitel eingelöst werden.

<u>1.5 Beispiel:</u> In Beispiel I. 1.6 war allen $x \in]0,1]$ ihre ni
abbrechende dyadische Entwicklung zugeordnet worden:

$$x = 0.\, d_1(x)\, d_2(x)\, d_3(x) \cdots$$

Die d_i sind dann meßbare Funktionen auf $]0,1]$, denn
$\{d_i \leq r\} = \emptyset$ für $r < 0$, $=]0,1]$ für $r \geq 1$, u
für $0 \leq r < 1$ ist

$$\{d_i \leq r\} = \{d_i = 0\}$$
$$=]0, 2^{-i}] \cup]2\cdot 2^{-i}, 3\cdot 2^{-i}] \cup \cdots \cup]1-2\cdot 2^{-i}, 1$$

Folglich sind auch $S_n := \frac{1}{n} \sum_{i=1}^{n} d_i$ meßbare Funktionen.
Ferner ist $\Omega := \{\lim_{n\to\infty} S_n \text{ existiert}\}$ meßbar, und die
Menge N der normalen Zahlen ist

$$\{x \in \Omega : \lim S_n(x) = \tfrac{1}{2}\}$$

ist eine Borelmenge.

Es ist also sinnvoll, nach $\lambda(N)$ zu fragen. Das starke Gesetz d
großen Zahlen liefert $\lambda(N) = 1$, siehe Kap. V.

S. 147 ff.

<u>1.6 Beispiel</u> $f : [0,1] \times [0,1] \to \mathbb{R}$ sei in jeder Variablen separat
stetig. Dann ist f meßbar. (Natürlich betrachten wir $\mathrm{Bor}([0,1]^2)$
$= \mathrm{Bor}(\mathbb{R}^2) \cap [0,1]^2$.) Zum Beweis betrachte zu $n \in \mathbb{N}$ die Funktion

$$f_n(x,y) = n \cdot \left(f\left(\tfrac{k}{n}, y\right) \cdot \left(x - \tfrac{k}{n}\right) + f\left(\tfrac{k+1}{n}, y\right) \left(\tfrac{k+1}{n} - x\right)\right)$$

falls $x \in [\frac{k}{n}, \frac{k+1}{n}]$. Man bestätigt, daß f_n (als Funktion vo
zwei Variablen) stetig ist und $f_n(x,y) \xrightarrow[n\to\infty]{} f(x,y)$ $\forall (x,y) \in [0,1$

gilt. Daher ist f meßbar.

(Das Beispiel stammt aus der 1. Publikation von Lebesgue aus dem Jahre 1898.)

Die folgende Definition und Satz 1.8 sind absolut fundamental für den Aufbau der Integrationstheorie.

<u>1.7 Definition</u> Sei $(S, \mathfrak{A})$ ein meßbarer Raum. Eine Treppenfunktion (genauer: $\mathfrak{A}$-Treppenfunktion) ist eine meßbare Funktion von S nach $\mathbb{R}$, die nur endlich viele Werte annimmt.

Jede Treppenfunktion f läßt sich also in der Form $\sum_{i=1}^{n} \alpha_i 1_{A_i}$ schreiben, wo $A_i = \{f = \alpha_i\}$ ist und $\alpha_1, \dots, \alpha_n$ die verschiedenen Werte von f durchläuft. Natürlich gibt es (i.a. unendlich viele) weitere Darstellungen von f als Summe von Indikatorfunktionen; z.B.

$$1_{]0,1]} + \tfrac{3}{2}\, 1_{]1,2]} + \tfrac{1}{2}\, 1_{]2,3]} = 1_{]0,2]} + \tfrac{1}{2}\, 1_{]1,3]}.$$

Beachte, daß im Fall $S = \mathbb{R}$ die „Stufen" einer Treppenfunktion wesentlich allgemeiner als Intervalle sein dürfen ($1_{\mathbb{Q}}$ ist noch ein recht harmloses Beispiel).

Nach 1.4 ist jeder punktweise Limes von Treppenfunktionen meßbar. Interessanterweise hat umgekehrt jede meßbare Funktion diese Gestalt.

1.8 Satz $(S, \mathcal{A})$ sei ein meßbarer Raum.

a) $f: S \to [0,\infty]$ sei meßbar. Dann existiert eine Folge von Treppenfunktionen (f_n) mit

$$0 \leq f_1(s) \leq f_2(s) \leq \cdots < \infty$$

und $\lim_{n\to\infty} f_n(s) = f(s)$

für alle $s \in S$.

b) Jede meßbare Funktion ist punktweiser Grenzwert einer Folge von Treppenfunktionen.

c) Für beschränkte Funktionen kann jeweils gleichmäßige Konvergenz erzielt werden.

Beweis: a) Zu $n \in \mathbb{N}$ setze

$$f_n(s) = \sum_{i=0}^{4^n-1} i \cdot 2^{-n} \cdot 1_{\{i\cdot 2^{-n} \leq f < (i+1)2^{-n}\}} + 2^n 1_{\{f \geq 2^n\}}$$

(Eine Skizze zeigt, was f_n tut!)

Nach Konstruktion ist (f_n) monoton wachsend, und für hinreichend großes n ist bei gegebenem s $|f_n(s) - f(s)| \leq 2^{-n}$ (nämlich sobald $2^n > f(s)$) bzw. $f_n(s) \geq 2^n$ (falls $f(s) = \infty$).

Damit ist für diesen Fall c) mit erledigt.

b) kann wegen $f = f^+ - f^-$ auf a) zurückgeführt werden.

(Zur Erinnerung: $f^+ = \max(f, 0)$, $f^- = \max(-f, 0)$ ≥ 0.)

c) ist bereits mitbewiesen.

Der folgende Satz zeigt eine überraschende Stetigkeitseigenschaft meßbarer Funktionen auf $[a,b]$.

1.9 Satz $f: [a,b] \to \mathbb{R}$ sei meßbar. Zu $\varepsilon > 0$ existiert dann eine kompakte Teilmenge $K \subset [a,b]$ mit

- $\lambda([a,b] \setminus K) \leq \varepsilon$
- die Restriktion von f auf K ist eine stetige Funktion von K nach $\mathbb{R}$.

Dieser Satz heißt in der Literatur meistens Satz von Lusin, obwohl er ca. 7 Jahre vor Lusin bereits von Vitali entdeckt wurde (1905).

Man darf in diesen Satz nicht mehr hineinlesen als drinsteht: Es ist nämlich nicht behauptet, daß f in den Punkten von K stetig ist (als Funktion auf $[a,b]$), sondern nur, daß f als auf K definierte Funktion stetig ist. (Teste dein Verständnis des Satzes am Beispiel der Dirichletschen Sprungfunktion!)

Mit Hilfe des Satzes von Urysohn aus der Topologie kann der Satz von Tietze Lusin auch so formuliert werden:

Zu $\varepsilon > 0$ existieren ein Kompaktum $K \subset [a,b]$ und eine stetige Funktion $g: [a,b] \to \mathbb{R}$ mit

- $\lambda([a,b] \setminus K) \leq \varepsilon$
- $f(x) = g(x) \quad \forall x \in K$.

Beweis von 1.9: O.E. ist f beschränkt (sonst betrachte $\arctan f$) und ≥ 0 (sonst betrachte $f + \text{const.}$). Für so eine Funktion beweisen wir zunächst eine Hilfsbehauptung:

$$(*)\quad \begin{cases} \text{Zu kompakten } F \subset [a,b],\ \varepsilon^* > 0 \text{ und } \delta > 0 \text{ existiert kompaktes} \\ \tilde F \subset F,\quad \lambda(F \setminus \tilde F) \le \varepsilon^*, \text{ mit} \\ \forall_{x \in \tilde F}\ \exists_{\eta > 0}\ \sup\limits_{\substack{|y-x|<\eta \\ y \in \tilde F}} f(y) - \inf\limits_{\substack{|y-x|<\eta \\ y \in \tilde F}} f(y) \le \delta. \end{cases}$$

⌈ Für $n \ge 0$ setze $A_n = \{ n\delta \le f < (n+1)\delta \} \cap F$.

Für genügend große N ist dann

$$\bigcup_{n=0}^{N} A_n = F, \quad \text{also} \quad \sum_{n=0}^{N} \lambda(A_n) = \lambda(F).$$

Da λ „regulär" ist (Satz I.4.5), existieren kompakte $C_n \subset A_n$ mit $\lambda(A_n \setminus C_n) \le 2^{-n} \cdot \frac{\varepsilon^*}{2}$, folglich ist für $\tilde F = \bigcup_{n=0}^{N} C_n$

$$\lambda(F \setminus \tilde F) = \sum_{n=0}^{N} \lambda(A_n \setminus C_n) \le \varepsilon^*.$$

Wegen der Kompaktheit der C_n ist auch $\tilde F$ kompakt, und es gilt $\eta := \min\limits_{n \le N} \operatorname{dist}(C_n, \tilde F \setminus C_n) > 0$.

(Der Abstand zweier Teilmengen ist bekanntlich

$$\operatorname{dist}(E_1, E_2) = \inf_{y \in E_1} d(y, E_2), \quad \text{vgl. S. 45 unten.)}$$

Nach Konstruktion tut η das, was verlangt war. ⌋

Die Hilfsbehauptung wird nun unendlich oft angewandt.

Setze $\varepsilon_n = \varepsilon \cdot 2^{-n}$, $\delta_n = 2^{-n}$ und $F_1 = [a,b]$. $(*)$ mit $F_1, \varepsilon_1, \delta_1$ angewandt beschert uns $\tilde F =: F_2$ wie beschrieben; $(*)$ mit $F_2, \varepsilon_2, \delta_2$ angewandt liefert F_3 etc. Für $K = \bigcap_{n=1}^{\infty} F_n$ gilt dann

$$\lambda([a,b] \setminus K) = \sum_{n=1}^{\infty} \lambda(F_n \setminus F_{n+1}) \le \sum_{n=1}^{\infty} \varepsilon_n = \varepsilon,$$

K ist kompakt (als Schnitt kompakter Mengen), und die Restriktion von f auf K ist nach Konstruktion stetig.

Abschließend eine Warnung: Für stetige $f: \mathbb{R}^k \to \mathbb{R}^m$ ist das Urbild einer Borelmenge wieder eine Borelmenge. Hingegen braucht das stetige Bild einer Borelmenge keine Borelmenge zu sein! Das ist gewissermaßen der GAU der Maßtheorie, Laurent Schwartz nennt dieses Phänomen die „Bildmaßkatastrophe". Das erste Gegenbeispiel wurde 1917 von Souslin produziert (es ist erwartungsgemäß ziemlich kompliziert), der damit das Studium der sog. analytischen Mengen eröffnete. Die explizite Konstruktion einer solchen nicht-borelschen Menge ist z.B. bei Behrends (Anhang) beschrieben, vgl. auch Cohn, Chap. 8.

II.2 Integrierbare Funktionen

Gegeben sei ein meßbarer Raum $(S, \mathfrak{A})$ und ein Maß μ auf $\mathfrak{A}$. (Das Tripel $(S, \mathfrak{A}, \mu)$ wird dann ein Maßraum genannt.) Das Integral einer meßbaren Funktion f wird in 3 Schritten eingeführt:

- für positive Treppenfunktionen,
- für meßbare positive Funktionen,
- für meßbare Funktionen.

Sei zunächst $f = \sum_{i=1}^{n} \alpha_i A_i$ eine positive Treppenfunktion, wo $\alpha_1, \dots, \alpha_n \ge 0$ den Wertebereich von f durchläuft und $A_i = \{ f = \alpha_i \}$ $(\in \mathfrak{A})$ ist.

2.1 Definition Für $E \in \mathfrak{A}$ setze

$$\int_E f \, d\mu = \sum_{i=1}^{n} \alpha_i \, \mu(A_i \cap E) \quad (\in [0, \infty]).$$

Es sei hier an die Konvention $0 \cdot \infty = 0$ erinnert!

Wichtige Bemerkung: Es gibt an dieser Stelle kein Wohldefiniertheitsproblem, da das Integral von f aus einer festen, nämlich der „minimalen", Darstellung berechnet wird. (Vgl. jedoch S. 68 oben.)

2.2 Definition Sei $f: S \to [0,\infty]$ meßbar. Setze

$$\int_E f \, d\mu = \sup \left\{ \int_E g \, d\mu : 0 \le g \le f, \ g \text{ Treppenfunktion} \right\}$$

$$(\in [0,\infty])$$

Für Treppenfunktionen f stimmen beide Definitionen überein, wie aus Teil c) des nächsten Lemmas folgt (f trägt selbst zum Supremum in 2.2 bei).

2.3 Lemma g und h seien positive Treppenfunktionen.

a) $\nu: E \mapsto \int_E g \, d\mu$ definiert ein Maß auf $\mathfrak{A}$.

b) $\int_E g \, d\mu + \int_E h \, d\mu = \int_E (g+h) \, d\mu \quad \forall E \in \mathfrak{A}$

c) $h \ge g \Rightarrow \int_E h \, d\mu \ge \int_E g \, d\mu \quad \forall E \in \mathfrak{A}$

d) $c \ge 0 \Rightarrow \int_E c \cdot g \, d\mu = c \int_E g \, d\mu$.

Beweis: a) Sei $g = \sum_{i=1}^n \alpha_i 1_{A_i}$ die minimale Darstellung mit $A_i := \{g = \alpha_i\}$, (α_i) paarweise verschieden. Ferner seien $E_1, E_2, \dots \in \mathfrak{A}$ paarweise disjunkt. Nach Definition gilt dann $\nu(\emptyset) = 0$ sowie $(E := \bigcup_{j=1}^\infty E_j)$

$$\nu(E) = \sum_{i=1}^n \alpha_i \, \mu(A_i \cap E)$$

$$= \sum_{i=1}^n \alpha_i \sum_{j=1}^\infty \mu(A_i \cap E_j)$$

$$= \sum_{j=1}^\infty \sum_{i=1}^n \alpha_i \, \mu(A_i \cap E_j)$$

$$= \sum_{j=1}^\infty \nu(E_j).$$

b) h sei analog als $h = \sum_{j=1}^m \beta_j 1_{B_j}$ dargestellt. Für $E_{ij} = A_i \cap B_j \cap E$ gilt dann definitionsgemäß

$$\int_{E_{ij}} g \, d\mu = \alpha_i \, \mu(E_{ij})$$

$$\int_{E_{ij}} h \, d\mu = \beta_j \, \mu(E_{ij})$$

$$\int_{E_{ij}} (g+h) \, d\mu = (\alpha_i + \beta_j) \, \mu(E_{ij})$$

Addition liefert wg. a) und $E = \bigcup_{i,j} E_{ij}$ die Behauptung:

$$\int_E (g+h) \, d\mu = \sum_{i,j} \int_{E_{ij}} (g+h) \, d\mu$$

$$\underset{\text{s.o.}}{=} \sum_{i,j} \left(\int_{E_{ij}} g \, d\mu + \int_{E_{ij}} h \, d\mu \right)$$

$$= \int_E g \, d\mu + \int_E h \, d\mu.$$

c) $f := h - g$ ist eine positive Treppenfunktion, also ist definitionsgemäß $\int_E f \, d\mu \ge 0$. Daher mit b)

$$\int_E h \, d\mu = \int_E f \, d\mu + \int_E g \, d\mu$$

$$\ge \int_E g \, d\mu.$$

d) ist klar. ✓

Aus 2.3 b) folgt noch, daß bei jeder Darstellung $\sum_{i=1}^{\tilde{n}} \tilde{\alpha}_i 1_{\tilde{A}_i}$ einer positiven Treppenfunktion (mit $\tilde{\alpha}_i \geq 0$ und $\tilde{A}_i \in \mathfrak{A}$) das Integral durch $\sum_{i=1}^{\tilde{n}} \tilde{\alpha}_i \, \mu(\tilde{A}_i \cap E)$ zu berechnen ist!

Sammeln wir nun einige elementare Konsequenzen aus der Definition des Integrals einer positiven meßbaren Funktion. Die Beweise ergeben sich unmittelbar aus 2.2, da die entsprechenden Aussagen (trivialerweise oder nach 2.3) für Treppenfunktionen gelten.

<u>2.4 Lemma</u> $f, g : S \to [0,\infty]$ seien meßbar, $E, F \in \mathfrak{A}$.

a) $0 \leq f|_E \leq g|_E \quad \Rightarrow \quad \int_E f \, d\mu \leq \int_E g \, d\mu$

b) $E \subset F \quad \Rightarrow \quad \int_E f \, d\mu \leq \int_F f \, d\mu$

c) $f|_E = 0 \quad \Rightarrow \quad \int_E f \, d\mu = 0$

d) $\mu(E) = 0 \quad \Rightarrow \quad \int_E f \, d\mu = 0$

(selbst wenn $f|_E = \infty$ ist!!)

Der folgende Satz ist die Basis aller Konvergenzsätze, für den Beweis ist die σ-Additivität von μ entscheidend.

<u>2.5 Satz</u> Seien f und $f_1, f_2, \dots : S \to [0,\infty]$ meßbar und gilt
$f_1(s) \leq f_2(s) \leq \dots$

sowie
$$\lim_{n\to\infty} f_n(s) = f(s) \quad \text{für alle } s \in S.$$

Dann gilt ($E \in \mathfrak{A}$)
$$\lim_{n\to\infty} \int_E f_n \, d\mu = \int_E f \, d\mu .$$

Beweis: Die Meßbarkeit von f braucht natürlich nicht extra gefordert zu werden, sie ergibt sich automatisch nach 1.4.

Da die Folge (f_n) monoton wächst, wächst die Folge ihrer Integrale nach 2.4 a) ebenfalls, so daß
$$\varphi := \lim_{n\to\infty} \int_E f_n \, d\mu$$
in $[0,\infty]$ existiert. Wegen $f_n \leq f$ ist auch stets $\int_E f_n \, d\mu \leq \int_E f \, d\mu$, daher $\varphi \leq \int_E f \, d\mu$.

Um die umgekehrte Ungleichung zu zeigen, beachte zunächst, daß nach Konstruktion $\varphi = \sup_n \int_E f_n \, d\mu$ ist. Es reicht daher, für alle Treppenfunktionen $0 \leq g \leq f$ die Ungleichung
$$\int_E g \, d\mu \leq \sup_n \int_E f_n \, d\mu \quad (= \varphi)$$
zu beweisen.

Beweis hierfür:

Sei $0 \leq c < 1$ beliebig und
$$E_n = \{ s \in E : f_n(s) \geq c \cdot g(s) \}.$$

Offenbar gilt $E_1 \subset E_2 \subset \dots$, und alle $E_i \in \mathfrak{A}$.

Des weiteren ist $\bigcup_{i=1}^{\infty} E_i = E$, da $(f_n(s))$ stets gegen $f(s)$ konvergiert. Es folgt ($n \in \mathbb{N}$ beliebig)
$$\begin{aligned} \varphi \geq \int_E f_n \, d\mu &\geq \int_{E_n} f_n \, d\mu && (2.4.b)) \\ &\geq \int_{E_n} c \cdot g \, d\mu && (2.4.a)) \\ &= c \cdot \int_{E_n} g \, d\mu && (2.3.d)) \end{aligned}$$

Da $c < 1$ beliebig war, ist sogar

$$\gamma \geq \int_{E_n} g \, d\mu .$$

Die σ-Additivität von μ (die ja die von $A \mapsto \int_A g \, d\mu$ impliziert, 2.3 a)) zeigt nun noch I.2.4

$$\gamma \geq \int_E g \, d\mu .$$

Satz 2.5 (bzw. das folgende Korollar 2.6) heißen Satz von der monotonen Konvergenz oder Satz von Beppo Levi.

2.6 Korollar $g_1, g_2, \dots : S \to [0,\infty]$ seien meßbar, und $g = \sum_{n=1}^{\infty} g_n$. Dann ist g meßbar, und für $E \in \mathfrak{A}$ gilt

$$\int_E g \, d\mu = \sum_{n=1}^{\infty} \int_E g_n \, d\mu .$$

Beweis: Betrachte $f_n = g_1 + \dots + g_n$!

Eine andere Methode, das Integral einer positiven meßbaren Funktion f zu definieren, besteht darin, f als Limes einer monotonen Folge (f_n) von Treppenfunktion darzustellen (1.8) und dann

$$\int_E f \, d\mu = \sup_n \int_E f_n \, d\mu$$

zu definieren. Hier muß man jedoch die Wohldefiniertheit nachweisen (was bei der obigen Def. 2.2 nicht nötig war) und dann zuerst das Analogon von 2.5 für Treppenfunktionen beweisen. Ein solcher Weg erscheint etwas umständlicher.

2.5 liefert auf eleganten Wege das

2.7 Lemma $f : S \to [0,\infty]$ meßbar, $E \in \mathfrak{A} \Rightarrow$

$$\int_E f \, d\mu = \int_S 1_E f \, d\mu$$

Beweis: Die Aussage ist offenbar (?) richtig für Treppenfunktionen. f kann nun monoton durch Treppenfunktionen f_n approximiert werden (1.8); dann gilt auch $1_E f_n \nearrow 1_E f$ [*)], also nach 2.5

$$\int_E f \, d\mu = \lim \int_E f_n \, d\mu = \lim \int_S 1_E f_n \, d\mu = \int_S 1_E f \, d\mu .$$

Freilich hätte man 2.7 auch direkt aus Def. 2.2 herleiten können; jedoch stellt die obige Methode ein typisches Beweisverfahren der Integrationstheorie dar: Eine zu zeigende Behauptung über meßbare Funktionen wird zuerst für Indikatorfunktionen, dann für Treppenfunktionen bewiesen [im Regelfall sind die Aussagen dafür trivial]; dann zeigt man durch Grenzübergang den allgemeinen Fall [dafür sind dann Sätze wie 2.5 oder die im 3. Abschnitt besprochenen Konvergenzsätze probate Mittel].

Eine weitere Anwendung dieser Methode folgt beim Beweis des sonst nur mühsam zu zeigenden

2.8 Lemma $f, g : S \to [0,\infty]$ seien meßbar. Dann gilt

$$\int_S (f+g) \, d\mu = \int_S f \, d\mu + \int_S g \, d\mu$$

*) $\nearrow$ bedeutet monotone Konvergenz

Beweis: Wähle monotone Folgen von Treppenfunktionen (f_n), (g_n), die f bzw. g punktweise approximieren (1.8). Mit 2.5 schließt man

$$\int_S (f+g)\,d\mu = \lim_n \int_S (f_n+g_n)\,d\mu \qquad (f_n+g_n \nearrow f+g)$$

$$= \lim_n \left(\int_S f_n\,d\mu + \int_S g_n\,d\mu \right) \qquad (2.3(i))$$

$$= \lim_n \int_S f_n\,d\mu + \lim_n \int_S g_n\,d\mu$$

$$= \int_S f\,d\mu + \int_S g\,d\mu . \qquad (2.5)$$

Kommen wir nun zum dritten Schritt bei der Definition des Integrals.

<u>2.9 Definition</u> $(S, \mathfrak{A}, \mu)$ sei ein Maßraum, und $f: S \to [-\infty, \infty]$ sei meßbar. f heißt integrierbar (genauer: μ-integrierbar), wenn die positiven (und meßbaren! [1.3]) Funktionen f^+ und f^- ein endliches Integral besitzen. Man setzt

$$\int_S f\,d\mu = \int_S f^+\,d\mu - \int_S f^-\,d\mu .$$

Einige <u>Bemerkungen</u> zu dieser Definition:

- Für eine positive meßbare Funktion f ist $\int_S f\,d\mu$ stets definiert. f ist integrierbar genau dann, wenn $\int_S f\,d\mu < \infty$ ist. Die Symbole $\int_S f\,d\mu$ aus Def. 2.2 und Def. 2.9 definieren dann dieselbe Zahl.
- Wegen 2.4 a) und 2.8 ist $f: S \to [-\infty, \infty]$ genau dann integrierbar, wenn f meßbar ist und $\int_S |f|\,d\mu < \infty$ ist (denn $0 \le f^+, f^- \le |f| = f^+ + f^-$). Achtung: die Meßbarkeit von $|f|$ impliziert nicht die Meßbarkeit von f [aber umgekehrt: 1.3].

- Analog wird für $E \in \mathfrak{A}$ „Integrierbarkeit über E" definiert. Wieder ist $\int_E f\,d\mu = \int_S 1_E f\,d\mu$.
- Weitere geläufige Symbole für das Integral sind $\int_E f(s)\,d\mu(s)$ und $\int_E f(s)\,\mu(ds)$.
- Ist $f: S \to [-\infty, \infty]$ (meßbar) als Differenz integrierbarer g und h dargestellt, so daß $\infty - \infty$ nicht vorkommt, so gilt
$$\int_S f = \int_S g - \int_S h .$$

Das folgt mit der Methode des nächsten Satzes, wo wir uns jedoch auf reellwertige Funktionen beschränken werden.

<u>2.10 Satz</u> $(S, \mathfrak{A}, \mu)$ sei ein Maßraum, $f, g: S \to \mathbb{R}$ seien integrierbar, $\alpha, \beta \in \mathbb{R}$. Dann ist $\alpha f + \beta g$ integrierbar mit

$$\int_S (\alpha f + \beta g)\,d\mu = \alpha \int_S f\,d\mu + \beta \int_S g\,d\mu .$$

Ferner gilt $\quad f \ge 0 \Rightarrow \int_S f\,d\mu \ge 0$

sowie $\quad \left| \int_S f\,d\mu \right| \le \int_S |f|\,d\mu .$

Anders gesagt, bildet die Menge der integrierbaren Funktionen einen Vektorraum, worauf $f \mapsto \int_S f\,d\mu$ ein positives lineares Funktional definiert. (Das sollte man vernünftigerweise von einer Integration erwarten.)

Beweis: $\alpha f + \beta g$ ist jedenfalls meßbar (1.3) und integrierbar nach der 2. Bemerkung, denn

$$\int_S |\alpha f + \beta g|\,d\mu \le \int_S (|\alpha|\,|f| + |\beta|\,|g|)\,d\mu \qquad (2.4\text{ a})$$

$$= \int_S |\alpha|\,|f|\,d\mu + \int_S |\beta|\,|g|\,d\mu \qquad (2.8)$$

$$= |\alpha| \int_S |f|\, d\mu + |\beta| \int_S |g|\, d\mu$$ (Beweis analog i))

$$< \infty.$$

Zeigen wir nun die Linearität der Integration. Es reicht, die beiden Fälle

i) $\alpha = \beta = 1$

ii) $\alpha \in \mathbb{R}$ beliebig, $\beta = 0$

zu behandeln.

i) Da $f + g = (f+g)^+ - (f+g)^-$

und $= (f^+ - f^-) + (g^+ - g^-)$

ist, ist $f^+ + g^+ + (f+g)^- = f^- + g^- + (f+g)^+$.

2.8 ergibt

$$\int_S f^+ d\mu + \int_S g^+ d\mu + \int_S (f+g)^- d\mu = \int_S f^- d\mu + \int_S g^- d\mu + \int_S (f+g)^+ d\mu,$$

woraus durch Subtraktion (alle Integrale sind endlich!)

$$\int_S f\, d\mu + \int_S g\, d\mu = \int_S (f+g)\, d\mu$$

folgt.

ii) • Im Falle $\alpha \geq 0$ und $f \geq 0$ folgt die Behauptung durch Grenzübergang aus 2.3 d).

• $\alpha \geq 0$, f beliebig: Es ist $\alpha f^+ = \alpha f + \alpha f^-$, daher nach (i) und dem 1. Fall:

$$\alpha \int_S f^+ d\mu = \int_S \alpha f^+ d\mu = \int_S \alpha f\, d\mu + \alpha \int_S f^- d\mu$$

$$\Rightarrow \quad \alpha \int_S f\, d\mu = \alpha \int_S f^+ d\mu - \alpha \int_S f^- d\mu = \int_S \alpha f\, d\mu$$

• $\alpha = -1$: $0 = \int_S (f + (-f))\, d\mu \underset{(i)}{=} \int_S f\, d\mu + \int_S (-f)\, d\mu$.

• $\alpha < 0$: Kombiniere Fall 2 und 3 zu

$$\int_S \alpha f\, d\mu = \int_S (-\alpha)\cdot(-f)\, d\mu = -\alpha \int_S (-f)\, d\mu = \alpha \int_S f\, d\mu$$

Der 1. Zusatz ist schon in 2.4 a) bewiesen, der 2. folgt daraus ($f, -f \leq |f|$!).

Es sei noch darauf hingewiesen, daß für integrierbare f die vertraute Formel

$$E_1 \cap E_2 = \emptyset \quad \Rightarrow \quad \int_{E_1 \cup E_2} f\, d\mu = \int_{E_1} f\, d\mu + \int_{E_2} f\, d\mu$$

gilt, da ja $\mathbb{1}_{E_1 \cup E_2} f = \mathbb{1}_{E_1} f + \mathbb{1}_{E_2} f$.

<u>2.11 Beispiele</u> a) Betrachte den klassischen Fall des Lebesguemaßes λ. Auf den Zusammenhang des Riemannschen und des Lebesgueschen Integrals wird in Kap. IV im Detail eingegangen; hier begnügen wir uns mit der Bemerkung, daß für stetige f $R-\int_a^b f(s)\, ds$ und das Lebesguesche $\int_{[a,b]} f\, d\lambda$ übereinstimmen. (Daher wird im Fall des Lebesguemaßes auch weiterhin bisweilen $\int_a^b f(s)\, ds$ geschrieben.)

In der Tat: O.E. darf $f \geq 0$ angenommen werden. Die gleichmäßige Stetigkeit von f garantiert, daß es Treppenfunktionen f_n gibt, die f von unten monoton approximieren, so daß die Stufen von f_n sogar Intervalle sind und die Konvergenz gleichmäßig ist. Nach Definition stimmen R- und L-Integral von f_n überein, und wegen der gleichmäßigen Konvergenz gilt nach einem Satz der Analysis

$$R-\int_a^b f_n(s)\, ds \longrightarrow R-\int_a^b f(s)\, ds.$$

Nach 2.5 gilt entsprechendes für die L-Integrale, also

$$R-\int_a^b f(s)\,ds \;=\; L-\int_a^b f(s)\,ds.$$

b) (Maße mit Dichten)

Es sei $(S, \mathfrak{A}, \mu)$ ein Maßraum, und $g \geq 0$ sei meßbar. Dann definiert $\nu: E \mapsto \int_E g\,d\mu$ ein Maß auf $\mathfrak{A}$ (eine Präform dieser Aussage ist in 2.3 a) formuliert).

[Das ist eine Konsequenz des Satzes von Beppo Levi (2.6): Seien $E_1, E_2, \ldots \in \mathfrak{A}$ paarweise disjunkt. Dann ist

$$1_{\cup E_i} = \sum_{i=1}^{\infty} 1_{E_i} \quad \text{(punktweise)},$$

also $\nu(\cup E_i) = \int_S (\Sigma 1_{E_i})\, g\,d\mu = \sum_i \int_{E_i} g\,d\mu = \Sigma_i$

Eine meßbare Funktion $f: S \to \mathbb{R}$ ist genau dann ν-integrierbar, wenn $\int_S |f|\cdot g\,d\mu < \infty$ ist, und es ist

$$(*) \qquad \int_S f\,d\nu = \int_S f\cdot g\,d\mu.$$

[$(*)$ stimmt

- wenn f eine Indikatorfunktion ist (Definition von ν),
- wenn f eine Treppenfunktion ist (Integration ist linear),
- wenn $f \geq 0$ und meßbar ist (wegen 1.8 und 2.5).

Also ist die erste Behauptung bewiesen, und nun folgt die zweite durch Differenzbildung.]

c) (Dirac-Maß)

Betrachte das Dirac-Maß δ_s (Bsp. I.2.2c)). Mit derselben Technik wie unter b), (Indikatorfunktionen ⇝ Treppenfunktionen ⇝ positive meßbare

Funktionen ⇝ integrierbare Funktionen) zeigt man

$$\int_S f\,d\delta_s = f(s).$$

d) (Summen als Integrale)

Betrachte den Maßraum $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \mu)$, wo μ das zählende Maß (Bsp. I.2.2 d)) ist. Eine Funktion $f: \mathbb{N} \to \mathbb{R}$ ist nichts anderes als eine Folge reeller Zahlen, und sie ist automatisch meßbar. Für $f = 1_{\{k\}}$ gilt offensichtlich $\int_{\mathbb{N}} f\,d\mu = 1$, daher zeigen die unter b) angegebenen Schritte:

$f: n \mapsto a_n$ ist genau dann μ-integrierbar, wenn $\sum_{n=1}^{\infty} |a_n| < \infty$ ist, und es ist dann $\int_{\mathbb{N}} f\,d\mu = \sum_{n=1}^{\infty} a_n$.

Zum Schluß sollen kurz komplexe Integranden besprochen werden.

$f: S \to \mathbb{C}$ heiße meßbar, wenn die reellen Funktionen $\operatorname{Re} f$ und $\operatorname{Im} f$ meßbar sind. (Identifiziert man $\mathbb{C}$ mit $\mathbb{R}^2$, so entspricht das dem üblichen Begriff der Borel-Meßbarkeit für $\mathbb{R}^2$-wertige Funktionen, siehe 1.2 c).) Satz 1.3 gilt entsprechend (max und min sind jetzt freilich sinnlos). Akzeptiert man komplexwertige Treppenfunktionen, so bleiben auch 1.8 b) und c) richtig.

Eine komplexe meßbare Funktion heißt μ-integrierbar, wenn $\operatorname{Re} f$ und $\operatorname{Im} f$ es sind. Wie nicht anders zu erwarten, setzt man

$$\int_S f\,d\mu = \int_S \operatorname{Re} f\,d\mu + i \int_S \operatorname{Im} f\,d\mu.$$

Offenbar gilt wieder für meßbares f: f integrierbar $\Leftrightarrow$ $|f|$ integrierbar, und 2.10 überträgt sich ohne Schwierigkeiten, ebenso wie die Beispiele 2.11.

II. 3 Konvergenzsätze

Das folgende sog. Lemma von Fatou kann als Verallgemeinerung des Satzes von Beppo Levi aufgefaßt werden.

3.1 Satz $(S, \mathfrak{A}, \mu)$ sei ein Maßraum, $f_n : S \to [0, \infty]$ seien meßbar. Dann gilt
$$\int_S \liminf f_n \, d\mu \le \liminf \int_S f_n \, d\mu$$

Beweis: $\liminf f_n$ ist in der Tat meßbar (1.4). Der Satz von Beppo Levi (2.5) impliziert:
$$\int_S \liminf f_n \, d\mu = \int_S \sup_k \inf_{n \ge k} f_n \, d\mu$$
$$= \sup_k \int_S \inf_{n \ge k} f_n \, d\mu$$
$$\le \sup_k \inf_{n \ge k} \int_S f_n \, d\mu$$
(denn: $\inf_{n \ge k} f_n \le f_m \quad \forall m \ge k$
$$\Rightarrow \int_S \inf_{n \ge k} f_n \, d\mu \le \int_S f_m \, d\mu \quad \forall m \ge k)$$
$$= \liminf \int_S f_n \, d\mu .$$

Damit erhält man ohne große Mühe den ersten zentralen Satz der Integrationstheorie, den Konvergenzsatz von Lebesgue (manchmal auch Satz von der majorisierten [oder auch dominierten] Konvergenz genannt.)

3.2 Theorem (Lebesgue)

$(S, \mathfrak{A}, \mu)$ sei ein Maßraum, die $f_n : S \to \mathbb{R}$ seien meßbar, und für alle $s \in S$ existiere $f(s) = \lim_{n\to\infty} f_n(s)$. Falls eine integrierbare Funktion $g : S \to [0, \infty]$ mit $|f_n| \le g$ existiert, so gelten:

a) die f_n und f sind integrierbar,

b) $\int_S |f_n - f| \, d\mu \to 0$,

c) $\lim_{n\to\infty} \int_S f_n \, d\mu = \int_S \lim_{n\to\infty} f_n \, d\mu \quad (= \int_S f \, d\mu).$

Beweis: a) $|f_n| \le g \Rightarrow \int_S |f_n| \, d\mu \le \int_S g \, d\mu < \infty$

Da auch $|f| \le g$ gilt, ist f ebenfalls integrierbar (die Meßbarkeit ergibt sich aus 1.4).

b) Es ist $0 \le |f_n - f| \le g + |f| =: \tilde{g}$ (alle $n \in \mathbb{N}$), wo $\tilde{g}$ integrierbar ist. Das Lemma von Fatou impliziert
$$\int_S \tilde{g} \, d\mu = \int_S \lim (\tilde{g} - |f_n - f|) \, d\mu$$
$$= \int_S \liminf \underbrace{(\tilde{g} - |f_n - f|)}_{\ge 0 !} d\mu$$
$$\le \liminf \int_S (\tilde{g} - |f_n - f|) \, d\mu$$
$$= \int_S \tilde{g} \, d\mu - \limsup \int_S |f_n - f| \, d\mu .$$
Durch Subtraktion der reellen Zahl (!) $\int_S \tilde{g} \, d\mu$ folgt
$$(0 \le) \limsup \int_S |f_n - f| \, d\mu \le 0 ,$$
was b) zeigt.

c) folgt jetzt aus $\left| \int_S f_n \, d\mu - \int_S f \, d\mu \right| \underset{2.10}{\le} \int_S |f_n - f| \, d\mu \to 0.$

Einfache Beispiele zeigen, daß man auf die Existenz so einer integrierbaren g in 3.2 nicht verzichten darf.

Ist $\mu(S) < \infty$, so sind die Voraussetzungen von 3.2 insbesondere erfüllt, wenn $M \geq 0$ mit

$$|f_n(s)| \leq M \qquad \forall s \in S,\ n \in \mathbb{N}$$

existiert (wähle dann $g = M \cdot 1_S$).

Nach Beispiel 2.11 a) ergibt diese letzte Version Integralkonvergenzsätze für das Riemann-Integral stetiger Funktionen. Explizit:

Sind f_n und $f : [a,b] \to \mathbb{R}$ stetig und existiert $M \geq 0$ mit

$$|f_n(s)| \leq M \qquad \forall s \in [a,b],\ n \in \mathbb{N},$$

so folgt aus $f_n(s) \to f(s) \quad \forall s \in [a,b]$

$$\int_a^b f_n(s)\,ds \longrightarrow \int_a^b f(s)\,ds.$$

(Ein analoger Satz gilt auch, wenn f_n und f bloß als R-integrierbar vorausgesetzt sind.) Beachte, daß die Integrierbarkeit von f hier Teil der Voraussetzung ist!

Der erwähnte Satz von Arzelà / Osgood*) ist im Rahmen der Riemannschen Theorie nur schwer zu beweisen; eine interessante Übersicht gibt

W. A. J. Luxemburg: Arzelà's dominated convergence theorem for the Riemann integral. Amer. Math. Monthly 78 (1971), 970-979.

Der Satz wird direkt aus der σ-Additivität des Jordanschen Inhalts auf $\mathbb{R}^q$ (unser Satz I.2.5) von

J. W. Lewin: A truly elementary approach to the bounded convergence theorem. Amer. Math. Monthly 93 (1986), 395-397

*) Arzelà (1885) für R-integrierbare Funktionen; Osgood (1898) hat den Satz dann wiederentdeckt, aber nur für stetige Funktionen.

hergeleitet. (Der Punkt in allen Beweisen ist, eine Spur σ-Additivität nachzuweisen; aber σ-Additivität ist dem Riemannschen Aufbau vollkommen fremd. Daher tun sich all diese Beweise recht schwer.)

Die weiteren Untersuchungen über die Konvergenz von Folgen meßbarer Funktionen haben mit dem Zauberwort „fast überall" zu tun. Zuerst also etwas dazu.

Sei $(S, \mathfrak{A}, \mu)$ ein Maßraum, und es sei (P) eine Eigenschaft, die ein Element $s \in S$ haben kann. Man sagt dann, (P) besteht fast überall (präziser μ-fast überall), wenn $N \in \mathfrak{A}$ mit $\mu(N) = 0$ *) und

$$s \notin N \Rightarrow s \text{ hat (P)}.$$

Die Feinheit, die hier zu beachten ist und leider zu ein paar technischen Verwicklungen führt, ist, daß nicht $\{s : s \text{ hat nicht (P)}\} \in \mathfrak{A}$ behauptet ist, sondern nur, daß $\{s : s \text{ hat nicht (P)}\}$ Teilmenge einer meßbaren Menge vom Maß 0 ist.

Einige Beispiele zum „fast überall"-Konzept.

a) Offenbar ist λ-fast jede Zahl rational. In Beispiel 1.5 wurde erwähnt, daß λ-fast jede Zahl in $]0,1]$ normal ist; leider hilft diese Information nicht zu entscheiden, ob eine vorgelegte Zahl (etwa $\frac{\pi}{4}$) normal ist oder nicht. Man kann nur sagen, daß das mit Wahrscheinlichkeit 1 („fast sicher") zutrifft.

b) Offensichtlich (aber trotzdem erwähnenswert) ist auch, daß der Bezug zum vorgelegten Maß entscheidend ist: Für das Dirac-Maß δ_0 auf $(\mathbb{R}, \mathfrak{P}(\mathbb{R}))$

*) Eine solche Menge heißt μ-Nullmenge.

ist fast jede Zahl $= 0$.

c) Sei M eine nicht-Borelsche Teilmenge der Cantormenge (vgl. S. 4…

Betrachtet man den Maßraum $(\mathbb{R}, \mathrm{Bor}(\mathbb{R}), \lambda)$, so liegt λ-fast jede…

außerhalb M. Hier ist ein Beispiel, wo $\{s : s \text{ hat nicht } (P)\}$ (wo…

$\notin M$) nicht meßbar ist, sondern nur Teilmenge einer meßbaren Nullmenge

ist. Für $(\mathbb{R}, \mathrm{Leb}(\mathbb{R}), \lambda)$ taucht dieses Nichtmeßbarkeitsdilemma nicht

auf, da λ auf $\mathrm{Leb}(\mathbb{R})$ vollständig ist (S. 40).

d) Auf $[0,1]$ konvergiert (t^n) λ-fast überall gegen 0.

e) $1_{\mathbb{Q}}$ stimmt λ-fast überall mit der konstanten Funktion 0 überein.

Dafür schreibt man kürzer

$$1_{\mathbb{Q}} = 0 \qquad \lambda\text{-f.ü.}$$

f) Entsprechend sind für Funktionen auf einem Maßraum $(S, \mathfrak{A}, \mu)$

$$f \geq g \qquad \mu\text{-f.ü.},$$

$$\lim_{n\to\infty} f_n = f \qquad \mu\text{-f.ü.} \qquad \text{etc.}$$

zu verstehen. Leider folgt i.a. aus „f meßbar und $g = f$ f.ü.“

nicht „g meßbar“ (im Beispiel c) ist $1_M = 0$ f.ü.), und a…

(*) $\quad \lim_{n\to\infty} f_n = g$ f.ü., alle f_n meßbar $\not\Rightarrow$ g meßbar.

Ein Ausweg aus diesem Dilemma ist, f mit Gewalt meßbar zu

machen, indem man zum vervollständigten Maßraum $(S, \mathfrak{A}_\mu, \hat{\mu})$ üb…

wo $\quad \mathfrak{A}_\mu = \{ B \triangle M : B \in \mathfrak{A}, \exists N \in \mathfrak{A};\ \mu(N) = 0 \ \&\ M \subset N\}$

$\quad \hat{\mu}(B \triangle M) = \mu(B)$.

Eine andere Möglichkeit besteht darin, nicht den Maßraum zu ändern,

sondern (in (*) etwa) g zu ändern: Wähle eine meßbare Nullm…

mit $\qquad s \notin N \Rightarrow \lim_{n\to\infty} f_n(s) = g(s)$.

Die Folge $(1_{N^c} f_n)$ konvergiert dann überall, und zwar gegen $1_{N^c} \cdot g$,

so daß $1_{N^c} g =: f$ meßbar ist und immer noch

$$\lim f_n = f \qquad \text{f.ü.}$$

gilt. Fazit: In (*) ist g ohne Einschränkung meßbar, evtl. ist g unwesentlich (nämlich auf der Teilmenge einer Nullmenge) abzuändern. Diesen Weg werden wir einschlagen.

Auf f.ü. bestehende Eigenschaften kommt man typischerweise durch eine Integration, andersherum ist es für eine Integration gleichgültig, wenn der Integrand auf einer Nullmenge geändert wird.

<u>3.3 Lemma</u> $\quad f : S \to [-\infty, \infty]$ sei μ-integrierbar.

a) $\quad |f| < \infty \qquad \mu\text{-f.ü.}$

b) $\quad \int_S |f| \, d\mu = 0 \ \Rightarrow \ f = 0 \ \ \mu\text{-f.ü.}$

c) $\quad g$ meßbar, $f = g$ μ-f.ü. $\Rightarrow$ g μ-integrierbar & $\int_S g \, d\mu = \int_S f \, d\mu$

Beweis: a) Für alle n ist $n 1_{\{|f| = \infty\}} \leq |f|$ (beachte $\{|f| = \infty\} \in \mathfrak{A}$),

daher $\forall n$: $\quad n \cdot \mu(\{|f| = \infty\}) = \int_S n \cdot 1_{\{|f| = \infty\}} \, d\mu \leq \int_S |f| \, d\mu < \infty$.

Es folgt $\mu(\{|f| = \infty\}) = 0$.

b) Sei $E_n = \{|f| \geq \frac{1}{n}\} \in \mathfrak{A}$. Dann

$$0 = \int_S |f| \, d\mu \geq \int_{E_n} |f| \, d\mu \geq \frac{1}{n} \mu(E_n)$$

$\Rightarrow$ alle $\mu(E_n) = 0 \quad \Rightarrow \quad \mu(\{f \neq 0\}) = \mu(\bigcup_{n=1}^{\infty} E_n) = 0$.

c) Ist N eine Nullmenge und $h \geq 0$ meßbar, so ist nach 2.4. d):
$$\int_N h \, d\mu = 0 .$$
Es folgt hier
$$\int_S |g| \, d\mu = \int_{\{f=g\}} (\ldots) + \int_{\{f \neq g\}} (\ldots)$$
$$= \int_{\{f=g\}} |f| \, d\mu + 0$$
$$= \int_{\{f=g\}} |f| \, d\mu + \int_{\{f \neq g\}} |f| \, d\mu$$
$$= \int_S |f| \, d\mu < \infty$$
Genauso zeigt man $\int_S f^+ \, d\mu = \int_S g^+ \, d\mu$, $\int_S f^- \, d\mu = \int_S g^- \, d\mu$,
folglich $\int_S f \, d\mu = \int_S g \, d\mu$.

Im folgenden werden wir auch f.ü. definierte Funktionen zulassen (z. B. $s \mapsto \frac{1}{s}$ auf $[0,1]$; für integrierbare f und g ist nach 3.3 a) $f-g$ f.ü. definiert). Es dürfte klar sein, wie solche Funktionen zu integrieren sind.

Ein letztes Beispiel soll zeigen, wie mit dem f.ü.-Kalkül operiert wird:

<u>3.4 Korollar</u> Sei $(S, \mathfrak{A}, \mu)$ ein Maßraum, $f_n : S \to [-\infty, \infty]$ seien meßbar, und $\lim\limits_{n\to\infty} f_n(s)$ existiere μ-fast überall. Falls eine integrierbare Funktion $g : S \to [0,\infty]$ mit
$$|f_n| \leq g \quad \text{f.ü.}$$
existiert, so gelten:

a) Die f_n sind integrierbar.

b) Es existiert integrierbares $f : S \to \mathbb{R}$ mit
$$\lim_{n\to\infty} f_n(s) = f(s) \quad \text{f.ü.}$$
$$\int_S |f_n - f| \, d\mu \to 0$$
$$\int_S f_n \, d\mu \to \int_S f \, d\mu$$

Beweis: Es sei $N_n = \{ f_n > |g| \}$, N_0 eine Nullmenge mit
$$s \notin N_0 \Rightarrow \lim f_n(s) \text{ existiert (in } [-\infty, \infty]).$$
$N = \bigcup\limits_{n=0}^{\infty} N_n$ ist dann eine Nullmenge. Die $\tilde{f}_n = 1_{N^c} f_n$ erfüllen die Voraussetzungen von 3.2, $\tilde{f} = \lim \tilde{f}_n$ existiert jetzt überall. $\tilde{f}$ ist daher integrierbar, und 3.3. a) gestattet, $\mathbb{R}$-wertiges integrierbares f mit $f = \tilde{f}$ f.ü. zu finden.

Die Behauptung folgt nun aus 3.2 und 3.3.

In Zukunft sollen f.ü.-Modifikationen etwas knapper dargestellt werden. Wesentlich interessanter ist:

<u>3.5 Satz</u> (Satz von Egoroff)

$(S, \mathfrak{A}, \mu)$ sei ein <u>endlicher</u> Maßraum, und die Folge (f_n) meßbarer Funktionen konvergiere f.ü. gegen 0. Zu $\varepsilon > 0$ existiert dann $E \in \mathfrak{A}$ mit $\mu(E^c) \leq \varepsilon$ derart, daß (f_n) auf E gleichmäßig gegen 0 konvergiert.

Beweis: Sei $k \in \mathbb{N}$ zunächst fest. Die Mengen $E_{k,m} = \bigcap_{n \ge m} \{|f_n| \le \frac{1}{k}\}$ sind offenbar der Schlüssel zum Beweis des Satzes.

Nach Voraussetzung existiert eine Nullmenge N mit $\bigcup_m E_{k,m} = S \setminus N$.

Da $E_{k,1} \subset E_{k,2} \subset \cdots$ gilt, folgt (I.2.4) $\mu(E_{k,m}) \xrightarrow[m\to\infty]{} \mu($

Wähle also $m (= m_k)$ mit $\mu(E^c_{k,m_k}) \le \varepsilon \cdot 2^{-k}$. ($\mu$ ist endlich!)

$E = \bigcap_k E_{k,m_k}$ ist nun die gesuchte Menge:

$$\mu(E^c) \le \sum_k \mu(E^c_{k,m_k}) \le \varepsilon$$

und $n \ge m_k$, $s \in E \Rightarrow |f_n(s)| \le \frac{1}{k}$ (da insbesondere $E \subset E_{k,m_k}$)

d.h. (f_n) konvergiert gleichmäßig auf E.

Der Satz gilt nicht für unendliche Maßräume (μ = zählendes Maß auf $\mathbb{N}$, $f_n = 1_{\{n\}}$) und nicht für $\varepsilon = 0$ ($f_n(t) = t^n$ auf $[0,1]$). (Bew

Als nächstes werden zwei neue Konvergenzbegriffe eingeführt.

3.6 Definition $(S, \mathfrak{A}, \mu)$ sei ein Maßraum, $f_n, f : S \to [-\infty, \infty]$ meßbare Funktionen.

a) (f_n) konvergiert gegen f „dem Maße nach" (oder „μ-stochastisch"), fa

$$\forall \varepsilon > 0 \quad \mu(\{|f_n - f| \ge \varepsilon\}) \xrightarrow[n\to\infty]{} 0.$$

Abkürzend schreibt man $f_n \xrightarrow{\mu} f$ oder $f_n \xrightarrow{L^0} f$.

b) f_n und f seien zusätzlich integrierbar. (f_n) konvergiert gegen f „im Mittel", falls

$$\int_S |f_n - f| \, d\mu \to 0.$$

Schreibweise: $f_n \xrightarrow{L^1} f$.

Die zunächst unanschauliche Symbolik $\xrightarrow{L^0}$ und $\xrightarrow{L^1}$ wird im nächsten Abschnitt (hoffentlich) verständlich.

Achtung: Manche Autoren (z.B. Bauer) definieren μ-stochastische Konvergenz anders, nämlich

$$\forall \varepsilon > 0 \quad \forall E \in \mathfrak{A}, \mu(E) < \infty: \quad \mu(\{|f_n - f| \ge \varepsilon\} \cap E) \to 0.$$

Auf endlichen Maßräumen stimmen beide Definitionen offenkundig überein.

Einfaches Beispiel: Auf $([0,1], \mathrm{Bor}([0,1]), \lambda)$ konvergiert $t \mapsto t^n$ dem Maße nach gegen 0, aber auch gegen $1_{\{1\}}$ (!).

Der Limes ist daher nicht eindeutig, aber „fast":

3.7 Lemma Alle im folgenden auftretenden Funktionen seien meßbar.

a) $f_n \xrightarrow{\mu} f$, $g_n \xrightarrow{\mu} g \Rightarrow f_n + g_n \xrightarrow{\mu} f + g$

b) $f_n \xrightarrow{\mu} f$, $\alpha \in \mathbb{R} \Rightarrow \alpha f_n \xrightarrow{\mu} \alpha f$

c) $f_n \xrightarrow{\mu} f$, $g_n \xrightarrow{\mu} g$; $\forall n: f_n = g_n$ f.ü. $\Rightarrow f = g$ f.ü.

Analoge Aussagen gelten für die Konvergenz im Mittel.

Beweis: a)

$$\begin{aligned}
&\mu(\{|(f_n + g_n) - (f + g)| \ge \varepsilon\}) \\
&\le \mu(\{|f_n - f| + |g_n - g| \ge \varepsilon\}) \\
&\le \mu(\{|f_n - f| \ge \tfrac{\varepsilon}{2}\}) + \mu(\{|g_n - g| \ge \tfrac{\varepsilon}{2}\}) \\
&\xrightarrow[n\to\infty]{} 0.
\end{aligned}$$

b) ist trivial.

c) Für alle $\varepsilon > 0$ gilt nach der Dreiecksungleichung

$$\mu(\{|f-g| \geq \varepsilon\})$$
$$\leq \mu(\{|f-f_n| + |f_n-g_n| + |g_n-g| \geq \varepsilon\})$$
$$\leq \mu(\{|f-f_n| \geq \tfrac{\varepsilon}{3}\}) + \mu(\{|f_n-g_n| \geq \tfrac{\varepsilon}{3}\}) + \mu(\{|g_n-g| \geq \tfrac{\varepsilon}{3}\})$$
$$\to 0 \quad \text{(der mittlere Term verschwindet).}$$

Folglich $\quad \mu(\{|f-g| \geq \frac{1}{k}\}) = 0. \quad \forall\, k \in \mathbb{N}$

und
$$\mu(\{f \neq g\}) = \mu\Big(\bigcup_k \{|f-g| \geq \tfrac{1}{k}\}\Big) \leq \sum_k \mu(\{|f-g| \geq \tfrac{1}{k}\}) = 0.$$

Der L^1-Fall ist sehr einfach und bleibt zur Übung.

Den Vergleich der verschiedenen Konvergenzarten bereiten wir durch eine klassische Ungleichung vor:

3.8 Lemma (Tschebyscheffsche Ungleichung)

Auf dem Maßraum $(S, \mathfrak{A}, \mu)$ sei g positiv und meßbar. Sei $r>0$.
Dann gilt für alle $\varepsilon > 0$
$$\mu(\{g \geq \varepsilon\}) \leq \frac{1}{\varepsilon^r} \int_S g^r \, d\mu$$

Beweis: Bemerke, daß g^r meßbar ist. Mit $A = \{g \geq \varepsilon\} \in \mathfrak{A}$ gilt
$$\int_S g^r \, d\mu \geq \int_A g^r \, d\mu \geq \int_A \varepsilon^r \, d\mu = \varepsilon^r \cdot \mu(A).$$

3.9 Satz Auf dem Maßraum $(S, \mathfrak{A}, \mu)$ seien f_n, f meßbar.

a) μ sei endlich. Dann:
$$f_n \to f \text{ f.ü.} \quad \Rightarrow \quad f_n \xrightarrow{\mu} f$$

b) μ sei beliebig. Dann:
$$f_n \xrightarrow{\mu} f \quad \Rightarrow \quad \exists \text{ ex. Teilfolge}: \ f_{n_k} \to f \text{ f.ü.}$$

c) μ sei beliebig, f_n, f zusätzlich integrierbar. Dann:
$$\int_S |f_n - f| \, d\mu \to 0 \quad \Rightarrow \quad f_n \xrightarrow{\mu} f$$

Beweis: a) ist eine einfache Konsequenz des Satzes von Egoroff (3.5). Seien $\varepsilon, \varepsilon^* > 0$ gegeben. Wähle $E \in \mathfrak{A}$ mit $\mu(E^c) \leq \varepsilon^*$, so daß (f_n) auf E gleichmäßig gegen f konvergiert. Für hinreichend großes n ist daher $\{|f_n - f| \geq \varepsilon\} \cap E = \emptyset$; für diese n ist also
$$\mu(\{|f_n-f| \geq \varepsilon\}) = \mu(\{|f_n-f| \geq \varepsilon\} \cap E^c) \leq \mu(E^c) \leq \varepsilon^*.$$

b) O.E. ist $f = 0$ und sind alle $f_n \geq 0$ (sonst gehe zu $|f_n - f|$ über).

Wähle induktiv $n_1 < n_2 < \cdots$ mit
$$\mu(\{f_{n_k} \geq 2^{-k}\}) \leq 2^{-k}.$$
Wegen $f_n \xrightarrow{\mu} 0$ ist das möglich. Abkürzend schreibe $A_k = \{f_{n_k} \geq 2^{-k}\}$.
Dann gelten mit $A = \bigcap_m \bigcup_{k \geq m} A_k$ ($=: \limsup A_k$):

a) $s \notin A \Rightarrow f_{n_k}(s) \to 0$

b) $\mu(A) = 0$.

[a) $s \notin A \Rightarrow s \in \bigcup_m \bigcap_{k \geq m} A_k^c$

$\Rightarrow \exists m \ \forall k \geq m \quad 0 \leq f_{n_k}(s) \leq 2^{-k} \to 0$.

b) Für alle $m \geq 0$ gilt

$$\mu(A) \leq \mu\Big(\bigcup_{k \geq m} A_k\Big) \leq \sum_{k \geq m} \mu(A_k) \xrightarrow[m \to \infty]{} 0 \ , \ \text{da } \Sigma_{\mu(A_k)}$$

[Die unter b) bewiesene Aussage

$$\Sigma \mu(A_k) < \infty \Rightarrow \mu\Big(\bigcap_m \bigcup_{k \geq m} A_k\Big) = 0$$

(d.h. $\mu(\{s : s \in A_k \text{ unendlich oft}\}) = 0$)

heißt 1. Borel - Cantelli - Lemma.]

c) folgt leicht aus der Tschebyscheffschen Ungleichung:

$$\mu(\{|f_n - f| \geq \varepsilon\}) \leq \frac{1}{\varepsilon} \int_S |f_n - f| \, d\mu \to 0 \quad \forall \varepsilon > 0$$

Die volle Umkehrung von 3.9 a) gilt nicht (nicht einmal für $\mu(S) \ldots$

Auf $(]0,1], \mathrm{Bor}, \lambda)$ betrachte nämlich

$1_{]0,1]}, \ 1_{]0,1/2]}, \ 1_{]1/2,1]}, \ 1_{]0,1/4]}, \ldots, 1_{]3/4,1]}, \ 1_{]0,1/8]}, \ldots$

Natürlich gilt 3.9 a) auch i.a. nicht mehr für $\mu(S) = \infty$:

$f_n = 1_{\{n\}}$ auf $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \text{zählendes Maß})$.

Der Beweis von 3.9 erstreckt sich im übrigen wörtlich auf $\mathbb{C}$-wertige Funktionen; dasselbe gilt für die übrigen Konvergenzsätze.

II.4 Die $\mathcal{L}^p$ - Räume

In diesem Abschnitt behandeln wir Vektorräume meßbarer Funktionen.

Im folgenden ist $(S, \mathcal{O}\!\!\iota, \mu)$ ein fester Maßraum.

4.1 Definition

a) $\mathcal{L}^0(S, \mathcal{O}\!\!\iota, \mu) := \{ f : S \to \mathbb{R}, \ f \text{ meßbar} \}$

b) Für $0 < p < \infty$, $p \in \mathbb{R}$, setze

$\mathcal{L}^p(S, \mathcal{O}\!\!\iota, \mu) := \{ f : S \to \mathbb{R}, \ f \text{ meßbar}, \ \int_S |f|^p \, d\mu < \infty \}$.

c) $\mathcal{L}^\infty(S, \mathcal{O}\!\!\iota, \mu) := \{ f : S \to \mathbb{R}, \ f \text{ meßbar}, \ \exists_{\alpha \geq 0} \ \mu(\{|f| > \alpha\}) = 0 \}$.

Der Buchstabe $\mathcal{L}$ (der bald in ein normales L verwandelt wird) soll an Lebesgue erinnern; daß der Exponent in allen Büchern p heißt, hat historische Gründe: p wie frz. "puissance" (= Potenz).

Statt $\mathcal{L}^p(S, \mathcal{O}\!\!\iota, \mu)$ wird in der Regel $\mathcal{L}^p(\mu)$ geschrieben; ist $S \subset \mathbb{R}^k$, $\mathcal{O}\!\!\iota = \mathrm{Bor}(S)$, μ = Lebesguemaß auf S, ist auch die Notation $\mathcal{L}^p(S)$ (z.B. $\mathcal{L}^p[0,1]$) gebräuchlich.

Die wichtigsten Fälle in Def. 4.1 sind $p \geq 1$ bzw. $p = \infty$, denn diese führen auf Banachräume (s.u.). Unter diesen haben wieder $p = 1, 2, \infty$ die größte Bedeutung.

Daß $\mathcal{L}^0(\mu)$ ein Vektorraum ist, wurde bereits in 1.3 bemerkt. (Die algebraischen Operationen sind hier natürlich punktweise definiert, also

$(f+g)(s) = f(s) + g(s)$, $\quad (c f)(s) = c \cdot f(s)$.)

Für $\mathcal{L}^\infty(\mu)$ ist die Vektorraumstruktur wegen

$$\mu(\{|f|>\alpha\})=0 \quad\Rightarrow\quad \mu(\{|cf|>|c|\cdot\alpha\})=0$$

$$\mu(\{|f|>\alpha\})=0=\mu(\{|g|>\beta\}) \quad\Rightarrow\quad \mu(\{|f+g|>\alpha+\beta\})=0$$

klar. Im Bereich $0<p<\infty$ schließt man so:

$$\int_S |f+g|^p \, d\mu \underset{\uparrow}{\le} \int_S 2^p \cdot \max\{|f(s)|^p, |g(s)|^p\} \, d\mu(s)$$

da $|f(s)+g(s)| \le |f(s)|+|g(s)| \le 2\cdot \max\{|f(s)|, |g(s)|\}$

$$\le 2^p \int_S (|f(s)|^p + |g(s)|^p) \, d\mu(s)$$

$$= 2^p \left[\int_S |f|^p \, d\mu + \int_S |g|^p \, d\mu\right] < \infty .$$

(Beachte, daß alle auftretenden Integranden wirklich meßbar sind!)

$$\int_S |cf|^p \, d\mu = |c|^p \int_S |f|^p \, d\mu < \infty .$$

In diesem Abschnitt wird gezeigt, wie die Räume $\mathcal{L}^p(\mu)$ zu vollständigen (halb-) metrischen Räumen gemacht werden können. Zur Erinnerung:

Eine Metrik auf einer Menge M ist eine Abbildung $d: M\times M \to \mathbb{R}$ mit den Eigenschaften:

(i) $x = y \iff d(x,y) = 0$

(ii) $d(x,y) = d(y,x) \quad \forall\, x,y \in M$

(iii) $d(x,z) \le d(x,y) + d(y,z) \quad \forall\, x,y,z \in M$.

Gilt in (i) nur „$\Rightarrow$", so heißt d Halbmetrik.

Eine Norm auf einem $\mathbb{R}$-Vektorraum X ist eine Abbildung $\|\cdot\|: X \to \mathbb{R}$ mit den Eigenschaften,

(i) $x = 0 \iff \|x\| = 0$

(ii) $\|cx\| = |c|\cdot\|x\| \quad \forall\, c\in\mathbb{R},\ x\in X$

(iii) $\|x+y\| \le \|x\| + \|y\| \quad \forall\, x,y \in X$.

[In $\mathbb{C}$-Vektorräumen hat man natürlich $c\in\mathbb{C}$ zu fordern.]

Gilt in (i) nur „$\Rightarrow$", so heißt $\|\cdot\|$ Halbnorm. (Übrigens folgt aus (ii) sowieso $\|0\| = \|0\cdot x\| = 0\cdot\|x\| = 0$.)

Jede (Halb-) Norm generiert auf natürliche Weise eine (Halb-) Metrik:

$$d(x,y) := \|x-y\| .$$

Sei nun $1 \le p < \infty$. Für $f\in\mathcal{L}^p(\mu)$ setze

$$\|f\|_{\mathcal{L}^p} := \left(\int_S |f|^p \, d\mu\right)^{1/p} .$$

Unser nächstes Ziel ist es, die Halbnormeigenschaften für $\|\cdot\|_{\mathcal{L}^p}$ nachzuweisen. (ii) ist hier klar; und (iii) (die „Dreiecksungleichung") ist im Fall $p=1$ ebenfalls trivial, im Fall $p>1$ jedoch ganz und gar nicht. Hier hilft folgende wichtige Ungleichung weiter.

4.2 Satz (Höldersche Ungleichung)

Sei $1<p<\infty$, $q = \frac{p}{p-1}$ (so daß $\frac{1}{p}+\frac{1}{q}=1$). Ist $f\in\mathcal{L}^p(\mu)$ und $g\in\mathcal{L}^q(\mu)$, so ist $f\cdot g \in \mathcal{L}^1(\mu)$ mit

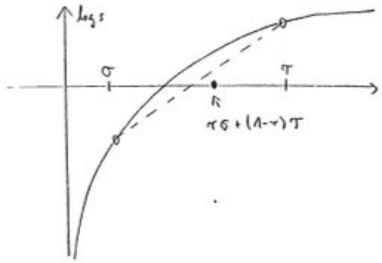
$$\left|\int_S f\cdot g \, d\mu\right| \le \|f\cdot g\|_{\mathcal{L}^1} \le \|f\|_{\mathcal{L}^p}\cdot\|g\|_{\mathcal{L}^q} .$$

Beweis: Wir erinnern zunächst an die „gewichtete Ungleichung vom arithmetischen und geometrischen Mittel“:

(*) Für $\sigma, \tau \geq 0$ und $0 < r < 1$ gilt $\sigma^r \cdot \tau^{1-r} \leq r\sigma + (1-r)\tau$.

⌈ Das ist klar, falls $\sigma = 0$ oder $\tau = 0$. Für $\sigma, \tau > 0$ ist (*) jedoch äquivalent zur Konkavität der Logarithmusfunktion:

$$\log(\sigma^r \tau^{1-r}) = r \log \sigma + (1-r) \log \tau \leq \log(r\sigma + (1-r)\tau)$$



Eine C^2-Funktion φ ist aber genau dann konkav, wenn $\varphi'' \leq 0$ gilt (Beweis?); und für $\varphi = \log$ ist wirklich

$$\varphi''(t) = -\frac{1}{t^2} \leq 0 .$$ ⌋

Zum Beweis der Hölderschen Ungleichung setze $A = \|f\|^p_{\mathcal{L}^p}$, $B = \|g\|^q_{\mathcal{L}^q}$. O.B.d.A. sind $A, B > 0$ (ist etwa $A = 0$, so ist $f = 0$ f.ü. und deshalb auch $f \cdot g = 0$ f.ü.). Wende nun (*) für

$$\sigma = \frac{|f(s)|^p}{A}, \quad \tau = \frac{|g(s)|^q}{B}, \quad r = \frac{1}{p} \quad (\Rightarrow 1-r = \tfrac{1}{q})$$

an und integriere:

$$\int_S \frac{|f(s)|}{A^{1/p}} \frac{|g(s)|}{B^{1/q}} \, d\mu(s) \leq \frac{1}{p}\left(\frac{1}{A}\int_S |f|^p \, d\mu\right) + \frac{1}{q}\left(\frac{1}{B}\int_S |g|^q \, d\mu\right)$$

$$= \frac{1}{p} + \frac{1}{q} = 1$$

$$\Rightarrow \left|\int_S f \cdot g \, d\mu\right| \underset{2.10}{\leq} \int_S |f \cdot g| \, d\mu \leq A^{1/p} \cdot B^{1/q} = \|f\|_{\mathcal{L}^p} \cdot \|g\|_{\mathcal{L}^q}$$

<u>4.3 Korollar</u> (Minkowskische Ungleichung)

Für $1 \leq p < \infty$ und $f, g \in \mathcal{L}^p(\mu)$ ist

$$\|f+g\|_{\mathcal{L}^p} \leq \|f\|_{\mathcal{L}^p} + \|g\|_{\mathcal{L}^p} .$$

Beweis: Wir dürfen $p > 1$ annehmen (für $p = 1$ ist die Sache trivial). Dann ist

$$\|f+g\|^p_{\mathcal{L}^p} = \int_S |f+g|^p \, d\mu$$
$$= \int_S |f+g| \cdot |f+g|^{p-1} \, d\mu$$
$$\leq \int_S |f| \cdot |f+g|^{p-1} \, d\mu + \int_S |g| \, |f+g|^{p-1} \, d\mu$$

Nun ist $|f+g|^{p-1} \in \mathcal{L}^q(\mu)$ $\left(\frac{1}{p} + \frac{1}{q} = 1\right)$, da ja $\int (|f+g|^{p-1})^q = \int |f+g|^p < \infty$, denn $\mathcal{L}^p(\mu)$ ist ein Vektorraum. Die Höldersche Ungleichung liefert daher

$$(\ldots) \leq \|f\|_{\mathcal{L}^p} \cdot \| \, |f+g|^{p-1} \|_{\mathcal{L}^q} + \|g\|_{\mathcal{L}^p} \cdot \| \, |f+g|^{p-1} \|_{\mathcal{L}^q}$$
$$= \left(\|f\|_{\mathcal{L}^p} + \|g\|_{\mathcal{L}^p}\right) \left(\int_S |f+g|^p \, d\mu\right)^{1/q}$$
$$= \left(\|f\|_{\mathcal{L}^p} + \|g\|_{\mathcal{L}^p}\right) \|f+g\|^{p-1}_{\mathcal{L}^p}$$

Also ist für $1 \leq p < \infty$ $(\mathcal{L}^p(\mu), \|\cdot\|_{\mathcal{L}^p})$ ein halbnormierter Raum. Konvergenz in $\mathcal{L}^p(\mu)$ wird – in Analogie zu 3.6 – Konvergenz im p-ten Mittel genannt.

Die entscheidende Eigenschaft der $\mathcal{L}^p$-Räume ist ihre <u>Vollständigkeit</u>.*)

Bekanntlich heißt ein (halb-) metrischer Raum vollständig, wenn jede

*) Dieser Vollständigkeitsbegriff darf nicht mit dem eines vollständigen Maßraums (S.40) verwechselt werden!

Cauchyfolge konvergiert. Wichtige <u>Bemerkung</u>: Da eine Cauchyfolge bereits konvergiert, wenn sie eine konvergente Teilfolge hat (nicht wahr?), reicht es, zum Beweis der Vollständigkeit diese etwas schwächere Bedingung zu zeigen. Tun wir das im Beweis von

<u>4.4 Theorem</u> (Fischer / Riesz)

Für $1 \le p < \infty$ ist $\mathcal{L}^p(\mu)$ ein vollständiger halbnormierter Raum.

Beweis: Sei (f_n) eine Cauchyfolge in $\mathcal{L}^p(\mu)$. Dann existiert eine Teilfolge mit

$$\| f_{n_{k+1}} - f_{n_k} \|_{\mathcal{L}^p} \le 2^{-k}.$$

Setze

$$g_m = \sum_{k=1}^{m} |f_{n_{k+1}} - f_{n_k}|$$

sowie

$$g = \sum_{k=1}^{\infty} |f_{n_{k+1}} - f_{n_k}| .$$

(g ist i.a. $[0,\infty]$-wertig.) Da $\mathcal{L}^p(\mu)$ ein Vektorraum ist, ist [illegible] und nach der Minkowskischen Ungleichung gilt

$$\|g_m\|_{\mathcal{L}^p} \le \sum_{k=1}^{m} \| |f_{n_{k+1}} - f_{n_k}| \|_{\mathcal{L}^p} = \sum_{k=1}^{m} \| f_{n_{k+1}} - f_{n_k} \|_{\mathcal{L}^p} \le \sum_{k=1}^{m} 2^{-k} \le 1.$$

Nun konvergiert (g_m) monoton gegen g, also auch $g_m^p \nearrow g^p$; der Satz von Beppo Levi (2.5) zeigt daher

$$\int_S g^p \, d\mu = \lim_{m\to\infty} \int_S g_m^p \, d\mu \le 1.$$

Konsequenz (3.3): g ist f.ü. endlich, sagen wir außerhalb einer Nullmenge N. Für $s \notin N$ konvergiert also die Reihe $\sum_{k=1}^{\infty} (f_{n_{k+1}}(s) - f_{n_k}(s))$ absolut, wegen der Vollständigkeit von $\mathbb{R}$ konvergiert sie daher! Setze

$$\tilde f(s) = \sum_{k=1}^{\infty} (f_{n_{k+1}}(s) - f_{n_k}(s)) \qquad \text{für } s \notin N$$

$$\tilde f(s) = 0 \qquad \text{für } s \in N.$$

Als punktweiser Limes meßbarer Funktionen ist $\tilde f$ meßbar; und es gilt $|\tilde f| \le g$, daher auch $|\tilde f|^p \le g^p$. Es folgt wegen

$$\int_S |\tilde f|^p \, d\mu \le \int_S g^p \, d\mu \le 1$$

$\tilde f \in \mathcal{L}^p(\mu)$. Zeigen wir jetzt

$$\sum_{k=1}^{\infty} (f_{n_{k+1}} - f_{n_k}) = \tilde f \qquad \text{bzgl. } \|\cdot\|_{\mathcal{L}^p}$$

Es ist also behauptet, daß die Partialsummen im p-ten Mittel gegen $\tilde f$ konvergieren. Nach Definition liegt f.ü.-Konvergenz vor. Für $h_n = |\tilde f - \sum_{k=1}^{n} (f_{n_{k+1}} - f_{n_k})|^p$ gilt daher $h_n \to 0$ f.ü. Andererseits ist

$$h_n \le \Big(\sum_{k>n} |f_{n_{k+1}} - f_{n_k}| \Big)^p \le g^p ,$$

was integrierbar ist. Der Lebesguesche Konvergenzsatz zeigt

$$\int_S h_n \, d\mu \to 0$$

was zu zeigen war.

Nun kommt endlich der Teleskoptrick: So eine Partialsumme ist ja nichts anderes als $f_{n_{m+1}} - f_{n_1}$, die Teilfolge (f_{n_m}) konvergiert also gegen $\tilde f + f_{n_1} \in \mathcal{L}^p(\mu)$ im p-ten Mittel. Das war zu zeigen.

Es ist üblich in der $\mathcal{L}^p$-Theorie, fast überall übereinstimmende Funktionen zu identifizieren. Mathematisch präzisiert bedeutet das den Übergang zu einem Quotientenvektorraum. Setze nämlich

$$V = \{ f: S \to \mathbb{R},\ f \text{ meßbar},\ f = 0 \quad \mu\text{-f.ü.} \}.$$

und betrachte den Quotientenraum

$$L^p(\mu) := \mathcal{L}^p(\mu) / V,$$

dessen Elemente (einstweilen) mit $[f]$ bezeichnet werden. Beachte, daß V mit dem „Kern" der Halbnorm $\|\cdot\|_{\mathcal{L}^p}$ übereinstimmt (3.3

Es folgt, daß

$$\|[f]\|_{L^p} := \|f\|_{\mathcal{L}^p}$$

eine wohldefinierte Abbildung liefert:

$$[f] = [g] \Rightarrow [f-g] = 0$$
$$\Rightarrow \|f\|_{\mathcal{L}^p} \le \|f-g\|_{\mathcal{L}^p} + \|g\|_{\mathcal{L}^p} = \|g\|_{\mathcal{L}^p}$$
$$\le \|g-f\|_{\mathcal{L}^p} + \|f\|_{\mathcal{L}^p} = \|f\|_{\mathcal{L}^p}$$

(hier haben wir $f = (f-g) + g$ und die Dreiecksungleichung verwandt)

Es ist klar, daß $\|\cdot\|_{L^p}$ eine Norm auf $L^p(\mu)$ ist:

z.B. $\|[f]+[g]\|_{L^p} = \|[f+g]\|_{L^p} = \|f+g\|_{\mathcal{L}^p}$
$$\le \|f\|_{\mathcal{L}^p} + \|g\|_{\mathcal{L}^p} = \|[f]\|_{L^p} + \|[g]\|_{L^p}$$

$$\|[f]\|_{L^p} = 0 \iff \|f\|_{\mathcal{L}^p} = 0 \iff f \in V$$

Auch die Vollständigkeit bleibt erhalten, denn $([f_n])$ ist Cauchyfolge (bzw. konvergiert) genau dann, wenn (f_n) Cauchyfolge ist (bzw. konvergiert).

Vollständige normierte Räume heißen Banachräume; es gilt also folgende Version des Riesz-Fischer Theorems:

4.4* Theorem Für $1 \le p < \infty$ ist $L^p(\mu)$ ein Banachraum.

Für den Umgang mit L^p-Räumen hat sich folgende pragmatische Nomenklatur als praktisch herausgestellt: Obwohl die Elemente von $L^p(\mu)$ Äquivalenzklassen von Funktionen sind, tut man so, als wären es echte Funktionen. Z.B. ist ein Satz von der Form „Wähle $f \in L^p[0,1]$ mit $\int_0^1 f\,d\lambda = 0$" in diesem Sinne zu verstehen. Nicht sinnvoll ist hingegen „Wähle $f \in L^p[0,1]$ mit $f(0) = 0$", denn verschiedene Repräsentanten derselben Äquivalenzklasse können sich bei $s=0$ sehr wohl unterscheiden, ihre Integrale sind jedoch stets dieselben.

In diesem Sinn gilt die Hölder'sche Ungleichung wörtlich in der L^p-Version.

Die Banachräume $L^p(\mu)$ bilden das fundamentale Beispielreservoir der Funktionalanalysis. Der Fall $p=2$ ist von besonderem Interesse, denn $L^2(\mu)$ ist ein Hilbertraum; d.h. seine Norm wird von einem Skalarprodukt, nämlich

$$\langle f, g \rangle_{L^2} = \int_S f(s)\,\overline{g(s)}\,d\mu(s)$$

induziert. (Der Querstrich bedeutet komplexe Konjugation; die L^p-Theorie für komplexe Funktionen ist natürlich vollständig analog.)

Für den Maßraum $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \text{zählendes Maß})$ schreibt man übrigens ℓ^p statt $L^p(\mu)$. Explizit ist (vgl. 2.11 d))

$$\ell^p = \left\{ (a_n) : \left(\sum_{n=1}^{\infty} |a_n|^p\right)^{1/p} < \infty \right\}.$$

Bevor wir einige Eigenschaften der L^p-Räume formulieren, sollen die übrigen in Def. 4.1 genannten Fälle behandelt werden.

Zuerst nun zu $p = \infty$.

Für $f \in \mathcal{L}^\infty(\mu)$ setze

$$\|f\|_{\mathcal{L}^\infty} = \inf\{\alpha : \mu(\{|f| > \alpha\}) = 0\}.$$

Auf S. 92 oben ist $\|\cdot\|_{\mathcal{L}^\infty}$ als Halbnorm erkannt worden. Offenbar ist

$$(*) \qquad \|f\|_{\mathcal{L}^\infty} = \inf_{\substack{N \in \mathfrak{A} \\ \mu(N) = 0}} \ \sup_{s \in S \setminus N} |f(s)| \ ;$$

daher wird $\|\cdot\|_{\mathcal{L}^\infty}$ auch „wesentliche Supremums(halb)norm" genannt.

Damit kann man nun den Grenzfall $p = 1$ der Hölderschen Un-gleichung formulieren:

$$f \in \mathcal{L}^1(\mu),\ g \in \mathcal{L}^\infty(\mu) \Rightarrow f \cdot g \in \mathcal{L}^1(\mu) \text{ mit } \|f \cdot g\|_{\mathcal{L}^1} \le \|f\|_{\mathcal{L}^1} \cdot \|g\|$$

(Beweis zur Übung!) Das ist ein Indiz, daß die Bezeichnung $\mathcal{L}^\infty$ w… Bedacht gewählt ist. (Ein anderes: Für ein endliches Maß ist

$$\lim_{p \to \infty} \|f\|_{\mathcal{L}^p} = \|f\|_{\mathcal{L}^\infty} \qquad \text{für } f \in \mathcal{L}^\infty(\mu).$$

Beweis ebenfalls zur Übung.)

Kommen wir zur Vollständigkeit von $\mathcal{L}^\infty(\mu)$.

Beginnen wir mit der Vorbemerkung, daß das inf in $(*)$ angenommen wird:

⌜ Wähle $N_k \in \mathfrak{A}$, $\mu(N_k) = 0$ mit

$$\|f\|_{\mathcal{L}^\infty} \ge \sup_{s \notin N_k} |f(s)| - \frac{1}{k}.$$

$N = \bigcup N_k$ ist dann eine Nullmenge mit

$$\|f\|_{\mathcal{L}^\infty} \ge \sup_{s \notin N} |f(s)| \qquad (\ge \|f\|_{\mathcal{L}^\infty}). \ ⌟$$

Sei nun (f_n) eine Cauchyfolge in $\mathcal{L}^\infty(\mu)$. Wir wählen Nullmengen $N_{n,m}$ nach Vorbemerkung, so daß

$$\|f_n - f_m\|_{\mathcal{L}^\infty} = \sup_{s \notin N_{n,m}} |f_n(s) - f_m(s)| \qquad \forall\, n,m \in \mathbb{N}.$$

Erst recht ist für die Nullmenge $N = \bigcup_{n,m} N_{n,m}$

$$\|f_n - f_m\|_{\mathcal{L}^\infty} = \sup_{s \notin N} |f_n(s) - f_m(s)| \qquad \forall\, n,m \in \mathbb{N}$$

⌜ $\le$ nach Def. von $\|\cdot\|_{\mathcal{L}^\infty}$, $\ge$ da $N_{n,m} \subset N$ ⌟

Mit $g_n = f_n \cdot 1_{N^c}$ erhält man

$f_n = g_n$ f.ü., g_n beschränkt und meßbar auf S.

(g_n) ist Cauchyfolge bzgl. der üblichen Supremumsnorm*)

Nun borgen wir uns das Resultat, daß die beschränkten Funktionen auf einer Menge S bzgl. der Supremumsnorm einen Banachraum bilden (z.B. Heuser, Analysis II, S. 16): (g_n) konvergiert gleichmäßig gegen eine beschränkte Funktion g. Diese muß wegen 1.4 meßbar sein, also $g \in \mathcal{L}^\infty(\mu)$. Bleibt zu zeigen, daß g (ein!) Limes von (f_n) bzgl. $\|\cdot\|_{\mathcal{L}^\infty}$ ist:

*) $\|g\|_\infty = \sup_{s \in S} |g(s)|$

$$\|f_n - g\|_{\mathcal{L}^\infty} \le \sup_{s\notin N} |f_n(s) - g(s)| = \sup_{s\in S} |g_n(s) - g(s)| \xrightarrow[n\to\infty]{} 0$$

Wie bei den L^p-Räumen mit $p<\infty$ geht man zum normierten Quotienten

$$L^\infty(\mu) := \mathcal{L}^\infty(\mu)/V$$

nach $V = \{f : S \to \mathbb{R}$ meßbar, $f = 0$ f.ü.$\}$ über. Die entsprechende Norm wird mit $\|\cdot\|_{L^\infty}$ bezeichnet, sie heißt wesentliche Supremumsnorm. Die Elemente von $L^\infty(\mu)$ werden wir im allgemeinen wieder als Funktionen (statt Äquivalenzklassen von Funktionen, wie es eigentlich korrekt wäre) ansehen, die bzgl. fast überall bestehender Gleichheit identifiziert werden.

Zusammengefaßt ist gezeigt:

4.5 Satz $L^\infty(\mu)$, versehen mit der wesentlichen Supremumsnorm $\|\cdot\|$, ist ein Banachraum.

Als nächstes zu $0<p<1$. Die entsprechenden L^p-Räume werden keine Banachräume werden und sind von weniger großer Bedeutung. Wir werden uns hier wesentlich kürzer fassen.

Auf $\mathcal{L}^p(\mu)$ mit $0<p<1$ betrachtet man

$$d_p(f,g) = \int_S |f-g|^p \, d\mu .$$

Es ist nicht schwer zu sehen, daß d_p eine Halbmetrik ist.

Man kann dann zeigen, daß $\mathcal{L}^p(\mu)$, versehen mit d_p, ein vollständiger halbmetrischer Raum ist; der Übergang zum Quotientenraum $L^p = \mathcal{L}^p/V$ liefert einen vollständigen metrischen Raum. L^p ist (fast) wie ein Banachraum im Bereich $0<p<1$.

Abschließend betrachten wir den Raum $\mathcal{L}^0(\mu)$.

Für $f,g \in \mathcal{L}^0(\mu)$ setze*)

$$d_0(f,g) = \inf\{\varepsilon>0 : \mu(\{|f-g| \ge \varepsilon\}) \le \varepsilon\}.$$

4.6 Satz a) d_0 ist eine Halbmetrik.**)

b) $d_0(f_n, f) \to 0 \iff f_n \xrightarrow{\mu} f$

c) (f_n) ist d_0-Cauchyfolge $\iff$ (f_n) ist Maßkonvergenz-Cauchyfolge
(d.h. $\forall \varepsilon>0\ \forall \tilde\varepsilon>0\ \exists N\ \forall n,m \ge N\ \mu(\{|f_n - f_m| \ge \varepsilon\}) \le \tilde\varepsilon$)

d) $d_0(f,0) = 0 \iff f = 0$ f.ü.

e) $(\mathcal{L}^0(\mu), d_0)$ ist vollständig.

Satz 4.6 enthält insbesondere ein Cauchy-Kriterium für Maßkonvergenz:
Jede μ-Cauchyfolge konvergiert dem Maße nach.

*) Für die Definition von $\mathcal{L}^0(\mu)$ war die Anwesenheit eines Maßes unerheblich. Erst d_0 nimmt explizit auf μ Bezug.

**) Das ist etwas geschummelt, da d_0 auch den Wert ∞ annehmen kann. Das ist zwar unerheblich, doch auch vermeidbar bei Bedarf: Wen das stört, sollte mit $\arctan d_0$ statt d_0 operieren!

Beweis: a) Nur die Dreiecksungleichung bedarf einer Überlegung.
Seien $\varepsilon_1 \geq d_0(f,g)$, $\varepsilon_2 \geq d_0(g,h)$ (sollte ein „Abstand" ∞ sein, ist nichts zu zeigen) mit

$$\mu(\{|f-g| \geq \varepsilon_1\}) \leq \varepsilon_1$$
$$\mu(\{|g-h| \geq \varepsilon_2\}) \leq \varepsilon_2 .$$

Da
$$\{|f-h| \geq \varepsilon_1+\varepsilon_2\} \subset \{|f-g| + |g-h| \geq \varepsilon_1+\varepsilon_2\} \subset \{|f-g| \geq \varepsilon_1\} \cup \{|g-h| \geq \varepsilon_2\}$$
folgt
$$\mu(\{|f-h| \geq \varepsilon_1+\varepsilon_2\}) \leq \varepsilon_1 + \varepsilon_2 ,$$
also
$$d_0(f,h) \leq \varepsilon_1+\varepsilon_2 .$$
Übergang zum Infimum zeigt die Behauptung.

b) $d_0(f_n, f) \to 0$ bedeutet ausführlich
$$\forall_{\eta>0}\ \exists N\ \forall n \geq N \quad \inf\{\cdots\} < \eta$$
d.h.
$$\forall_{\eta>0}\ \exists N\ \forall_{n\geq N}\ \exists \varepsilon<\eta \quad \mu(\{|f_n-f| \geq \varepsilon\}) \leq \varepsilon .$$
$f_n \xrightarrow{\mu} f$ bedeutet ausführlich
$$\forall \varepsilon_{1/2}>0\ \exists N\ \forall n\geq N \quad \mu(\{|f_n-f| \geq \varepsilon_1\}) \leq \varepsilon_2$$
Damit „$\Leftarrow$", indem das $\xrightarrow{\mu}$-Kriterium für $\varepsilon_1 = \varepsilon_2 < \eta$ angewandt wird.
„$\Rightarrow$": Wegen $\mu(\{|f_n-f| \geq \eta\}) \leq \mu(\{|f_n-f| \geq \varepsilon\})$ hat man im $\xrightarrow{d_0}$-Kriterium sogar
$$\underset{\eta>0}{\forall}\ \underset{N}{\exists}\ \underset{n\geq N}{\forall} \quad \mu(\{|f_n-f| \geq \eta\}) \leq \eta .$$

Wendet man das auf $\eta = \min\{\varepsilon_1, \varepsilon_2\}$ an, bekommt man das $\xrightarrow{\mu}$-Kriterium.

c) geht analog.

d) kann aus 3.7 c) geschlossen werden: $d_0(f,0) = 0$ bedeutet, daß die konstante Folge $f, f, f, \cdots$ bzgl. $\xrightarrow{d_0}$ gegen 0 konvergiert(!). Da f natürlich selbst ebenfalls Limes ist, folgt aus b) und 3.7. c) $f = 0$ f.ü. Die Umkehrung ist klar.

e) Wegen der Bemerkung auf S. 96 oben genügt es, aus einer d_0-Cauchyfolge eine d_0-konvergente Teilfolge auszusondern.
Nun ist eine d_0-Cauchyfolge eine μ-Cauchyfolge, und wie in 3.9 kann man zeigen, daß es eine f.ü.-Cauchyteilfolge gibt. Diese konvergiert f.ü. gegen eine (o.E.) meßbare Funktion f. Die Vollständigkeit von $\mathcal{L}^0(\mu)$ folgt nun (durch Übergang zu $f_n - f$) aus der folgenden Hilfsbehauptung:

(f_n) μ-Cauchy, $f_n \to 0$ f.ü. $\Rightarrow$ es ex. Teilfolge $f_{n_k} \xrightarrow{\mu} 0$.

Der Beweis hierfür ähnelt dem von 3.9 b).[*)]

Die Cauchy-Eigenschaft liefert eine Teilfolge mit
$$\mu(\{|f_{n_{k+1}} - f_{n_k}| \geq 2^{-k}\}) \leq 2^{-k}$$
Setze $A_m = \bigcup_{k\geq m} \{|f_{n_{k+1}} - f_{n_k}| \geq 2^{-k}\}$, so daß $\mu(A_m) \leq 2 \cdot 2^{-m}$.
Auf A_m^c hingegen konvergiert $\sum_{k=m}^{\infty} (f_{n_{k+1}} - f_{n_k})$ gleichmäßig (da $\sum_{k=m}^{\infty} |f_{n_{k+1}} - f_{n_k}| \leq \sum_{k=m}^{\infty} 2^{-k}$ auf A_m^c), also (Teleskopsumme) konvergiert (f_{n_k}) dort gleichmäßig, und die Limesfunktion f muß f.ü. $= 0$ sein, da $f_n \to 0$ f.ü.
Bei gegebenem $\varepsilon>0$ ist daher
$$\mu(\{|f_{n_k}| \geq \varepsilon\} \cap A_m^c) = 0$$

*) Für endliches μ ist übrigens nur 3.9 a) anzuwenden.

für hinreichend große $k \geq m$, und für solche k ist dann

$$\mu\{|f_{n_k}| \geq \varepsilon\} = \mu(\{|f_{n_k}| \geq \varepsilon\} \cap A_m) \leq \mu(A_m) \leq 2^{-m} \xrightarrow[m\to\infty]{} 0.$$

Wie in den übrigen Fällen auch, verschiebt man abschließend d_0 auf den Quotientenraum $\mathcal{L}^0(\mu)/V$ (V wie oben) und erhält so den vollständigen Raum $L^0(\mu)$.

Am Schluß dieses Abschnitts folgen noch zwei bisweilen nützliche Dichtheitsaussagen. Eine Teilmenge D eines metrischen Raums M heißt bekanntlich dicht, wenn ihr Abschluß $\overline{D} = M$ ist.

4.7 Satz Für $1 \leq p \leq \infty$ liegen die Treppenfunktionen in $\mathcal{L}^p(\mu)$ dicht in $L^p(\mu)$ (bzgl. $\|\cdot\|_{L^p}$).

Beweis: Für $p = \infty$ folgt das sofort aus 1.8. Wir dürfen also $p < \infty$ annehmen. Ist $f \geq 0$, so existieren Treppenfunktionen $0 \leq f_n \nearrow f$. Wegen

$$|f - f_n|^p \leq (|f| + |f_n|)^p \underset{0 \leq f_n \leq f}{\leq} 2^p f^p$$

und $f \in \mathcal{L}^p(\mu)$ zeigt der Lebesguesche Konvergenzsatz

$$\|f - f_n\|_{L^p} = \left(\int_S |f - f_n|^p \, d\mu\right)^{1/p} \to 0.$$

Im allgemeinen Fall zerlege in $f^+ - f^-$ und wende den 1. Teil an; komplexe Funktionen zerlege vorher in $\operatorname{Re} f + i \operatorname{Im} f$.

Die Aussage von 4.7 gilt auch für $0 \leq p < 1$.

Etwas schärfer ist das nächste Resultat, das hier nur für endliche Maßräume formuliert werden soll (es gibt auch eine Version im σ-endlichen Fall):

4.8 Satz $(S, \mathfrak{A}, \mu)$ sei ein endlicher Maßraum, und $\mathfrak{A}_0 \subset \mathfrak{A}$ sei eine Algebra mit $\sigma(\mathfrak{A}_0) = \mathfrak{A}$. Dann liegen die $\mathfrak{A}_0$-Treppenfunktionen (d.h. die lineare Hülle von $\{1_A : A \in \mathfrak{A}_0\}$) dicht in $L^p(\mu)$, falls $1 \leq p < \infty$.

Beweis: Zunächst zeigen wir

$$\forall \varepsilon > 0 \ \forall A \in \mathfrak{A} \ \exists B \in \mathfrak{A}_0 : \ \mu(A \,\Delta\, B) \leq \varepsilon.$$

Zur Erinnerung: $A \,\Delta\, B = A \setminus B \cup B \setminus A = (A \cup B) \setminus (A \cap B)$.

Fassen wir in $\overline{\mathfrak{A}_0}$ die Mengen aus $\mathfrak{A}$ zusammen, für die die Behauptung stimmt. Wegen $\mathfrak{A}_0 \subset \overline{\mathfrak{A}_0}$ ist nur zu zeigen, daß $\overline{\mathfrak{A}_0}$ eine σ-Algebra ist:

(i) $\emptyset, S \in \overline{\mathfrak{A}_0}$: klar

(ii) $A \in \overline{\mathfrak{A}_0} \Rightarrow A^c \in \overline{\mathfrak{A}_0}$: wegen $A \,\Delta\, B = A^c \,\Delta\, B^c$

(iii) $A_1, A_2, \dots \in \overline{\mathfrak{A}_0} \Rightarrow A := \bigcup_{n=1}^{\infty} A_n \in \overline{\mathfrak{A}_0}$:

Wähle n_0 so groß, daß $\mu(A \setminus \bigcup_{n=1}^{n_0} A_n) < \frac{\varepsilon}{2}$. Alsdann wähle zu A_n $(n \leq n_0)$ $B_n \in \mathfrak{A}_0$ mit $\mu(A_n \,\Delta\, B_n) \leq \frac{\varepsilon}{2} \cdot 2^{-n}$. Setze $B = \bigcup_{n=1}^{n_0} B_n \in \mathfrak{A}_0$ (!).

Es ist

$$\mu(A \,\Delta\, B) \leq \mu\Big(A \,\Delta\, \bigcup_{n=1}^{n_0} A_n\Big) + \mu\Big(\bigcup_{n=1}^{n_0} A_n \,\Delta\, B\Big)$$

[denn $A \,\Delta\, B = (A \,\Delta\, C) \,\Delta\, (C \,\Delta\, B) \subset (A \,\Delta\, C) \cup (C \,\Delta\, B)$]

$$\le \mu\Big(A \setminus \bigcup_{n=1}^{n_0} A_n\Big) + \mu\Big(\bigcup_{n=1}^{n_0} (A_n \,\Delta\, B)\Big)$$

$$\Big[\text{ denn } \big(\bigcup A_n\big) \,\Delta\, B \subset \bigcup (A_n \,\Delta\, B) \Big]$$

$$\le \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \sum_{n=1}^{n_0} 2^{-n}$$

$$\le \varepsilon .$$

Wegen $\| 1_A - 1_B \|_{L^p} = \big(\mu(A \,\Delta\, B)\big)^{1/p}$ heißt das:
Jede $\mathcal{A}$-Indikatorfunktion ist $\| \ \|_{L^p}$-Limes von $\mathcal{A}_0$-Indikatorfunktionen. Folglich ist auch jede $\mathcal{A}$-Treppenfunktion $\| \ \|_{L^p}$-Limes von $\mathcal{A}_0$-Treppenfunktionen (leichte Konsequenz aus der Dreiecksungleichung). Ein Blick auf 4.7 zeigt schließlich 4.8.

Für $p = \infty$ ist 4.8 falsch, genauso wie das folgende Korollar:

<u>4.9 Korollar</u> a) $C[a,b]$ (= die stetigen Funktionen) liegt dicht in $L^p[a,b]$,
b) $C^\infty[a,b]$ (= die unendlich häufig differenzierbaren Funktionen) liegt dicht in $L^p[a,b]$,
jeweils für $1 \le p < \infty$.

Beweis: a) $\mathcal{A}_0 = \mathcal{F}_1 \cap [a,b]$ ist eine Algebra, die $\mathrm{Bor}([a,b])$ erzeugt (I.1.9). 4.8 impliziert, daß es reicht, für $\varepsilon > 0$ und $a \le \alpha \le \beta \le b$ eine stetige Funktion f mit $\| f - 1_{[\alpha,\beta]} \|_{L^p}$ zu finden, was keine Kunst ist (solange $p < \infty$!).
b) folgt aus a), da nach dem Approximationssatz von Weierstraß jede stetige Funktion auf $[a,b]$ gleichmäßiger Limes von Polynomen ist, erst recht also $\| \ \|_{L^p}$-Limes.

<u>II.5 Bildmaße</u>

Seien $(S_1, \mathcal{A}_1)$ und $(S_2, \mathcal{A}_2)$ meßbare Räume und μ ein Maß auf $\mathcal{A}_1$. Ist $T : S_1 \to S_2$ eine meßbare Abbildung, so kann man das Maß auf $\mathcal{A}_2$ hinübertransportieren:

<u>5.1 Definition</u> $\mu \circ T^{-1} : \mathcal{A}_2 \to [0,\infty]$, $F \mapsto \mu(T^{-1}(F))$
heißt Bildmaß von μ unter T.

Es ist klar, daß das Bildmaß wirklich ein Maß ist.

<u>5.2 Beispiele</u> a) Besonders wichtig ist der Fall $(S_2, \mathcal{A}_2) = (\mathbb{R}, \mathrm{Bor}(\mathbb{R}))$. In diesem Fall heißt das Bildmaß <u>Verteilung von T</u>.
b) z.B. sei $S_1 = [0,1]$, $\mathcal{A}_1 = \mathrm{Bor}[0,1]$, $\mu = \lambda$ und $(S_2, \mathcal{A}_2) = (\mathbb{R}, \mathrm{Bor}(\mathbb{R}))$, $T : [0,1] \to \mathbb{R}$ die konstante Funktion $T(s) = c$. Dann ist die Verteilung von T
$\lambda \circ T^{-1} = \delta_c$, das Dirac-Maß bei c.
c) Es folgt aus I.4.4, daß für eine bijektive lineare Abbildung $T : \mathbb{R}^k \to \mathbb{R}^k$

$$\lambda^k \circ T^{-1} = |\det T|^{-1} \cdot \lambda^k$$

gilt.

Für die Integration bzgl. des Bildmaßes gilt das einfache, aber nützliche Resultat:

5.3 Satz (Transformationsformel)

Unter den Bezeichnungen von Def. 5.1 ist eine meßbare Funktion $f: S_2 \to \bar{\mathbb{R}}$ genau dann $\mu \circ T^{-1}$-integrierbar, wenn $f \circ T$ μ-integrierbar ist. In diesem Fall gilt

$$\int_F f \, d(\mu \circ T^{-1}) = \int_{T^{-1}(F)} f \circ T \, d\mu \qquad \forall F \in \mathfrak{A}_2.$$

Beweis: Der erste Teil ist mit

$$\int_{S_2} g \, d(\mu \circ T^{-1}) = \int_{S_1} g \circ T \, d\mu \qquad \text{für positive meßbare}$$

gezeigt.

Das ist nach Definition richtig, wenn g eine Indikatorfunktion ist, denn $1_F \circ T = 1_{T^{-1}(F)}$. Deshalb stimmt die auch für Treppenfunktionen (Integration ist linear) und für Limiten von Treppenfunktionen (Beppo Levi, 2.5), also für alle $g \geq 0$ (1.8).

Die zweite Aussage zeigt man nach demselben Schema zuerst für f^+

Beispiel: T sei reellwertig, dann ist mit $f(t) = |t|^r$

$$\int_{S_1} |T|^r \, d\mu = \int_{\mathbb{R}} |t|^r \, d(\mu \circ T^{-1})(t)$$

So können abstrakte Integrale in Integrale auf $\mathbb{R}$ überführt werden. Deshalb ist das Studium von Maßen auf $\mathbb{R}$ besonders wichtig.

II-6 Produktmaße und der Satz von Fubini

Gegeben sei eine meßbare Funktion $f: \mathbb{R}^2 \to \mathbb{R}$. Häufig trifft man in der Analysis auf das iterierte Integral $\int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x,y) \, dx \right) dy$ und dann das Problem, hier die Integrationsreihenfolge zu vertauschen. In diesem Abschnitt werden wir leicht verifizierbare Kriterien für die Vertauschbarkeit angeben, und zwar nicht nur für das Lebesguemaß, sondern für beliebige σ-endliche Maße. En passant treffen wir dabei auf die auch für sich interessante Konstruktion der sog. Produktmaßes.

6.1 Definition Seien $(S_1, \mathfrak{A}_1)$ und $(S_2, \mathfrak{A}_2)$ meßbare Räume. Die auf dem kartesischen Produkt $S_1 \times S_2$ von den „meßbaren Rechtecken" $A_1 \times A_2$ ($A_i \in \mathfrak{A}_i$) erzeugte σ-Algebra heißt Produkt-σ-Algebra und wird mit $\mathfrak{A}_1 \otimes \mathfrak{A}_2$ bezeichnet.

6.2 Lemma Sei $\pi_i: S_1 \times S_2 \to S_i$, $(s_1, s_2) \mapsto s_i$ die kanonische Projektion. Eine Abbildung T von einem meßbaren Raum $(S, \mathfrak{A})$ nach $(S_1 \times S_2, \mathfrak{A}_1 \otimes \mathfrak{A}_2)$ ist genau dann meßbar, wenn die $\pi_i \circ T$ $(\mathfrak{A}, \mathfrak{A}_i)$-meßbar sind.

Beweis: Die π_i sind jedenfalls meßbar, denn z.B. $\pi_1^{-1}(A_1) = A_1 \times S_2 \in \mathfrak{A}_1 \otimes \mathfrak{A}_2$.

Daher: T meßbar $\Rightarrow$ $\pi_i \circ T$ meßbar

Umgekehrt ist $T^{-1}(A_1 \times A_2) \in \mathfrak{A} \quad \forall A_i \in \mathfrak{A}_i$ zu zeigen (vgl. 1.2.a)). Beachte nur $A_1 \times A_2 = \pi_1^{-1}(A_1) \cap \pi_2^{-1}(A_2)$, um

$$T^{-1}(A_1 \times A_2) = (\pi_1 T)^{-1}(A_1) \cap (\pi_2 T)^{-1}(A_2) \in \mathfrak{A}$$

zu erhalten.

<u>6.3 Beispiele</u> a) Es ist $\mathrm{Bor}(\mathbb{R}^k) \otimes \mathrm{Bor}(\mathbb{R}^l) = \mathrm{Bor}(\mathbb{R}^{k+l})$.
„$\supset$" gilt, da $\mathrm{Bor}(\mathbb{R}^{k+l})$ von speziellen meßbaren Rechtecken, nämlich d Intervallen, erzeugt wird. Die Aussage „$\subset$" werden wir zunächst so übers daß 6.2 mit Gewinn angewandt werden kann: „$\subset$" bedeutet, daß d Identität auf $\mathbb{R}^{k+l}$ $\mathrm{Bor}(\mathbb{R}^{k+l})$ – $\mathrm{Bor}(\mathbb{R}^k) \otimes \mathrm{Bor}(\mathbb{R}^l)$ - meßbar ist, was nach 6.2 äquivalent zur Borel-Meßbarkeit der Projektionen von $\mathbb{R}^{k+l}$ a $\mathbb{R}^k$ bzw. $\mathbb{R}^l$ ist. Die sind stetig, also stimmt „$\subset$"!

b) Eine analoge Aussage für Lebesguemengen ist falsch (Übungsaufg

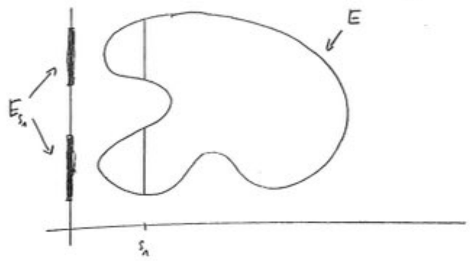
c) $\mathfrak{P}(\mathbb{N}) \otimes \mathfrak{P}(\mathbb{N}) = \mathfrak{P}(\mathbb{N} \times \mathbb{N})$, denn jedes $E \subset \mathbb{N} \times \mathbb{N}$ ist abzählbar, und Punkte liegen in $\mathfrak{P}(\mathbb{N}) \otimes \mathfrak{P}(\mathbb{N})$.
Hingegen gilt für Mengen S, deren Kardinalität größer als die von $\mathbb{R}$ ist, $\mathfrak{P}(S) \otimes \mathfrak{P}(S) \neq \mathfrak{P}(S \times S)$. Das ist eine etwas schwierigere Übungsaufgabe, vgl. etwa Behrends, S. 120.

Sei nun $E \subset S_1 \times S_2$. Für $s_1 \in S_1$ setze

$$E_{s_1} = \{ s_2 \in S_2 : (s_1, s_2) \in E \}$$

sowie für $s_2 \in S_2$

$$E^{s_2} = \{ s_1 \in S_1 : (s_1, s_2) \in E \}.$$



<u>6.4 Lemma</u> a) Für $E \in \mathfrak{A}_1 \otimes \mathfrak{A}_2$ gilt

$$E_{s_1} \in \mathfrak{A}_2 \quad \forall s_1 \in S_1, \qquad E^{s_2} \in \mathfrak{A}_1 \quad \forall s_2 \in S_2.$$

b) Ist $f : S_1 \times S_2 \to \overline{\mathbb{R}}$ eine $\mathfrak{A}_1 \otimes \mathfrak{A}_2$ - meßbare Funktion, so sind alle partiellen Funktionen*)

$$f(\cdot, s_2) \quad \mathfrak{A}_1\text{- meßbar} \quad \forall s_2 \in S_2,$$
$$f(s_1, \cdot) \quad \mathfrak{A}_2\text{- meßbar} \quad \forall s_1 \in S_1.$$

Beweis: a) Zu $s_1 \in S_1$ betrachte $\varphi_{s_1} : s_2 \mapsto (s_1, s_2)$. φ_{s_1} ist meßbar nach 6.2, und $E_{s_1} = \varphi_{s_1}^{-1}(E)$.
b) gilt wegen $f(s_1, \cdot) = f \circ \varphi_{s_1}$.
[Die übrigen Aussagen sind analog zu behandeln.]

Nehmen wir nun an, es seien Maße μ_i auf $\mathfrak{A}_i$ definiert. Ziel ist, jetzt ein Maß π auf $\mathfrak{A}_1 \otimes \mathfrak{A}_2$ mit

$$\pi(A_1 \times A_2) = \mu_1(A_1) \cdot \mu_2(A_2) \qquad \forall A_i \in \mathfrak{A}_i$$

zu definieren. Wir beschränken uns hier auf den Fall σ-endlicher μ_i.
(Einige Bemerkungen über den allgemeinen Fall folgen auf S. 124.)

*) $f(\cdot, s_2) : s_1 \mapsto f(s_1, s_2)$

6.5 Lemma $(S_1, \mathfrak{A}_1, \mu_1)$ und $(S_2, \mathfrak{A}_2, \mu_2)$ seien σ-endlich.
Für $E \in \mathfrak{A}_1 \otimes \mathfrak{A}_2$ sind dann die Funktionen

$$s_1 \mapsto \mu_2(E_{s_1}) \qquad \mathfrak{A}_1\text{-meßbar}$$

bzw. $$s_2 \mapsto \mu_1(E^{s_2}) \qquad \mathfrak{A}_2\text{-meßbar}.$$

Beweis: Die Funktionen sind wohldefiniert nach 6.4.
Aus Symmetriegründen reicht es, die erste der beiden Funktionen zu behandeln.
Zuerst sei angenommen, daß μ_2 sogar endlich ist. Wir werden zeigen, daß

$$\mathcal{D} := \{E \in \mathfrak{A}_1 \otimes \mathfrak{A}_2 : \; s_1 \mapsto \mu_2(E_{s_1}) \text{ meßbar}\}$$

ein Dynkinsystem ist (Def I.3.9), das alle meßbaren Rechtecke umfaßt. Da diese einen $\cap$-stabilen Erzeuger von $\mathfrak{A}_1 \otimes \mathfrak{A}_2$ bilden (denn

$$A_1 \times A_2 \cap B_1 \times B_2 = (A_1 \cap B_1) \times (A_2 \cap B_2) \quad)$$

zeigt Satz I.3.10 dann die Behauptung im endlichen Fall.

⌐ Schreibe abkürzend $\mu_2^E(s_1) = \mu_2(E_{s_1})$.

- $\mathcal{D}$ ist Dynkinsystem:

(i) $\emptyset, S_1 \times S_2 \in \mathcal{D}$: wegen $\mu_2^{\emptyset} = 0$, $\mu_2^{S_1 \times S_2} = \mu_2(S_2)$

(ii) $E \in \mathcal{D} \Rightarrow E^c \in \mathcal{D}$: $\mu_2^{E^c} = \mu_2(S_2) - \mu_2^E$
(hier braucht man die Endlichkeit von μ_2!)

(iii) $E_1, E_2, \dots \in \mathcal{D}$, die E_i paarweise disjunkt
$\Rightarrow E := \bigcup_{i=1}^{\infty} E_i \in \mathcal{D}$:

$$\mu_2^E = \sum_{i=1}^{\infty} \mu_2^{E_i} \quad \text{(wegen der Disjunktheit der } E_i)$$

ist meßbar als punktweiser Limes meßbarer Funktionen.

- $A_1 \times A_2 \in \mathcal{D}$ für $A_i \in \mathfrak{A}_i$:
da $\mu_2^{A_1 \times A_2} = \mu_2(A_2) \cdot 1_{A_1}$! ┘

Nun sei μ_2 σ-endlich. Wähle also $B_1 \subset B_2 \subset \dots \in \mathfrak{A}_2$ mit $\bigcup_n B_n = S_2$ und $\mu_2(B_n) < \infty$. Für die endlichen Maße

$$\nu_n(A_2) = \mu_2(A_2 \cap B_n)$$

wissen wir, daß $s_1 \mapsto \nu_n(E_{s_1})$ stets meßbar ist. Wegen

$$\mu_2(E_{s_1}) = \sup_n \nu_n(E_{s_1})$$

folgt die Behauptung nun aus 1.4.

6.6 Satz Seien $(S_1, \mathfrak{A}_1, \mu_1)$ und $(S_2, \mathfrak{A}_2, \mu_2)$ σ-endliche Maßräume.
Dann existiert genau ein Maß $\mu_1 \otimes \mu_2$ auf $\mathfrak{A}_1 \otimes \mathfrak{A}_2$ mit

$$(*) \qquad (\mu_1 \otimes \mu_2)(A_1 \times A_2) = \mu_1(A_1) \cdot \mu_2(A_2) \qquad \forall A_i \in \mathfrak{A}_i.$$

Es hat die Eigenschaft

$$\mu_1 \otimes \mu_2(E) = \int_{S_1} \mu_2(E_{s_1}) \, d\mu_1(s_1) = \int_{S_2} \mu_1(E^{s_2}) \, d\mu_2(s_2)$$

für alle $E \in \mathfrak{A}_1 \otimes \mathfrak{A}_2$.

Das in Satz 6.6 beschriebene Maß heißt Produktmaß von μ_1 und μ_2. Es ist ebenfalls σ-endlich nach $(*)$.

Beweis: Die Eindeutigkeit ist eine unmittelbare Konsequenz von I.3.11, denn jedes Maß, das $(*)$ erfüllt, ist σ-endlich auf dem $\cap$-stabilen Erzeuger der meßbaren Rechtecke.

Zur Existenz: Definiere (die Integranden sind meßbar nach 6.5!)

$$\tau_1(E) = \int_{S_1} \mu_2(E_{s_1}) \, d\mu_1(s_1),$$

$$\tau_2(E) = \int_{S_2} \mu_1(E^{s_2}) \, d\mu_2(s_2) \qquad (E \in \mathfrak{A}_1 \otimes \mathfrak{A}_2).$$

Dann ist $\tau_1(\emptyset) = 0$, und für eine disjunkte Vereinigung $\bigcup_{i=1}^{\infty} E_i$ ist

$$\tau_1\Big(\bigcup_{i=1}^{\infty} E_i\Big) = \int_{S_1} \mu_2\Big(\big(\bigcup_{i=1}^{\infty} E_i\big)_{s_1}\Big) \, d\mu_1(s_1)$$

$$= \int_{S_1} \mu_2\Big(\bigcup_{i=1}^{\infty} (E_i)_{s_1}\Big) \, d\mu_1(s_1)$$

$$= \int_{S_1} \sum_{i=1}^{\infty} \mu_2\big((E_i)_{s_1}\big) \, d\mu_1(s_1) \qquad \text{(disj. Vereinigung)}$$

$$= \sum_{i=1}^{\infty} \int_{S_1} \mu_2\big((E_i)_{s_1}\big) \, d\mu_1(s_1) \qquad \text{(Beppo Levi, 2}\ldots)$$

$$= \sum_{i=1}^{\infty} \tau_1(E_i).$$

Damit ist τ_1 ein Maß, das (*) erfüllt:

$$\tau_1(A_1 \times A_2) = \int_{S_1} 1_{A_1}(s_1) \cdot \mu_2(A_2) \, d\mu_1(s_1)$$

$$= \mu_1(A_1) \cdot \mu_2(A_2).$$

Dasselbe gilt für τ_2, und die bereits gezeigte Eindeutigkeit liefert τ_1 ...

6.7 Beispiele a) $\lambda^k \otimes \lambda^l = \lambda^{k+l}$ auf $\mathrm{Bor}(\mathbb{R}^k) \otimes \mathrm{Bor}(\mathbb{R}^l) = \mathrm{Bor}(\mathbb{R}^{k+l})$.
(*) zeigt insbesondere, daß $\lambda^k \otimes \lambda^l$ mit dem Jordanschen Inhalt auf Intervallen übereinstimmt. Da der Jordansche Inhalt eindeutig zu einem Maß auf den Borelmengen fortsetzbar ist (I.4.1), folgt die Behauptung.

b) Produktmaße haben eine große Bedeutung in der Wahrscheinlichkeitstheorie. Sei $(S, \mathfrak{A}, \mu)$ ein Wahrscheinlichkeitsraum, also ein Maßraum mit $\mu(S) = 1$. $f_1, f_2 : S \to \mathbb{R}$ seien meßbare Funktionen, sog. Zufallsvariable. μ_i sei die Verteilung von f_i (also das Bildmaß von μ unter f_i, Bsp. 5.2.a)). f_1 und f_2 werden unabhängig genannt, wenn ihre gemeinsame Verteilung auf $\mathbb{R}^2$, d.h. das Bildmaß von μ unter $s \mapsto (f_1(s), f_2(s))$, das Produktmaß $\mu_1 \otimes \mu_2$ ist. (Zurückübersetzt bedeutet das nämlich

$$\mu(\{f_1 \in A_1\} \cap \{f_2 \in A_2\}) = \mu(\{f_1 \in A_1\}) \cdot \mu(\{f_2 \in A_2\});$$

das ist die „elementare" Definition der Unabhängigkeit.)

c) Die in 6.6 aufgezeigte Möglichkeit, z.B. Volumina durch Integration ihrer 2-dimensionalen Schnitte zu berechnen $(\lambda^3 = \lambda^2 \otimes \lambda^1)$, heißt Cavalierisches Prinzip. Es impliziert z.B. für $E, F \in \mathrm{Bor}(\mathbb{R}^3) = \mathrm{Bor}(\mathbb{R}^2) \otimes \mathrm{Bor}(\mathbb{R})$

$$\lambda^2(E^s) = \lambda^2(F^s) \quad \forall\, s \in \mathbb{R} \quad \Rightarrow \quad \lambda^3(E) = \lambda^3(F).$$

Kommen wir nun zur Integration bzgl. des Produktmaßes. Der entsprechende Satz (Satz von Fubini) gehört zu den Eckpfeilern der Integrationstheorie.

6.8 Lemma $(S_1, \mathfrak{A}_1, \mu_1)$ und $(S_2, \mathfrak{A}_2, \mu_2)$ seien σ-endlich, $f : S_1 \times S_2 \to [0, \infty]$ sei $\mathfrak{A}_1 \otimes \mathfrak{A}_2$-meßbar.
Dann ist

$$s_1 \mapsto \int_{S_2} f(s_1, \cdot) \, d\mu_2 \qquad \mathfrak{A}_1\text{-meßbar}$$

bzw.

$$s_2 \mapsto \int_{S_1} f(\cdot, s_2) \, d\mu_1 \qquad \mathfrak{A}_2\text{-meßbar}.$$

Beweis: Die Integranden sind jedenfalls meßbar nach 6.4.

Nach 6.5 ist die Behauptung richtig, wenn f eine Indikatorfunktion ist, denn $\int_{S_2} 1_E(s_1,\cdot)\, d\mu_2 = \mu_2(E_{s_1})$.

Daher gilt die Behauptung auch für Treppenfunktionen (Integration ist linear) und nach 1.8 sowie Beppo Levi (2.5) allgemein.

6.9 Satz (Satz von Tonelli)

Unter den Voraussetzungen von 6.8 gilt

$$\int_{S_1\times S_2} f\, d\mu_1\otimes\mu_2 = \int_{S_1}\Big(\int_{S_2} f(s_1,s_2)\, d\mu_2(s_2)\Big)\, d\mu_1(s_1)$$

$$= \int_{S_2}\Big(\int_{S_1} f(s_1,s_2)\, d\mu_1(s_1)\Big)\, d\mu_2(s_2)$$

Beweis: Die Integranden sind meßbar wegen 6.8. Die Aussage stimmt für Indikatorfunktionen, da z.B.

$$\int_{S_1}\Big(\int_{S_2} 1_E(s_1,s_2)\, d\mu_2(s_2)\Big)\, d\mu_1(s_1)$$

$$= \int_{S_1} \mu_2(E_{s_1})\, d\mu_1(s_1)$$

$$\underset{6.6}{=} (\mu_1\otimes\mu_2)(E).$$

Mit der üblichen Methode (s.o.) erhält man die Behauptung allgemein.

Wir machen jetzt den Schritt von den positiven meßbaren zu den integrierbaren Funktionen. Die Formulierung des folgenden Satzes ist leider etwas schwerfällig. (Eine griffigere Formulierung folgt auf S. 121.)

6.10 Theorem (Satz von Fubini)

$(S_1,\mathcal{A}_1,\mu_1)$ und $(S_2,\mathcal{A}_2,\mu_2)$ seien σ-endlich, $f: S_1\times S_2 \to \overline{\mathbb{R}}$ sei $\mathcal{A}_1\otimes\mathcal{A}_2$-meßbar und $\mu_1\otimes\mu_2$-integrierbar.

a) Für μ_1-fast alle s_1 ist $f(s_1,\cdot)$ μ_2-integrierbar,
für μ_2-fast alle s_2 ist $f(\cdot,s_2)$ μ_1-integrierbar.

b) Sei

$$I_f(s_1) = \begin{cases} \int_{S_2} f(s_1,\cdot)\, d\mu_2 & \text{falls } f(s_1,\cdot)\ \mu_2\text{-integrierbar} \\ 0 & \text{sonst} \end{cases}$$

$$J_f(s_2) = \begin{cases} \int_{S_1} f(\cdot,s_2)\, d\mu_1 & \text{falls } f(\cdot,s_2)\ \mu_1\text{-integrierbar} \\ 0 & \text{sonst} \end{cases}$$

Dann sind I_f und J_f meßbar und μ_1- bzw. μ_2-integrierbar.

c) $$\int_{S_1\times S_2} f\, d\mu_1\otimes\mu_2 = \int_{S_1} I_f\, d\mu_1 = \int_{S_2} J_f\, d\mu_2$$

Beweis: a) Die Meßbarkeit dieser Funktionen wurde in 6.4 beobachtet. Die Behauptung ergibt sich aus 6.9 und 3.3 wegen

$$\int_{S_1}\Big(\int_{S_2} |f(s_1,\cdot)|\, d\mu_2\Big)\, d\mu_1(s_1) \underset{6.9}{=} \int_{S_1\times S_2} |f|\, d\mu_1\otimes\mu_2 < \infty$$

$$\underset{3.3\,a)}{\Rightarrow} \int_{S_2} |f(s_1,\cdot)|\, d\mu_2 < \infty \quad \mu_1\text{-f.ü.}$$

b) Da der „sonst"-Fall auf der meßbaren Menge

$$\{s_1 : \int_{S_2} |f(s_1,\cdot)|\, d\mu_2 = \infty\}$$

eintritt (beachte 6.8!), folgt die Meßbarkeit von I_f durch

Zerlegung $f = f^+ - f^-$ aus der entsprechenden Aussage über positive Funktionen (6.8). 6.9 zeigt die Integrierbarkeit:

$$\int_{S_1} |I_f| \, d\mu_1 \le \int_{S_1}\Big(\int_{S_2} |f(s_1,\cdot)| \, d\mu_2\Big) \, d\mu_1(s_1)$$

$$= \int_{S_1\times S_2} |f| \, d(\mu_1\otimes\mu_2) < \infty.$$

(J_f geht natürlich genauso.)

c) stimmt für f^+ und f^- nach 6.9, also auch für f, da z.B. $I_{f^+} - I_{f^-} = I_{f^+ - f^-}$.

Um den Satz von Fubini anwenden zu können, muß man sich zuerst von der Integrierbarkeit von f überzeugen; nach dem Satz von Tonelli muß man dazu „nur" eines der iterierten Integrale von $|f|$ berechnen! Als Konsequenz erhält man die Übereinstimmung der iterierten Integrale von f:

$$(*) \quad \int_{S_1}\Big(\int_{S_2} f(s_1,s_2) \, d\mu_2(s_2)\Big) \, d\mu_1(s_1) = \int_{S_2}\Big(\int_{S_1} f(s_1,s_2) \, d\mu_1(s_1)\Big) \, d\mu_2(s_2)$$

wobei zu bemerken ist, daß in dieser Formulierung die Integranden evtl. nur fast überall definiert sind.

Stellen wir also die für Anwendungen gebräuchlichste Variante des Satzes von Fubini / Tonelli auf:

Satz von Fubini / Tonelli:

Falls
- $(S_1, \mathfrak{A}_1, \mu_1)$ und $(S_2, \mathfrak{A}_2, \mu_2)$ σ-endlich,
- $f : S_1 \times S_2 \to \overline{\mathbb{R}}$ $\mathfrak{A}_1\otimes\mathfrak{A}_2$ - meßbar,
- $\int_{S_1}\Big(\int_{S_2} |f| \, d\mu_2\Big) \, d\mu_1 < \infty$ oder $\int_{S_2}\Big(\int_{S_1} |f| \, d\mu_1\Big) \, d\mu_2 < \infty$,

so gilt $(*)$ von S. 120.

Als Anwendung des Satzes können wir bequem

$$\lim_{R\to\infty} \int_0^R \frac{\sin x}{x} \, dx$$

berechnen.[*] Wegen $\frac{1}{x} = \int_0^\infty e^{-ux} \, du$ ist nämlich

$$\int_0^R \frac{\sin x}{x} \, dx = \int_0^R \int_0^\infty \sin x \, e^{-ux} \, du \, dx.$$

Um die Integrationsreihenfolge zu vertauschen, checken wir die obigen Voraussetzungen:

- $[0,R]$ und $[0,\infty[$, jeweils mit Borelmengen und Lebesguemaß versehen, sind σ-endlich.
- $f(x,u) = \sin x \cdot e^{-ux}$ ist stetig, also meßbar (beachte Beispiel 6.3 a)).

[*] In Beispiel 2.11 a) war gezeigt, daß für stetige Integranden auf kompakten Intervallen R- und L-Integral übereinstimmen. Falls z.B. R-$\int_0^\infty |f(x)| \, dx$ existiert, gilt die Übereinstimmung auch für uneigentliche R- und L-Integrale (verwende den Lebesgueschen Konvergenzsatz). (Beachte jedoch $\int_0^\infty \left|\frac{\sin x}{x}\right| dx = \infty$.) Alle folgenden Integrale können daher als R-Integrale berechnet werden.

• $\int_0^R \int_0^\infty |f(x,u)|\,du\,dx = \int_0^R |\sin x| \cdot \frac{1}{x}\,dx \underset{|\sin x| \le x}{\le} \int_0^R dx = R < \infty$

Also
$$\int_0^R \frac{\sin x}{x}\,dx = \int_0^\infty du \int_0^R e^{-ux} \sin x\,dx$$
$$= \int_0^\infty du \left[\frac{1}{1+u^2}\left(1 - e^{-uR}(u \sin R + \cos R)\right)\right]$$
(durch zweimalige partielle Integration)
$$= \int_0^\infty \frac{du}{1+u^2} - \int_0^\infty \frac{u \sin R + \cos R}{1+u^2}\, e^{-uR}\,du$$

Das erste Integral ist $= \frac{\pi}{2}$, das zweite kann betragsmäßig nach oben abgeschätzt werden gegen
$$(\dots) \le \int_0^\infty (u+1)\, e^{-uR}\,du \underset{u+1 \le e^u}{\le} \int_0^\infty e^{-u(R-1)}\,du = \frac{1}{R-1} \xrightarrow[R\to\infty]{} 0.$$

Zusammen:
$$\lim_{R\to\infty} \int_0^R \frac{\sin x}{x}\,dx = \frac{\pi}{2}.$$

Kannst du diesen Beweis auch mit deinen Kenntnissen aus Analysis II über Integralvertauschung bei iterierten uneigentlichen R-Integralen führen?

Abschließend wird die Rolle der 3 Voraussetzungen bei Fubini/Tonelli beleuchtet. Beachte, daß die Formulierung auf S. 121 keinen Bezug auf das Produktmaß nimmt, dessen Existenz wir ja nur im σ-endlichen Fall bewiesen hatten*). (Der Beweis freilich sehr wohl!) Sobald für eine Funktion f die iterierten Integrale existieren, drängt sich das Vertauschungsproblem (gilt (*) von S. 120?) auf.

*) vgl. hierzu S. 124.

Wir werden sehen, daß alle drei Voraussetzungen wesentlich sind:

a) Aus der Analysis sind Beispiele bekannt, wo die iterierten Integrale existieren, aber verschieden sind, etwa für $f(x,y) = \frac{x^2-y^2}{(x^2+y^2)^2}$, $(x,y) \in]0,1]^2$, gilt
$$\int_0^1 \left(\int_0^1 f(x,y)\,dy\right) dx = \frac{\pi}{4},$$
$$\int_0^1 \left(\int_0^1 f(x,y)\,dx\right) dy = -\frac{\pi}{4}.$$
(Hier ist $f \notin L^1(]0,1]^2)$.)

b) Sei $S_1 = S_2 = [0,1]$, $\mathfrak{A}_1 = \mathfrak{A}_2 = \mathrm{Bor}([0,1])$, $\mu_1 = \mu :=$ zählendes Maß, $\mu_2 = \lambda :=$ Lebesguemaß. Sei
$$\Delta = \{(x,x) : x \in [0,1]\}$$
die „Diagonale". Da Δ abgeschlossen ist, ist $\Delta \in \mathrm{Bor}([0,1]^2) = \mathfrak{A}_1 \otimes \mathfrak{A}_2$. Also ist $f = 1_\Delta$ meßbar. Beide iterierten Integrale existieren und sind verschieden:
$$\int_0^1 \left(\int_0^1 1_\Delta(x,y)\,d\lambda(y)\right) d\mu(x) = \int_0^1 0 \cdot d\mu = 0$$
$$\int_0^1 \left(\int_0^1 1_\Delta(x,y)\,d\mu(x)\right) d\lambda(y) = \int_0^1 1 \cdot d\lambda = 1 \quad !$$

(Hier ist μ_1 nicht σ-endlich.)

c) Das folgende von Sierpinski stammende Beispiel zeigt, daß man auf die ⊗-Meßbarkeit nicht verzichten darf, selbst bei endlichen Maßen und endlichem f.
Wähle $(S_i, \mathfrak{A}_i, \mu_i) = ([0,1], \mathrm{Bor}[0,1], \lambda)$. Die Konstruktion des Gegenbeispiels beruht auf der Kontinuumshypothese. Als

Konsequenz davon gibt es eine bijektive Abbildung φ von $[0,1]$ auf eine wohlgeordnete Menge $(W, \prec)$ ($\prec$ bezeichne die Ordnung in W), so daß $\varphi(x)$ stets nur abzählbar viele Vorgänger in W hat. (Es ist nicht unbedingt leicht, sich so etwas vorzustellen!) Setze

$$Q = \{ (x,y) \in [0,1]^2 : \varphi(x) \prec \varphi(y) \}$$

und

$$f = 1_Q .$$

Für jedes $x \in [0,1]$ ist Q_x co-abzählbar, folglich eine Borelmenge mit $\lambda(Q_x) = 1$. Umgekehrt ist für jedes $y \in [0,1]$ der Schnitt abzählbar, folglich eine Borelmenge mit $\lambda(Q^y) = 0$. Damit gilt

$$\int_0^1 \left(\int_0^1 f(x,y)\, dx \right) dy = \int_0^1 0 \, dy = 0 ,$$

$$\int_0^1 \left(\int_0^1 f(x,y)\, dy \right) dx = \int_0^1 1 \, dx = 1 .$$

Streifen wir nun noch kurz die Produktmaßfrage im nicht σ-endlichen Fall. In der Tat läßt sich für zwei beliebige Maße stets ein Produktmaß konstruieren. (Idee: Betrachte den Ring der disjunkten endlichen Vereinigungen von meßbaren Rechtecken und definiere darauf $\tau\left(\bigcup_{i=1}^{n} (A_i \times B_i)\right) = \sum_{i=1}^{n} \tau(A_i) \dots$ Zeige, daß es sich um ein wohldefiniertes Prämaß handelt, und betrachte das zugehörige äußere Maß auf $\mathfrak{A}_1 \otimes \mathfrak{A}_2$.) Bezeichnen wir dieses à la Carathéodory konstruierte Maß mit $\mu_1 \otimes \mu_2$ (in der Regel wird es noch weitere Maße ν auf $\mathfrak{A}_1 \otimes \mathfrak{A}_2$ mit $\nu(A_1 \times A_2) = \mu_1(A_1) \cdot \mu_2(A_2)$ geben, wenn keine σ-Endlichkeit vorliegt), so kann man für dieses Produktmaß den Satz von Fubini in der Textversion 6.10 zeigen. Die

Brauchbarkeit des Produktmaßes und des Satzes von Fubini wird jedoch erheblich dadurch eingeschränkt, daß der Satz von Tonelli i.a. nicht mehr gilt! Einzelheiten entnehme man z.B. den Büchern von Mukherjea / Pothoven oder Royden. Vgl. auch Behrends, S. 94 ff., der das in Beispiel 6) (S. 123) entstehende Produktmaß analysiert.

Eine wesentlich nützlichere Verallgemeinerung ist die auf Produkte endlich vieler σ-endlicher Maßräume. Die wesentlichen Resultate (Existenz des Produktmaßes, Satz von Fubini / Tonelli) folgen problemlos induktiv aus dem Fall zweier Faktoren. Aus der Eindeutigkeit des Produktmaßes folgt noch, daß die Produktbildung hier assoziativ ist. Eine explizite Darstellung findet man beispielsweise im Buch von Bauer.

Unendliche Produkte von Wahrscheinlichkeitsräumen werden in Kapitel V behandelt.

III. Differentiation von Maßen

III.1 Der Satz von Lebesgue - Radon - Nikodym

Wir betrachten einen meßbaren Raum $(S, \mathfrak{A})$; auf diesem meßbaren Raum werden wir nun mehrere Maße in Beziehung zueinander setzen. Wir benötigen zwei neue Begriffe:

1.1 Definition ν und μ seien Maße auf $\mathfrak{A}$.

a) ν heißt absolutstetig bezüglich μ (in Zeichen: $\nu \ll \mu$), falls

$$A \in \mathfrak{A}, \quad \mu(A) = 0 \quad \Rightarrow \quad \nu(A) = 0$$

b) ν und μ heißen singulär (in Zeichen: $\nu \perp \mu$), falls

$$\exists A \in \mathfrak{A} \quad \text{mit} \quad \mu(A) = 0 \quad \text{und} \quad \nu(A^c) = 0.$$

Offenbar sind Absolutstetigkeit und Singularität stark entgegengesetzte Eigenschaften, die einander ausschließen:

$$\nu \ll \mu \;\; \& \;\; \nu \perp \mu \quad \Rightarrow \quad \nu = 0$$

1.2 Beispiele a) $f: S \to [0, \infty]$ sei meßbar und ν das Maß mit der „Dichte f" (Bsp. II.2.11 b))

$$\nu(A) = \int_A f \, d\mu.$$

Dann gilt $\nu \ll \mu$. Im fundamentalen Satz von Radon-Nikodym (1.4. b)) wird die Umkehrung dieses Sachverhalts untersucht.

b) Für einen Punkt $t \in \mathbb{R}$ ist δ_t singulär zum Lebesguemaß.

c) Ein weiteres Beispiel für ein λ-singuläres Maß liefert das Lebesgue-Stieltjes Maß μ_F aus I.4.10 f): Es ist für die Cantormenge $\lambda(C) = 0$ und $\mu_F(\mathbb{R} \setminus C) = 0$.

d) Auf $Bor(\mathbb{R}^2)$ betrachte $\mu(A) = \lambda^1(A \cap \mathbb{R} \times \{0\})$. Offenbar ist $\mu \perp \lambda^2$.

Absolutstetige Maße haben folgende Stetigkeitseigenschaft (die allerdings nicht der Anlaß zu dieser Terminologie war; dazu vgl. IV.2):

1.3 Lemma ν sei ein endliches, μ ein beliebiges Maß auf $(S, \mathfrak{A})$. Dann sind äquivalent:

(i) $\nu \ll \mu$

(ii) $\forall \varepsilon > 0 \;\; \exists \delta > 0 \;\; \forall A \in \mathfrak{A}: \quad \mu(A) < \delta \Rightarrow \nu(A) < \varepsilon.$

Beweis: (ii) $\Rightarrow$ (i) klar

(i) $\Rightarrow$ (ii) Falls nicht, existieren $\varepsilon_0 > 0$ und eine Folge (A_n) in $\mathfrak{A}$ mit

$$\mu(A_n) < 2^{-n}, \quad \text{aber} \quad \nu(A_n) \geq \varepsilon_0$$

für alle $n \in \mathbb{N}$. Sei $A = \bigcap_{n=1}^{\infty} \bigcup_{m \geq n} A_m \in \mathfrak{A}$. Dann gilt:

$\mu(A) = 0$ [vgl. S. 90, Borel-Cantelli-Lemma]

$\nu(A) \geq \varepsilon_0$

[denn $\nu(A) \underset{\substack{\uparrow \\ \text{I.2.4} \\ \nu \text{ ist endlich!}}}{=} \inf_n \nu\left(\bigcup_{m \geq n} A_m\right) \geq \inf_n \nu(A_n) \geq \varepsilon_0$]

Der folgende Satz ist ein weiterer Eckpfeiler der Maßtheorie.

1.4 Theorem ν und μ seien σ-endliche Maße auf $(S,\mathfrak{A})$.

a) (Lebesguescher Zerlegungssatz)

Es existieren eindeutig bestimmte Maße ν_a und ν_s mit

$$\nu = \nu_a + \nu_s \,, \qquad \nu_a \ll \mu \quad \& \quad \nu_s \perp \mu .$$

b) (Satz von Radon - Nikodym)

Falls ν absolutstetig bezüglich μ ist, besitzt ν eine μ-f.ü. eindeutig bestimmte Dichte f. M.a.W., $\nu \ll \mu$ impliziert

$$\nu(A) = \int_A f \, d\mu \qquad \text{für alle } A \in \mathfrak{A}$$

für eine meßbare Funktion $f: S \to [0,\infty[$.

Beweis: Wir verfahren in 4 Schritten:

1. Beweis der Existenzaussage in b) für endliche ν und μ, die sogar $\nu(A) \leq \mu(A) \quad \forall A \in \mathfrak{A}$ erfüllen.
2. Simultaner Beweis der Existenzaussagen in a) und b) für endliche ν und μ.
3. Ausdehnung auf den σ-endlichen Fall.
4. Eindeutigkeitsbeweis.

Zu 1. Das ist der Kern des Beweises, gelte also zusätzlich

$$\nu(A) \leq \mu(A) \qquad \forall A \in \mathfrak{A}$$

sowie

$$\mu(S) < \infty .$$

Sei $Z = \{A_1, \dots, A_n\}$ eine Zerlegung von S in paarweise disjunkte meßbare Mengen. Definiere eine meßbare Funktion h_Z durch

$$h_Z(s) = \begin{cases} \dfrac{\nu(A_i)}{\mu(A_i)} & \text{für } s \in A_i \text{, falls } \mu(A_i) \neq 0 \\ 0 & \text{sonst.} \end{cases}$$

Offensichtlich gilt für $A \in \sigma(Z)$ (d.h. A ist Vereinigung gewisser A_i)

$$(*) \qquad \nu(A) = \int_A h_Z \, d\mu ,$$

und es ist $\quad 0 \leq h_Z \leq 1$.

Ist nun Z' eine Verfeinerung von Z (d.h. $Z \subset \sigma(Z')$), so gilt

$$(**) \qquad \int_S h_{Z'}^2 \, d\mu = \int_S h_Z^2 \, d\mu + \int_S (h_{Z'} - h_Z)^2 \, d\mu \geq \int_S h_Z^2 \, d\mu .$$

Wegen $(*)$ (angewandt auf Z') ist

$$\nu(A) = \int_A h_{Z'} \, d\mu = \int_A h_Z \, d\mu$$

für alle $A \in Z$. h_Z ist auf A konstant (etwa $= a$),

also $\displaystyle\int_A h_Z h_{Z'} \, d\mu = a \int_A h_{Z'} \, d\mu = a \int_A h_Z \, d\mu = \int_A h_Z^2 \, d\mu$,

also $\displaystyle\int_S h_Z h_{Z'} \, d\mu = \int_S h_Z^2 \, d\mu$. Das liefert sofort $(**)$.

Nun sei

$$c := \sup \int_S h_Z^2 \, d\mu ,$$

wo das Supremum über alle endlichen meßbaren Zerlegungen zu erstrecken ist. Offensichtlich ist wegen $0 \leq h_Z \leq 1$

$$0 \leq c \leq \mu(S).$$

Wähle Zerlegungen Z_n mit

$$\int_S h_{Z_n}^2 \, d\mu \geq c - \left(\tfrac{1}{4}\right)^n .$$

O.E. dürfen wir annehmen, daß Z_{n+1} feiner als Z_n ist (sonst gehe zur gemeinsamen Verfeinerung $\{A \cap B : A \in Z_n, B \in Z_{n+1}\}$ über und beachte (∗∗)). (∗∗) zeigt dann

$$\int_S (h_{Z_{n+1}} - h_{Z_n})^2 \, d\mu = \int_S h^2_{Z_{n+1}} \, d\mu - \int_S h^2_{Z_n} \, d\mu$$
$$\leq c - (c - (\tfrac{1}{4})^n)$$
$$= (\tfrac{1}{4})^n ,$$

d.h.
$$\| h_{Z_{n+1}} - h_{Z_n} \|_{L^2(\mu)} \leq (\tfrac{1}{2})^n .$$

Der Teleskopsummentrick liefert, daß (h_{Z_n}) eine $\|\cdot\|_{L^2(\mu)}$-Cauchyfolge ist, und wegen der Vollständigkeit von $L^2(\mu)$ (II.4.4.) existiert eine meßbare Funktion h mit

$$\| h_{Z_n} - h \|_{L^2(\mu)} \to 0 .$$

Zeigen wir also
$$\nu(A) = \int_A h \, d\mu \qquad \text{für } A \in \mathfrak{A}.$$

Nach der Hölderschen Ungleichung (II.4.2) ist stets

$$\left| \int_A h \, d\mu - \int_A h_{Z_n} \, d\mu \right|$$
$$\leq \int_S 1_A \cdot |h - h_{Z_n}| \, d\mu$$
$$\leq \| 1_A \|_{L^2} \cdot \| h - h_{Z_n} \|_{L^2}$$
$$\leq \mu(S)^{1/2} \cdot \| h - h_{Z_n} \|_{L^2}$$
$$\xrightarrow[n \to \infty]{} 0 ,$$

d.h.
$$\int_A h \, d\mu = \lim_{n \to \infty} \int_A h_{Z_n} \, d\mu .$$

Ist nun $A \in Z_{n_0}$ für ein n_0, also nach unserer Annahme über (Z_n) $A \in Z_n$ für $n \geq n_0$, so ist die rechte Seite nach (∗) $= \nu(A)$ für $n \geq n_0$, und

die Behauptung ist bewiesen. Im allgemeinen Fall gehen wir bei gegebenem $A \in \mathfrak{A}$ von Z_n zu $\tilde{Z}_n$, der gemeinsamen Verfeinerung von Z_n und $\{A, A^c\}$, über. Wir wissen nach dem gerade Gezeigten $\nu(A) = \lim_{n\to\infty} \int_A h_{\tilde{Z}_n} \, d\mu$, und es bleibt nun

$$\int_A (h_{\tilde{Z}_n} - h_{Z_n}) \, d\mu \to 0$$

zu zeigen. Wieder zeigt die Höldersche Ungleichung, daß hierfür $\| h_{\tilde{Z}_n} - h_{Z_n} \|_{L^2} \to 0$ hinreichend ist, und das folgt aus (∗∗), da

$$\| h_{\tilde{Z}_n} - h_{Z_n} \|^2_{L^2} \underset{(**)}{=} \int_S h^2_{\tilde{Z}_n} \, d\mu - \int_S h^2_{Z_n} \, d\mu$$
$$\leq c - (c - (\tfrac{1}{4})^n)$$
$$= (\tfrac{1}{4})^n \to 0 .$$

Wir bemerken noch, daß für die oben konstruierte Funktion

$$0 \leq h \leq 1 \qquad \mu\text{-f.ü.}$$

gilt. (Eine Teilfolge der (h_{Z_n}) konvergiert gegen h f.ü. [II.3.9 b)&c)], und die h_{Z_n} sind $[0,1]$-wertig.) Nach Abänderung von h auf einer Nullmenge dürfen wir sogar o.E.

$$0 \leq h(s) \leq 1 \qquad \forall s \in S$$

annehmen.

Zu 2. Wir setzen nun ν und μ als endlich voraus und zeigen a) und b) auf einmal. Wir wenden den 1. Schritt mit ν und $\nu+\mu$ bzw. μ und $\nu+\mu$ an. Man erhält $[0,1]$-

wertige meßbare Funktionen h_ν, h_μ mit

$$\nu(A) = \int_A h_\nu \, d(\nu+\mu),$$

$$\mu(A) = \int_A h_\mu \, d(\nu+\mu) \qquad (A \in \mathfrak{A}).$$

Sei $N = \{h_\mu = 0\}$ und

$$h(s) = \begin{cases} \dfrac{h_\nu(s)}{h_\mu(s)} & s \notin N \\ 0 & s \in N. \end{cases}$$

Dann ist $\mu(N) = 0$, also ν_s mit $\nu_s(A) := \nu(A \cap N)$ μ-singulär.
Für alle $A \in \mathfrak{A}$ gilt schließlich

$$\begin{aligned} \nu(A) &= \nu(A \cap N^c) + \nu_s(A) \\ &= \int_{A \cap N^c} h_\nu \, d(\nu+\mu) + \nu_s(A) \\ &= \int_{A \cap N^c} h\, h_\mu \, d(\nu+\mu) + \nu_s(A) \\ &= \int_{A \cap N^c} h \, d\mu + \nu_s(A) \qquad [\text{nach Def. von } h_\mu \text{ und I.2.11 b)}] \\ &= \int_A h \, d\mu + \nu_s(A). \end{aligned}$$

Zu 3. Da nun μ σ-endlich ist, existiert eine Folge (E_n) paarweise disjunkter meßbarer Mengen mit $\bigcup E_n = S$ und $\mu(E_n) < \infty$ $\forall\, n \in \mathbb{N}$. Sei (F_n) eine analoge Folge für ν. Ordnet man die abzählbare Menge der $E_n \cap F_m$, $n, m \in \mathbb{N}$, in eine Folge $A_1, A_2, \dots$, so gilt $\bigcup A_n = S$ und $\mu(A_n) < \infty$, $\nu(A_n) < \infty$ $\forall\, n \in \mathbb{N}$. Für die Maße $A \mapsto \mu(A \cap A_n)$ und $A \mapsto \nu(A \cap A_n)$ kann der 2. Schritt angewandt

werden; man erhält

$$\nu(A \cap A_n) = \int_A 1_{A_n} \cdot h_n \, d\mu + \nu_{s,n}(A).$$

Setzt man noch $h = \sum_n 1_{A_n} h_n$ und $\nu_s = \sum_n \nu_{s,n}$, dann ist

$$\nu(A) = \int_A h \, d\mu + \nu_s(A) \quad \text{und} \quad \nu_s \perp \mu.$$

Zu 4. Wegen der im 3. Schritt vorgestellten Konstruktion reicht es, die Eindeutigkeitsaussagen für den Fall endlicher μ und ν zu zeigen.
Falls $\nu(A) = \int_A h_1 \, d\mu = \int_A h_2 \, d\mu$ für alle $A \in \mathfrak{A}$ gilt, so wäre $h := h_1 - h_2$ eine integrierbare (!) Funktion mit $\int_A h \, d\mu = 0$ $\forall\, A \in \mathfrak{A}$. So eine Funktion ist nach I.33 μ-f.ü. $= 0$.
Seien nun $\nu = \nu_a + \nu_s = \tilde{\nu}_a + \tilde{\nu}_s$ zwei Lebesgue-Zerlegungen. Wir können $\nu_a(A) = \int_A h \, d\mu$ und $\nu_s(A) = \nu(A \cap N)$ (und analog für $\tilde{\nu}_a, \tilde{\nu}_s$) schreiben, wo h und $\tilde{h}$ μ-integrierbar und N und $\tilde{N}$ μ-Nullmengen sind. Es folgt für $A \subset N \cup \tilde{N}$ $\nu_a(A) = \tilde{\nu}_a(A) = 0$ und daher $\nu_s = \tilde{\nu}_s$. Dann muß auch $\nu_a = \tilde{\nu}_a$ sein, wie oben gezeigt.

Abschließend noch einige Bemerkungen.
Die im Satz von Radon-Nikodym konstruierte Funktion h wird auch Radon-Nikodym-Ableitung genannt und mit $\frac{d\nu}{d\mu}$ bezeichnet.
Diese Bezeichnung kann formal durch die Gleichheit (I.2.11 b))

$$\int_A f \, d\nu = \int_A f\, h \, d\mu = \int_A f \, \frac{d\nu}{d\mu} \, d\mu$$

(wo nun $d\mu$ „gekürzt" werden kann) suggeriert werden;

jedoch sollte man sich die Beweisstrategie des 1. Schritts im Spezialfall μ = Lebesguemaß auf $[a,b]$, $\nu(A) = \int_A f \, d\mu$ mit stetigem f verdeutlichen, um wirklich einen Differentiationsprozeß zu erkennen! Mehr dazu unter III.4 und IV.2. Dort wird gezeigt, daß der Satz von Radon-Nikodym als allgemeine Version des Hauptsatzes der Differential- und Integralrechnung angesehen werden darf.

Jenseits der σ-Endlichkeit braucht Theorem 14 nicht mehr zu gelten, z.B. hat das zählende Maß μ auf $([0,1], \mathrm{Bor}\,[0,1])$ keine Lebesguezerlegung bzgl. des Lebesguemaßes λ, andererseits gilt $\lambda \ll \mu$, ohne daß es eine Dichte gäbe. (Beweis?)

Hingegen gilt der Satz von Radon-Nikodym noch für beliebige ν, solange μ σ-endlich ist.*)

Beweisskizze: Gemäß der Standardtechnik (vgl. 3. Schritt) genügt es, endliche μ zu behandeln. Wir betrachten $\mathcal{E} = \{A \in \mathfrak{A} : \nu(\cdot \cap A) \ \sigma\text{-endlich}\}$ und $\alpha = \sup_{A \in \mathcal{E}} \mu(A)$. Dann existiert $A_0 \in \mathcal{E}$ mit $\mu(A_0) = \alpha$.
Für $B \subset A_0^c$, $B \in \mathfrak{A}$, gilt:

$$\mu(B) = 0 \Rightarrow \nu(B) = 0 \qquad [\text{denn } \nu \ll \mu]$$

$$\mu(B) > 0 \Rightarrow \nu(B) = \infty \qquad [\text{sonst Widerspruch zur Wahl von } A_0]$$

In beiden Fällen ist $\nu(B) = \int_B \infty \, d\mu \qquad [0 \cdot \infty = 0]$.

Setze also

$$h(s) = \begin{cases} \dfrac{d\,\nu(\cdot \cap A_0)}{d\mu}(s) & s \in A_0 \\ \infty & s \notin A_0 \end{cases}.$$

*) Man muß jetzt aber zulassen, daß h den Wert ∞ annimmt.

2. Signierte und komplexe Maße

Unser bisheriger Maßbegriff war darauf ausgerichtet, positive Größen (z.B. Länge, Fläche, Volumen, Masse, Wahrscheinlichkeiten etc.) zu messen. Eine elektrische Ladung kann jedoch durchaus negativ sein. Vor diesem Hintergrund ist folgende Erweiterung des Maßbegriffs natürlich:

2.1 Definition $(S, \mathfrak{A})$ sei ein meßbarer Raum. Eine σ-additive Abbildung $\mu : \mathfrak{A} \to \mathbb{R}$ (bzw. $\mathbb{C}$) heißt signiertes (bzw. komplexes) Maß.

Ausführlich bedeutet diese Forderung

$$(A_i) \text{ paarweise disjunkt} \Rightarrow \mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i)$$

Da es hier auf die Reihenfolge der Summanden nicht ankommen darf ($\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{\infty} A_{\sigma(i)}$ für jede Permutation $\sigma : \mathbb{N} \to \mathbb{N}$), muß die Reihe $\sum_{i=1}^{\infty} \mu(A_i)$ unbedingt, also absolut konvergieren.

Beachte, daß die Werte $\pm\infty$ explizit für ein signiertes Maß ausgeschlossen sind (manche Autoren gestatten, daß einer von beiden angenommen wird); daher ergibt sich hier aus $\emptyset = \emptyset \cup \emptyset \cup \dots$ automatisch die Normierung

$$\mu(\emptyset) = 0.$$

Es wird sich zeigen, daß das Studium signierter oder komplexer Maße weitgehend auf positive endliche Maße zurückgeführt werden kann.

(Hier ist μ ein positives Maß.)

<u>2.2 Beispiele</u> a) $f: S \to \mathbb{R}$ sei μ-integrierbar. Dank des Konvergenzsatzes von Lebesgue definiert $\nu(A) = \int_A f \, d\mu$ ein signiertes Maß.

b) μ_1 und μ_2 seien positive endliche Maße. Dann ist $\mu_1 - \mu_2$ ein signiertes Maß. Wir werden sehen, daß jedes signierte Maß auf diese Weise entsteht (Jordan-Zerlegung, s.u.).

c) Offenbar kann jedes komplexe Maß als $\mu_{re} + i\,\mu_{im}$ mit signierten Maßen μ_{re}, μ_{im} geschrieben werden. (Natürlich ist $\mu_{re}(A) = \mathrm{Re}\,\mu(A)$.)

Für signierte Maße gilt i.a. nicht $\mu(A) = 0$, $B \subset A \underset{(*)}{\Rightarrow} \mu(B) = 0$. (Beispiel 2) Mengen mit $\mu(A) = 0$ sind also für signierte Maße nicht unbedingt vernachlässigbar. Das sind sie erst, wenn wirklich $(*)$ gilt, wenn also im Sinn der folgenden Definition $|\mu|(A) = 0$ ist.

<u>2.3 Definition</u> μ sei ein signiertes oder komplexes Maß. Für $A \in \mathfrak{A}$ setze

$$|\mu|(A) = \sup \sum_{i=1}^{\infty} |\mu(E_i)| \quad ,$$

wo das Supremum über alle disjunkten Zerlegungen $A = \bigcup_{i=1}^{\infty} E_i$ (mit $E_i \in \mathfrak{A}$) zu erstrecken ist. $|\mu|$ heißt Variation von μ.

<u>2.4 Satz</u> Für ein signiertes oder komplexes Maß μ definiert $|\mu|$ ein endliches positives Maß.

Der Wertebereich von μ ist also eine beschränkte Teilmenge von $\mathbb{R}$ oder $\mathbb{C}$! Man sagt auch, μ sei von beschränkter Variation.

Beweis: Offensichtlich ist $|\mu|(\emptyset) = 0$. Wir zeigen nun die σ-Additivität. Seien die $A_i \in \mathfrak{A}$ paarweise disjunkt, und sei $A := \bigcup_{i=1}^{\infty} A_i$. Sei $(E_j)_{j=1}^{\infty}$ eine Zerlegung von A wie in Def. 2.3., dann ist $(A_i \cap E_j)_{j=1}^{\infty}$ eine solche für A_i und $(A_i \cap E_j)_{i=1}^{\infty}$ eine solche für E_j. Es folgt

$$\begin{aligned} \sum_j |\mu(E_j)| &= \sum_j \Big| \sum_i \mu(A_i \cap E_j) \Big| \\ &\le \sum_j \sum_i |\mu(A_i \cap E_j)| \\ &= \sum_i \sum_j |\mu(A_i \cap E_j)| \\ &\le \sum_i |\mu|(A_i). \end{aligned}$$

Übergang zum Supremum zeigt

$$|\mu|(A) \le \sum_i |\mu|(A_i).$$

Zum Beweis der umgekehrten Ungleichung sei $t_i < |\mu|(A_i)$. Dann existiert eine Zerlegung $(A_{ij})_{j=1}^{\infty}$ von A_i wie in 2.3 mit

$$t_i \le \sum_{j=1}^{\infty} |\mu(A_{ij})|.$$

$(A_{ij})_{i,j=1}^{\infty}$ ist eine Zerlegung von A mit

$$\sum_i t_i \le \sum_{i,j} |\mu(A_{ij})| \le |\mu|(A).$$

Es folgt

$$\sum_i |\mu|(A_i) \le |\mu|(A).$$

Nehmen wir nun an, es sei μ ein signiertes Maß mit $|\mu|(S) = \infty$. Wir werden paarweise disjunkte $A_1, A_2, \dots$ mit $|\mu(A_i)| \ge 1 \quad \forall i \in \mathbb{N}$ konstruieren, so daß $\sum_i \mu(A_i)$ nicht konvergiert. Dazu benötigen wir die

<u>Behauptung</u>: Ist $E \in \mathfrak{A}$ mit $|\mu|(E) = \infty$, so existieren $E', E'' \in \mathfrak{A}$ mit $E = E' \cup E''$, $E' \cap E'' = \emptyset$, $|\mu(E')| \ge 1$, $|\mu(E'')| \ge 1$.

Sei $t = 2 + 2\,|\mu(E)|$. Es existiert eine disjunkte Zerlegung $(E_i)_{i=1}^{\infty}$ von E mit

$$\sum_{i=1}^{\infty} |\mu(E_i)| \geq t .$$

Ist $P = \{i : \mu(E_i) \geq 0\}$ und $N = \{i : \mu(E_i) < 0\}$, so folgt mit $E' = \bigcup_{i \in P} E_i$, $E'' = \bigcup_{i \in N} E_i$

$$E' \cap E'' = \emptyset, \quad E' \cup E'' = E \quad \text{und}$$

$$|\mu(E')| + |\mu(E'')| = \Big|\sum_P \mu(E_i)\Big| + \Big|\sum_N \mu(E_i)\Big| = \sum_P |\mu(E_i)| + \sum_N |\mu(E_i)| \geq t,$$

daher o.B.d.A. $|\mu(E')| \geq \frac{t}{2} \geq 1$. Dann ist auch

$$|\mu(E'')| = |\mu(E) - \mu(E')| \geq |\mu(E')| - |\mu(E)| \geq \frac{t}{2} - |\mu(E)| = 1.$$

Da $|\mu|(E') + |\mu|(E'') = |\mu|(E) = \infty$ ist, ist einer der beiden Summanden $= \infty$, und die Behauptung kann erneut auf diesen angewandt werden etc.

Beginnt man mit $E = S$, erhält man so die gewünschte Folge (A_i).

Die Ungleichung $|\mu|(S) \leq |\mathrm{Re}\,\mu|(S) + |\mathrm{Im}\,\mu|(S)$, die leicht aus der für komplexe Zahlen gültigen Ungleichung $|z| \leq |\mathrm{Re}\,z| + |\mathrm{Im}\,z|$ folgt, zeigt schließlich, daß auch jedes komplexe Maß endliche Variation besitzt.

2.5 Satz (Hahn-Jordan-Zerlegung)

Zu jedem signierten Maß μ existiert genau ein Paar positiver endlicher Maße μ^+ und μ^- mit

$$\mu = \mu^+ - \mu^- \quad \text{und} \quad \mu^+ \perp \mu^- .$$

Es gilt dann $|\mu| = \mu^+ + \mu^-$.

Beweis: Eindeutigkeit: Gelte $\mu = \mu_1 - \mu_2 = \tilde\mu_1 - \tilde\mu_2$ mit $\mu_1 \perp \mu_2$, $\tilde\mu_1 \perp \tilde\mu_2$.

Seien N und $\tilde N$ meßbare Mengen mit

$$\mu_1(N) = 0 = \mu_2(N^c), \quad \tilde\mu_1(\tilde N) = 0 = \tilde\mu_2(\tilde N^c).$$

Für $A \in \mathfrak{A}$ gilt dann

$$\begin{aligned}
\mu_1(A) &\geq \mu_1(A \cap \tilde N^c) = \tilde\mu_1(A \cap \tilde N^c) - \tilde\mu_2(A \cap \tilde N^c) + \mu_2(A \cap \tilde N^c) \\
&= \tilde\mu_1(A \cap \tilde N^c) + \mu_2(A \cap \tilde N^c) \\
&\geq \tilde\mu_1(A \cap \tilde N^c) \\
&= \tilde\mu_1(A).
\end{aligned}$$

Aus Symmetriegründen ergibt sich $\tilde\mu_1(A) \geq \mu_1(A)$ und analog $\mu_2 = \tilde\mu_2$.

Existenz: Setze $\mu^+ = \frac{1}{2}(|\mu| + \mu)$, $\mu^- = \frac{1}{2}(|\mu| - \mu)$.

Das sind positive Maße, da stets $|\mu(A)| \leq |\mu|(A)$ gilt, und nach Konstruktion ist $\mu = \mu^+ - \mu^-$, $|\mu| = \mu^+ + \mu^-$.

Da $\mu^+ \leq |\mu|$ ist, existiert nach dem Satz von Radon-Nikodym (1.4) eine meßbare Funktion h, so daß

$$\mu^+(A) = \int_A h \, d|\mu| \qquad \forall\, A \in \mathfrak{A}.$$

Im 1. Beweisschritt von 1.4 wurde noch angemerkt, daß man h mit

$$0 \leq h(s) \leq 1 \qquad \forall\, s \in S$$

wählen kann. Aus $\mu = 2\mu^+ - |\mu|$ folgen dann mit $g := 2h - 1$

$$\mu(A) = \int_A g \, d|\mu| \qquad \forall A \in \mathfrak{A}$$

und
$$|g(s)| \le 1 \qquad \forall s \in S.$$

Zeigen wir nun $|g| = 1$ $\quad |\mu|$-f.ü.:

Sei $A_n = \{ |g| \le 1 - \frac{1}{n} \}$. Für $B \subset A_n$ gilt dann

$$|\mu(B)| = \left| \int_B g \, d|\mu| \right| \le \int_B \left(1 - \tfrac{1}{n}\right) d|\mu| = \left(1 - \tfrac{1}{n}\right) |\mu|(B)$$

und deshalb ($|\mu|$ ist σ-additiv!)

$$|\mu|(A_n) \le \left(1 - \tfrac{1}{n}\right) |\mu|(A_n),$$

d.h. $\quad |\mu|(A_n) = 0$.

Daraus folgt die Behauptung.

O.B.d.A. dürfen wir daher

$$|g(s)| = 1 \qquad \forall s \in S$$

annehmen. Setze

$$A^+ = \{ g = 1 \}, \qquad A^- = \{ g = -1 \}.$$

Dann ist $\quad \mu^+(A) = \int_A \frac{g+1}{2} \, d|\mu| = \mu(A \cap A^+)$

und $\quad \mu^-(A) = -\mu(A \cap A^-)$.

Das zeigt $\mu^+ \perp \mu^-$.

Mit den Bezeichnungen des letzten Beweises haben wir den Raum S in $A^+ \cup A^-$ disjunkt zerlegt, wo $(B \in \mathfrak{A})$

$$B \subset A^+ \Rightarrow \mu(B) \ge 0$$

$$B \subset A^- \Rightarrow \mu(B) \le 0.$$

A^+ nennt man aus naheliegenden Gründen eine Positivmenge für μ und A^- eine Negativmenge.

Man kann nun spielend bzgl. eines signierten (oder komplexen) Maßes integrieren: Eine meßbare Funktion f heißt μ-integrierbar, wenn sie $|\mu|$-integrierbar ist, und man setzt

$$\int_S f \, d\mu = \int_S f \, d\mu^+ - \int_S f \, d\mu^-.$$

Ist ν ein signiertes oder komplexes Maß und μ ein positives Maß, so heiße ν absolutstetig oder singulär bzgl. μ, wenn $|\nu|$ diese Eigenschaft hat. Der Satz von Lebesgue-Radon-Nikodym gilt dann sinngemäß (für σ-endliches μ!) und kann durch entsprechende Zerlegung leicht auf 1.4 zurückgeführt werden.

Ein wichtiges Ergebnis in diesem Zusammenhang ist der

Satz von F. und M. Riesz: Ein komplexes Maß ν auf $\mathrm{Bor}\,[0, 2\pi]$ mit

$$\int_0^{2\pi} e^{in\vartheta} \, d\nu(\vartheta) = 0 \qquad \forall n = 0, 1, 2, \dots$$

ist absolutstetig bzgl. des Lebesguemaßes.

(Beweis z.B. in Rudin, Chap. 17.)

3. Martingale

In diesem Abschnitt besprechen wir die allgemeine Strategie, die hinter der Beweismethode von 1.4 steckt. Damit haben wir gleichzeitig Gelegenheit, einen der Kernbegriffe der modernen Wahrscheinlichkeitstheorie kennenzulernen.

Im weiteren bezeichnet $(S, \mathfrak{A}, \mu)$ einen Wahrscheinlichkeitsraum (also $\mu(S)=1$

3.1 Definition Es seien $\mathfrak{A}_1 \subset \mathfrak{A}_2 \subset \cdots$ Unter-σ-Algebren von $\mathfrak{A}$, und (f_n) eine Folge integrierbarer Funktionen. (f_n) heißt ein Martingal, falls

f_n bzgl. $\mathfrak{A}_n$ meßbar

ist und

$$\int_A f_{n+1}\, d\mu = \int_A f_n \, d\mu \qquad \forall A \in \mathfrak{A}_n$$

gilt (alle $n \in \mathbb{N}$).

3.2 Beispiele a) Sei ν ein weiteres Maß auf $\mathfrak{A}$ mit $\nu \leq \mu$. Im Beweis von 1.4 haben wir eine Folge immer feiner werdender Zerlegungen $Z_1, Z_2, \ldots$ sowie meßbare Funktionen

$$f_n \; (= h_{Z_n}) = \sum_{E \in Z_n} \frac{\nu(E)}{\mu(E)} 1_E$$

(mit der momentanen Konvention $\frac{0}{0} = 0$) betrachtet. Setzt man $\mathfrak{A}_n = \sigma(Z_n)$, so zeigt $(*)$ auf S. 129 die Martingaleigenschaft von (f_n).

b) Auf $S = \,]0,1]$ sei $\mathfrak{A}_n$ die von den dyadischen Intervallen

$$]0, 2^{-n}], \;]2^{-n}, 2\cdot 2^{-n}], \; \ldots, \;]1-2^{-n}, 1].$$

erzeugte (endliche!) σ-Algebra. $f_n = 2^n \, 1_{]0, 2^{-n}]}$ definiert dann ein Martingal. (Hier sei natürlich $\mu =$ Lebesguemaß.)

c) Seien $(S, \mathfrak{A}, \mu)$ und $\mathfrak{A}_n$ wie unter b). Seien $d_1, d_2, \ldots$ die in Beispiel I.1.6 und II.1.5 betrachteten meßbaren Funktionen. Wir symmetrisieren und normalisieren die d_n folgendermaßen:

$$r_n := 2 d_n - 1.$$

(Skizze!) Die r_n heißen Rademacherfunktionen und erfüllen offenbar

$$\int_0^1 r_n \, d\lambda = 0, \qquad \int_0^1 r_n^2 \, d\lambda = 1, \qquad \int_0^1 r_n r_m \, d\lambda = 0 \quad (n \neq m)$$ *)

Ferner ist r_n $\mathfrak{A}_n$-meßbar. Das zeigt, daß

$$f_n = \sum_{k=1}^n \frac{r_k}{k}$$

ein Martingal ist.

3.3 Satz $\mathfrak{A}_0$ sei eine Unter-σ-Algebra von $\mathfrak{A}$ und f eine integrierbare Funktion. Dann existiert eine bis auf f.ü.-Gleichheit eindeutig bestimmte $\mathfrak{A}_0$-meßbare integrierbare Funktion g mit

$$\int_A f \, d\mu = \int_A g \, d\mu \qquad \forall A \in \mathfrak{A}_0.$$

Beweis: Sei $\nu: \mathfrak{A}_0 \to \mathbb{R}$ das bzgl. $\mu|_{\mathfrak{A}_0}$ absolutstetige signierte Maß

$$\nu(A) = \int_A f \, d\mu \qquad (A \in \mathfrak{A}_0).$$

Wende nun den Satz von Radon-Nikodym für signierte Maße an!

3.4 Definition Unter den Voraussetzungen von 3.3 heißt g die bedingte Erwartung von f bzgl. $\mathfrak{A}_0$. Bezeichnung: $\mathbb{E}(f \mid \mathfrak{A}_0)$.

(Das ist eigentlich nicht ganz korrekt, da g nicht wirklich eindeutig bestimmt ist; wohl aber die Äquivalenzklasse von g in $L^1(S, \mathfrak{A}_0, \mu|_{\mathfrak{A}_0})$, was diese kleine Schlamperei akzeptabel macht.)

*) Die letzten beiden Bedingungen besagen, daß die r_n ein Orthonormalsystem bezüglich des Skalarprodukts von $L^2(\lambda)$ bilden!

Mit Def. 3.4 kann die Martingaleigenschaft auch durch

$$\mathbb{E}(f_{n+1} \mid \mathfrak{A}_n) = f_n \qquad \forall n \in \mathbb{N}$$

beschrieben werden. Einfache Eigenschaften der bedingten Erwartung sind:

<u>3.5 Lemma</u> $\mathfrak{A}_0$ sei eine Unter-σ-Algebra von $\mathfrak{A}$.

a) $\mathbb{E}(\cdot \mid \mathfrak{A}_0)$ ist eine lineare Abbildung.

b) $f \geq 0$ f.ü. $\Rightarrow$ $\mathbb{E}(f \mid \mathfrak{A}_0) \geq 0$ f.ü. („$\mathbb{E}(\cdot \mid \mathfrak{A}_0)$ ist positiv")

c) h $\mathfrak{A}_0$-meßbar und beschränkt $\Rightarrow$ $\mathbb{E}(f \cdot h \mid \mathfrak{A}_0) = h \cdot \mathbb{E}(f \mid \mathfrak{A}_0)$

d) $\| \mathbb{E}(f \mid \mathfrak{A}_0) \|_{L^1} \leq \| f \|_{L^1}$

$$\left.\begin{aligned} &\text{e)} \quad \int_S f^2 \, d\mu = \int_S (f - \mathbb{E}(f \mid \mathfrak{A}_0))^2 \, d\mu + \int_S (\mathbb{E}(f \mid \mathfrak{A}_0))^2 \, d\mu \\ &\text{f)} \quad \| \mathbb{E}(f \mid \mathfrak{A}_0) \|_{L^2} \leq \| f \|_{L^2} \end{aligned}\right\} \text{für } f \in L^2(\mu)$$

Beweis: a) und b) sind klar.

c) beweist man zuerst für $h = 1_A$ mit $A \in \mathfrak{A}_0$, dann für $\mathfrak{A}_0$-Treppenfunktionen und zum Schluß für deren gleichmäßige Limiten. (Beachte II. 1.8 c))

d)
$$\begin{aligned} \int_S |\mathbb{E}(f \mid \mathfrak{A}_0)| \, d\mu &= \int_S | \mathbb{E}(f^+ - f^- \mid \mathfrak{A}_0) | \, d\mu \\ &\overset{a)}{=} \int_S | \mathbb{E}(f^+ \mid \mathfrak{A}_0) - \mathbb{E}(f^- \mid \mathfrak{A}_0) | \, d\mu \\ &\leq \int_S \left(|\mathbb{E}(f^+ \mid \mathfrak{A}_0)| + |\mathbb{E}(f^- \mid \mathfrak{A}_0)| \right) d\mu \\ &\overset{b)}{=} \int_S \left(\mathbb{E}(f^+ \mid \mathfrak{A}_0) + \mathbb{E}(f^- \mid \mathfrak{A}_0) \right) d\mu \\ &\overset{a)}{=} \int_S \mathbb{E}(f^+ + f^- \mid \mathfrak{A}_0) \, d\mu \\ &\overset{S \in \mathfrak{A}_0}{=} \int_S (f^+ + f^-) \, d\mu \\ &= \int_S |f| \, d\mu . \end{aligned}$$

e)[*] (Das sollte an (**) von S. 129 erinnern!)

$$\begin{aligned} \int_S (f - \mathbb{E}(f \mid \mathfrak{A}_0))^2 \, d\mu &= \int_S f^2 \, d\mu - 2 \underbrace{\int_S f \cdot \mathbb{E}(f \mid \mathfrak{A}_0) \, d\mu}_{= \int_S \mathbb{E}(f \mid \mathfrak{A}_0)^2 \, d\mu} + \int_S \mathbb{E}(f \mid \mathfrak{A}_0)^2 \, d\mu \\ &\qquad [\text{c) mit } h = \mathbb{E}(f \mid \mathfrak{A}_0), \text{ beachte } S \in \mathfrak{A}_0] \\ &= \int_S f^2 \, d\mu - \int_S \mathbb{E}(f \mid \mathfrak{A}_0)^2 \, d\mu \end{aligned}$$

f) folgt aus e).

[Der obige Beweis für e) und f) ist zunächst nur korrekt, wenn f (und folglich (!) $\mathbb{E}(f \mid \mathfrak{A}_0)$) beschränkt sind. Der allgemeine Fall ergibt sich aus diesem durch ein Limesargument; siehe dazu II. 4.7.]

Nun können wir den Hauptsatz dieses Abschnitts formulieren und beweisen.

<u>3.6 Satz</u> Sei (f_n) ein L^2-beschränktes Martingal (d.h. $\sup_n \|f_n\|_{L^2} < \infty$). Dann existiert eine integrierbare Funktion f mit

a) $\lim_{n \to \infty} f_n = f$ f.ü.

b) $\|f_n - f\|_{L^2} \to 0$

c) $\|f_n - f\|_{L^1} \to 0$

d) $\mathbb{E}(f \mid \mathfrak{A}_n) = f_n$.

Beweis: (Vgl. den ersten Schritt im Beweis von 1.4!)
Wir zeigen zuerst, daß (f_n) eine L^2-Cauchyfolge ist.

---

[*] In der Sprache der Hilbertraumtheorie bedeutet e), daß $\mathbb{E}(\cdot \mid \mathfrak{A}_0)$ eine Orthogonalprojektion ist.

Sei etwa $n \geq m$. Dann ist

$$\|f_n - f_m\|_{L^2}^2 \underset{f_m = \mathbb{E}(f_n | \mathfrak{A}_m)}{=} \|f_n\|_{L^2}^2 - \|f_m\|_{L^2}^2 \qquad (3.5\ e))$$

$$\underset{n,m \to \infty}{\longrightarrow} 0 ,$$

denn $(\|f_n\|_{L^2}^2)$ ist monoton wachsend (3.5 f)) und beschränkt (nach Voraussetzung).

Da $L^2(\mu)$ vollständig ist (II.4.4*), existiert $f \in L^2(\mu)$ mit b).
Wegen $\|\cdot\|_{L^1} \leq \|\cdot\|_{L^2}$ ($\mu(S) = 1$, Höldersche Ungleichung) ist auch $f \in L^1(\mu)$, und es gilt c). d) folgt nun sehr einfach. (Wie?)
Es bleibt a) zu zeigen. Dazu benötigen wir die

Doobsche Maximalungleichung: Für ein Martingal (φ_n) gilt ($p \in \mathbb{N}$)

$$\mu(\{\max_{n \leq p} |\varphi_n| > \alpha\}) \leq \frac{1}{\alpha} \int_S |\varphi_p| \, d\mu .$$

Sei $E = \{\max_{n \leq p} |\varphi_n| > \alpha\}$ und sei

$$E_\ell = \{|\varphi_1| \leq \alpha, \dots, |\varphi_{\ell-1}| \leq \alpha, |\varphi_\ell| > \alpha\}.$$

Dann ist $E = \bigcup_{\ell=1}^{p} E_\ell$ (disjunkte Vereinigung), und $E_\ell \in \mathfrak{A}_\ell$. Es folgt

$$\int_S |\varphi_p| \, d\mu \geq \int_E |\varphi_p| \, d\mu = \sum_{\ell=1}^{p} \int_{E_\ell} |\varphi_p| \, d\mu$$
$$\underset{3.5\,d)}{\geq} \sum_{\ell=1}^{p} \int_{E_\ell} |\varphi_\ell| \, d\mu$$
$$\geq \sum_{\ell=1}^{p} \int_{E_\ell} \alpha \, d\mu$$
$$= \alpha \sum_{\ell=1}^{p} \mu(E_\ell)$$
$$= \alpha \, \mu(E)$$

Wähle nun Indizes m_k ($k \in \mathbb{N}$) mit $\sum_{k=1}^{\infty} k \cdot \|f - f_{m_k}\| < \infty$, was nach c) möglich ist. Nun wende die Doobsche Ungleichung auf das Martingal (!) $(f_n - f_{m_k})_{n \geq m_k}$ an und lasse $p \to \infty$ streben. Es folgt dann

$$\mu(\{\sup_{m_k \leq n} |f_n - f_{m_k}| > \tfrac{1}{k}\}) \leq k \cdot \int_S |f - f_{m_k}| \, d\mu$$

$$\Rightarrow \sum_{k=1}^{\infty} \mu(\{\sup_{m_k \leq n} |f_n - f_{m_k}| > \tfrac{1}{k}\}) < \infty .$$

Mit dem Borel-Cantelli-Lemma (S. 90) folgt für μ-fast alle s:

$$\exists k_0 \; \forall k \geq k_0 \quad \sup_{n \geq m_k} |f_n(s) - f_{m_k}(s)| \leq \frac{1}{k} .$$

D.h., $(f_n(s))$ ist eine Cauchyfolge für diese s. [Für so ein s wähle k_0 wie oben. Ist $k \geq k_0$ und sind $n, m \geq m_k$, so ist ja

$$|f_n(s) - f_m(s)| \leq \frac{2}{k} . \;]$$

Also existiert $\lim_{n \to \infty} f_n(s)$ fast überall. Da andererseits eine Teilfolge von (f_n) gewiß f.ü. gegen f konvergiert (I.3.9), folgt a).

Es ist auch richtig, daß L^1-beschränkte Martingale (d.h. $\sup_n \|f_n\|_{L^1} < \infty$) fast überall konvergieren (etwa gegen f), unter dieser schwächeren Voraussetzung gilt jedoch i.a. nicht mehr $\|f_n - f\|_{L^1} \to 0$. (Vgl. etwa das Beispiel 3.2 b).)

Abschließend werde der Martingalkonvergenzsatz 3.6 auf das Beispiel 3.2 c) angewandt, um das auf S. 60 angekündigte Resultat zu beweisen. Dort wurde erwähnt, daß die Menge N der „normalen" Zahlen in $]0,1]$ Lebesguemaß 1 hat.

Den Beweis dieser Aussage beginnen wir mit dem Nachweis, daß das Martingal (f_n) aus 3.2 c) L^2-beschränkt ist:

$$\|f_n\|_{L^2}^2 = \int_0^1 \Big(\sum_{k=1}^n \frac{r_k}{k}\Big)^2 \, d\lambda$$
$$= \int_0^1 \Big(\sum_{k=1}^n \frac{r_k}{k}\Big)\Big(\sum_{l=1}^n \frac{r_l}{l}\Big) \, d\lambda$$
$$= \sum_{k,l=1}^n \frac{1}{k \cdot l} \int_0^1 r_k r_l \, d\lambda$$
$$= \sum_{k=1}^n \frac{1}{k^2} \qquad \text{(vgl. S. 143 oben)}$$
$$\leq \sum_{k=1}^\infty \frac{1}{k^2} \qquad (< \infty).$$

Nach dem Martingalkonvergenzsatz existiert

$$\lim_{n\to\infty} \sum_{k=1}^n \frac{r_k}{k} \qquad \text{fast überall.}$$

Daraus folgt

$$\frac{1}{n}\sum_{k=1}^n r_k \to 0 \qquad \text{fast überall.}$$

Es gilt nämlich das Lemma von Kronecker:
Sei (a_n) eine Folge reeller Zahlen, für die $\sum_{k=1}^\infty \frac{a_k}{k}$ konverg
Dann gilt $\frac{1}{n}\sum_{n=1}^n a_k \to 0$.

Zum Beweis des Lemmas setze $s_n = \sum_{k=1}^n \frac{a_k}{k}$ und $s = \sum$
In einer unausweichlichen Übungsaufgabe muß jeder Student im 1. Semester beweisen, daß aus $s_n \to s$ auch

$$\frac{1}{n}\sum_{k=1}^n s_k \to s$$

folgt. Das Lemma ergibt sich jetzt aus

$$\frac{1}{n}\sum_{k=1}^n a_k = s_n - \frac{1}{n}\sum_{k=1}^{n-1} s_k.$$

Übersetzung in die ursprünglichen d_k zeigt

$$\frac{1}{n}\sum_{k=1}^n d_k \to \frac{1}{2} \qquad \text{fast überall,}$$

was zu zeigen war.

Mit derselben Methode kann auch das sog. starke Gesetz der großen Zahlen der Wahrscheinlichkeitstheorie bewiesen werden, wovon die obige Aussage ein Spezialfall ist.

4. Maße auf dem $\mathbb{R}^k$

Im folgenden Abschnitt wird eine alternative Beschreibung der Radon-Nikodym-Dichte eines bzgl. λ^k absolutstetigen Maßes gegeben. Dabei tritt der Differentiationsaspekt besonders klar zutage; die Methoden werden im Abschnitt IV.2 zu Lebesgues Version des Hauptsatzes der Differential- und Integralrechnung führen.
Wir werden hier λ für das k-dimensionale Lebesguemaß schreiben; ferner bezeichnet $B(x,r)$ die offene Kugel um $x \in \mathbb{R}^k$ mit Radius r bzgl. der euklidischen Norm (jede andere würde es auch tun):

$$B(x,r) = \{ y \in \mathbb{R}^k : \|x-y\| < r \}.$$

μ bezeichnet ein signiertes Maß auf $\mathrm{Bor}(\mathbb{R}^k)$.

<u>4.1 Definition</u> a) $\displaystyle (Q_r\mu)(x) = \frac{\mu(B(x,r))}{\lambda(B(x,r))}$ $\quad (x \in \mathbb{R}^k,\ r>0)$

b) $\displaystyle (D\mu)(x) = \lim_{r\to 0} (Q_r\mu)(x)$ (falls dieser Limes existiert)

c) $\displaystyle (Q^*\mu)(x) = \sup_{0<r<\infty} (Q_r|\mu|)(x)$.

($Q^*\mu$ heißt Hardy - Littlewood - Maximalfunktion.)

<u>4.2 Lemma</u> Für $\alpha \in \mathbb{R}$ ist $\{Q^*\mu > \alpha\}$ offen. Folglich ist $Q^*\mu$ meßbar

Beweis: O.E. ist $\mu \geq 0$. Sei $E = \{Q^*\mu > \alpha\}$. Für $x \in E$ existiert dann $r > 0$ mit $(Q_r\mu)(x) > \alpha$. Sei zuerst $\alpha > 0$. Wähle $\delta > 0$ mit

$$(r+\delta)^k < r^k \frac{(Q_r\mu)(x)}{\alpha}.$$

Wir möchten $B(x,\delta) \subset \{Q^*\mu > \alpha\}$ zeigen.

Sei also $\|y-x\| < \delta$. Da nach der Dreiecksungleichung $B(x,r) \subset B(y, r+\delta)$ gilt, folgt

$$(Q^*\mu)(y) \geq \frac{\mu(B(y,r+\delta))}{\lambda(B(y,r+\delta))} \geq \frac{\mu(B(x,r))}{\lambda(B(y,r+\delta))}$$

$$= \frac{\mu(B(x,r))}{\left(\frac{r+\delta}{r}\right)^k \lambda(B(x,r))}$$ (denn λ ist translationsinvariant, und $\lambda(B(0,r)) = r^k \lambda(B(0,1))$ nach Kor. I.4.4 [$T = r\,\mathrm{Id}$]

$$= \left(\frac{r}{r+\delta}\right)^k (Q_r\mu)(x)$$

$$> \alpha.$$

Daher ist $\{Q^*\mu > \alpha\}$ offen, falls $\alpha > 0$. Das stimmt dann auch für $\alpha = 0$, denn $\{Q^*\mu > 0\} = \bigcup_{n=1}^{\infty} \{Q^*\mu > \frac{1}{n}\}$; und für $\alpha < 0$ sowieso, da in diesem Fall $\{Q^*\mu > \alpha\} = \mathbb{R}^k$ ist.

<u>4.3 Lemma</u> $W \subset \mathbb{R}^k$ sei beschränkt, gelte etwa $W \subset \bigcup_{i=1}^{n} B(x_i, r_i)$.

Dann existiert eine Indexmenge $I \subset \{1, \dots, n\}$ mit:

a) Die $B(x_i, r_i)$ $(i \in I)$ sind paarweise disjunkt.

b) $W \subset \bigcup_{i\in I} B(x_i, 3r_i)$

c) $\lambda(W) \leq 3^k \sum_{i\in I} \lambda(B(x_i,r_i)) = 3^k \lambda\left(\bigcup_{i\in I} B(x_i,r_i)\right)$

Beweis: Man darf $r_1 \geq r_2 \geq \dots \geq r_n$ annehmen.

Setze $i_1 = 1$. Sei i_2 der kleinste Index i, für den $B(x_i,r_i)$ zu $B(x_{i_1}, r_{i_1})$ disjunkt ist, sei i_3 der kleinste Index $i \geq i_2$, für den $B(x_i,r_i)$ zu $B(x_{i_1}, r_{i_1})$ und $B(x_{i_2}, r_{i_2})$ disjunkt ist etc. Das Verfahren bricht nach endlich vielen (etwa p) Schritten ab; setze $I = \{i_1, \dots, i_p\}$.

Nach Konstruktion ist dann a) erfüllt. b) sieht man so:

Ist $i_0 \in I$, so gilt natürlich $B(x_{i_0}, r_{i_0}) \subset B(x_{i_0}, 3r_{i_0})$.

Ist $i_0 \notin I$, so existiert $i < i_0$, $i \in I$ mit $B(x_i,r_i) \cap B(x_{i_0}, r_{i_0}) \neq \emptyset$.

Daher $B(x_{i_0}, r_{i_0}) \subset B(x_i, 3r_i)$, da $r_i \geq r_{i_0}$.

(Dreiecksungleichung! Skizze!) In jedem Fall ist

$$B(x_{i_0}, r_{i_0}) \subset \bigcup_{i\in I} B(x_i, 3r_i)$$

$$\Rightarrow \quad W \subset \bigcup_{i=1}^{n} B(x_i,r_i) \subset \bigcup_{i\in I} B(x_i, 3r_i).$$

c) ist eine Folge von b) und a). ($\lambda(B(x,\alpha r)) = \alpha^k \lambda(B(x,r))$; I.4.4!)

Dieses Lemma ermöglicht die entscheidende technische Abschätzung:

4.4 Satz Für $\alpha > 0$ ist

$$\lambda(\{Q^*\mu > \alpha\}) \le \frac{3^k \cdot |\mu|(\mathbb{R}^k)}{\alpha}$$

Beweis: O.E. ist $\mu \ge 0$. Sei K eine kompakte Teilmenge von $\{Q^*\mu > \alpha\}$. Für jede $x \in K$ gibt es dann eine offene Kugel $B_x = B(x, r_x)$ mit $\mu(B_x) > \alpha\, \lambda(B_x)$. Diese Kugeln bilden natürlich eine Überdeckung von K, und wegen der Kompaktheit von K reichen bereits endlich viele der Kugeln zur Überdeckung aus:

$$K \subset \bigcup_{i=1}^{n} B_{x_i}$$

4.3 c) liefert dann für eine geeignete Teilmenge $I \subset \{1, \dots, n\}$

$$\begin{aligned} \lambda(K) &\le 3^k \sum_{i \in I} \lambda(B_{x_i}) \\ &\le \frac{3^k}{\alpha} \sum_{i \in I} \mu(B_{x_i}) \\ &\underset{4.3c)}{=} \frac{3^k}{\alpha}\, \mu\Big(\bigcup_{i \in I} B_{x_i}\Big) \\ &\le \frac{3^k}{\alpha}\, \mu(\mathbb{R}^k) \end{aligned}$$

Die Regularität von λ (I.4.5) impliziert die Behauptung; beachte mit

Ist μ absolutstetig (bzgl. λ) mit Dichte f, also $\mu(A) = \int_A f \, d\lambda$, so setzen wir

$$Q^* f := Q^* \mu$$

Da aus der Hahn-Jordan-Zerlegung schnell

$$(\otimes) \qquad |\mu|(A) = \int_A |f| \, d\lambda$$

folgt, erhält man:

4.5 Korollar $\lambda(\{Q^* f > \alpha\}) \le \frac{3^k}{\alpha} \cdot \|f\|_{L^1}$ $\quad (\alpha > 0)$

Wäre $Q^* f \in L^1(\mathbb{R}^k)$ (was i.a. nicht stimmt), folgte eine Abschätzung vom Typ (*) $\lambda(\{Q^* f > \alpha\}) \le \frac{\text{konst.}}{\alpha}$ aus der Tschebyscheff-Ungleichung I.3.8. Die Aussage in 4.5 ist daher etwas schwächer; man sagt, der „Maximaloperator" Q^* sei vom schwachen Typ 1-1. (da er L^1 in den „schwachen L^1" *) abbildet). Die Doobsche Maximalungleichung (S. 146) ist ebenfalls von diesem Typ.

4.6 Definition Sei $f: \mathbb{R}^k \to \mathbb{R}$ λ-integrierbar. $x \in \mathbb{R}^k$ heißt Lebesguepunkt von f, falls

$$\lim_{r \to 0} \frac{1}{\lambda(B(x,r))} \int_{B(x,r)} |f(x) - f(y)| \, d\lambda(y) = 0$$

Offenbar ist für stetige f jeder Punkt Lebesguepunkt (Beweis?). Lebesguepunkt zu sein ist daher als abgemilderte Stetigkeitseigenschaft zu verstehen.

(Was sind z.B. die Lebesguepunkte von $f = 1_{[0,1]}$?) — Interessanterweise hat jede integrierbare Funktion Lebesguepunkte. Mehr noch:

*) Damit ist die Menge der meßbaren Funktionen g gemeint, die eine Abschätzung (*) gestatten, d.h. $\sup_{\alpha > 0} \alpha \cdot \lambda(\{g > \alpha\}) < \infty$.

4.7 Satz Ist f λ-integrierbar, so ist λ-fast jedes $x \in \mathbb{R}^k$ Lebesguepunkt von f.

Beweis: Setze

$$(T_r f)(x) = \frac{1}{\lambda(B(x,r))} \int_{B(x,r)} |f(x) - f(y)| \, d\lambda(y)$$

$$(Tf)(x) = \limsup_{r \to 0} (T_r f)(x).$$

Wir haben $Tf = 0$ f.ü. zu zeigen.

Dazu benötigen wir die Aussage:

Die stetigen integrierbaren Funktionen liegen dicht in $L^1(\mathbb{R}^k)$.

[Das ist allgemeiner als II. 4. 9 a) und wird in größerer Allgemeinheit in V.2 gezeigt. Hier sei ein direkter Beweis skizziert:

Sei $f \in L^1(\mathbb{R}^k)$. Wähle zuerst $\tilde{f} \in L^1(\mathbb{R}^k)$ mit

$\exists R > 0: \quad \tilde{f}(x) = 0 \quad$ f.ü. falls $\|x\| \geq R$

und $\quad \|\tilde{f} - f\|_{L^1} \leq \varepsilon$

Sei des weiteren $\varphi \geq 0$ eine stetige Funktion mit

$$\varphi(x) = 0 \quad \forall \, \|x\| \geq 1$$

$$\int_{\mathbb{R}^k} \varphi \, d\lambda = 1.$$

Sei $\quad \tilde{f}_n(x) = \int_{\mathbb{R}^k} \tilde{f}(y) \, \varphi_n(y - x) \, d\lambda(y),$

wo $\quad \varphi_n(x) := n^k \varphi(nx).$

Man weist dann nach, daß $\tilde{f}_n$ stetig und integrierbar ist und daß $\|\tilde{f}_n - \tilde{f}\|_{L^1} \to 0$ gilt.

Sind nun $\varepsilon > 0$ und $n \in \mathbb{N}$ gegeben, so existiert eine stetige Funktion g mit

$$\|f - g\|_{L^1} \leq \frac{1}{n}.$$

Nach der Vorbemerkung gilt $\quad Tg = 0.$

Mit $h = f - g$ erhält man aus

$$(T_r h)(x) \leq \frac{1}{\lambda(B(x,r))} \int_{B(x,r)} |h| \, d\lambda + |h(x)|$$

$$\leq (Q^* h)(x) + |h|(x)$$

sowie $\quad T_r f \leq T_r(h) + T_r(g) \quad$ (Dreiecksungleichung)

die Ungleichung

$$Tf \leq Q^* h + |h| + Tg = Q^* h + |h|.$$

Deshalb ist (bemerke, daß h von n abhängt!).

$$\{Tf > 2\varepsilon\} \subset \{Q^* h > \varepsilon\} \cup \{|h| > \varepsilon\} =: E_{\varepsilon,n}.$$

Der Satz ist bewiesen, wenn $E_\varepsilon := \bigcap_{n \in \mathbb{N}} E_{\varepsilon,n}$ als Nullmenge erkannt ist (E_ε ist meßbar nach 4.2!), denn

$$\{Tf \neq 0\} \subset \bigcup_{m \in \mathbb{N}} E_{1/m} :$$

$$\lambda(E_{\varepsilon,n}) \leq \frac{3^k}{\varepsilon} \|h\|_{L^1} + \frac{1}{\varepsilon} \|h\|_{L^1}$$

(nach 4.5 und der Tschebyscheff-Ungleichung II.3.8)

$$\leq \frac{3^k + 1}{\varepsilon \cdot n} \xrightarrow[n \to \infty]{} 0$$

$$\Rightarrow \lambda(E_\varepsilon) = 0.$$

(Bemerke, daß nichts über die Meßbarkeit von Tf ausgesagt wurde!)

4.8 Korollar: Sei $\mu \ll \lambda$ mit Radon-Nikodym-Dichte f. Wenn $x \in \mathbb{R}^k$ Lebesguepunkt von f ist, existiert $D\mu(x)$ [Def. 4.1]. Ferner gilt

$$D\mu = f \qquad \lambda\text{-f.ü.}$$

Beweis: Da μ endliche Variation hat, ist f integrierbar ((⊛) von 5.1(2)). Sei $x \in \mathbb{R}^k$ Lebesguepunkt von f. Dann gilt

$$f(x) = \lim_{r\to 0} \frac{1}{\lambda(B(x,r))} \int_{B(x,r)} f \, d\lambda$$

$$= \lim_{r\to 0} \frac{\mu(B(x,r))}{\lambda(B(x,r))}$$

$$= (D\mu)(x).$$

Da nach 4.7 fast jeder Punkt Lebesguepunkt ist, folgt $f = D\mu$ f.ü.

4.8 eröffnet eine Beschreibung von f, die die Bezeichnung Radon-Nikodym-Ableitung und das Symbol $f = \frac{d\mu}{d\lambda}$ plausibel macht.

Satz 4.7 wird in IV.2/3 weiter ausgebaut.

IV. Anwendungen in der klassischen Analysis

IV.1. Das Riemannintegral

Für die Indikatorfunktion eines Intervalls stimmen nach Definition Riemann- und Lebesgueintegral überein. Daraus haben wir in Beispiel II.2.11 a) geschlossen, daß auch für stetige f auf kompakten Intervallen

$$R\text{-}\int_a^b f(t)\,dt = L\text{-}\int_a^b f(t)\,dt$$

gilt. Wir wollen zeigen, daß dieses Resultat für alle (beschränkten!) R-integrierbaren Funktionen gilt.

1.1 Satz: Ist $f: [a,b] \to \mathbb{R}$ Riemann-integrierbar, so ist f $\mathrm{Leb}([a,b])$-meßbar, und es gilt

$$R\text{-}\int_a^b f(t)\,dt = L\text{-}\int_a^b f(t)\,dt.$$

Beispiel 1.3 wird zeigen, daß eine Riemann-integrierbare Funktion nicht $\mathrm{Bor}([a,b])$-meßbar zu sein braucht!

Beweis: Ist f Riemann-integrierbar, so existieren Zerlegungen Z_n von $[a,b]$, so daß die zugehörigen Folgen der Ober- und Untersummen gegen $R\text{-}\int_a^b f(t)\,dt$ konvergieren. Jede Obersumme kann nun in naheliegender Weise als Integral einer Treppenfunktion $\psi \geq f$, deren Stufen Intervalle sind, interpretiert werden. Analoges gilt für Untersummen, wobei diesmal die Treppenfunktion $\varphi \le f$ ist. Die Riemann-Integrabilität von f bedeutet in dieser Interpretation, daß es Intervall-Treppenfunktionen φ_n, ψ_n gibt mit

$$\varphi_n \le f \le \psi_n$$
$$\int_a^b \varphi_n(t)\,dt \longrightarrow R\text{-}\int_a^b f(t)\,dt$$
$$\int_a^b \psi_n(t)\,dt \longrightarrow R\text{-}\int_a^b f(t)\,dt.$$

Sei $\varphi = \sup_n \varphi_n$, $\psi = \inf_n \psi_n$, dann sind φ und ψ Borel-meßbar (II.1.4). Außerdem ist natürlich $\varphi \le f \le \psi$.

O.E. dürfen wir annehmen, daß die Folge (φ_n) monoton wächst und die Folge (ψ_n) monoton fällt. (Dem entspricht, daß stets die [illegible] als Verfeinerung von Z_n angenommen werden darf.) Der Satz von Beppo Levi impliziert dann

$$L\text{-}\int_a^b \varphi(t)\,dt = \sup_n \int_a^b \varphi_n(t)\,dt = R\text{-}\int_a^b f(t)\,dt$$

und

$$L\text{-}\int_a^b \psi(t)\,dt = \inf_n \int_a^b \psi_n(t)\,dt = R\text{-}\int_a^b f(t)\,dt.$$

Also ist $\psi - \varphi$ eine positive Borel-meßbare Funktion mit

$$L\text{-}\int_a^b (\psi-\varphi)\,d\lambda = 0,$$

woraus nach II.3.3 $\varphi = \psi$ f.ü. folgt.

Das heißt aber $\varphi = f = \psi$ f.ü., so daß f f.ü. mit einer Borel-meßbaren Funktion übereinstimmt. Punkt f) von S. 82 mit I.3.8 zeigen, daß f Lebesgue-meßbar ist, ferner gilt nun

$$L\text{-}\int_a^b f(t)\,dt = L\text{-}\int_a^b \varphi(t)\,dt = R\text{-}\int_a^b f(t)\,dt.$$

1.2 Satz Eine beschränkte Funktion $f : [a,b] \to \mathbb{R}$ ist genau dann Riemann-integrierbar, wenn sie fast überall stetig ist (d.h. die Menge der Unstetigkeitspunkte eine λ-Nullmenge ist).

Vergleiche diesen Satz mit dem Satz von Lusin (II.1.9) und Satz III.4.7!

Übrigens ist die Menge Ω_f der Unstetigkeitspunkte stets eine Borelmenge. Um das einzusehen, definiere die Oszillation osz_f von f durch

$$\operatorname{osz}_f(t) = \inf_U \sup_{x,y\in U} |f(x)-f(y)|,$$

wo das Infimum über alle Umgebungen U von t zu erstrecken ist. Man zeigt dann leicht

- $\Omega_f = \{\operatorname{osz}_f > 0\}$
- $\forall \alpha > 0 \quad \{\operatorname{osz}_f < \alpha\}$ offen ($\Leftrightarrow \{\operatorname{osz}_f \ge \alpha\}$ abgeschlossen).

Beweis von 1.2: Zuerst sei f Riemann-integrierbar. Mit den Bezeichnungen des Beweises von 1.1 setzen wir

$$N = \{\varphi \ne \psi\} \cup \bigcup_n \Omega_{\varphi_n} \cup \bigcup_n \Omega_{\psi_n} \cup \{a,b\}.$$

Das ist eine Nullmenge, denn jedes φ_n (bzw. ψ_n) hat ja nur endlich viele Unstetigkeitsstellen. Wir zeigen nun

$$t \notin N \Rightarrow f \text{ stetig bei } t,$$

so daß f f.ü. stetig ist.

Sei $t \notin N$. Dann gilt nach Konstruktion

$$f(t) = \sup_n \varphi_n(t) = \inf_n \psi_n(t).$$

Zu $\varepsilon > 0$ wähle $m \in \mathbb{N}$ mit

$$\psi_m(t) - \varepsilon \le f(t) \le \varphi_m(t) + \varepsilon$$

In einer δ-Umgebung U von t sind ψ_m und φ_m konstant (denn t ist kein Teilpunkt der zugehörigen Zerlegungen), also gilt für alle $s \in U$ wegen $\varphi_m \le f \le \psi_m$

$$\begin{aligned} f(s) - \varepsilon \le \psi_m(s) - \varepsilon &= \psi_m(t) - \varepsilon \\ &\le f(t) \\ &\le \varphi_m(t) + \varepsilon = \varphi_m(s) + \varepsilon \le f(s) + \varepsilon \end{aligned}$$

$$\Rightarrow \quad |f(s) - f(t)| \le \varepsilon.$$

∎

Umgekehrt sei f f.ü. stetig. Wir betrachten die Zerlegungen

$$Z_n = \{a,\ a + 2^{-n}(b-a),\ a + 2 \cdot 2^{-n}(b-a),\ \dots,\ b\}$$

sowie die zugehörigen Ober- und Untersummen bzw. – wie im Beweis von 1.1 – die zugehörigen Intervall-Treppenfunktionen ψ_n und φ_n. Es gilt dann

$$\varphi_1 \le \varphi_2 \le \dots \le f$$
$$\psi_1 \ge \psi_2 \ge \dots \ge f$$

Setze wieder $\varphi = \sup_n \varphi_n$, $\psi = \inf_n \psi_n$, so daß

$$\varphi \le f \le \psi.$$

Behauptung: f stetig bei $t \Rightarrow \varphi(t) = \psi(t)$.

Zu $\varepsilon > 0$ gibt es dann $m \in \mathbb{N}$ mit

$$f(t) - \varepsilon \le \varphi_m(t)$$
$$f(t) + \varepsilon \ge \psi_m(t)$$

$$\Rightarrow \quad f(t) - \varepsilon \le \varphi(t)$$
$$f(t) + \varepsilon \ge \psi(t),$$

und da $\varepsilon > 0$ beliebig war

$$\psi(t) \le f(t) \le \varphi(t).$$

Wegen $\varphi \le \psi$ heißt das $\varphi(t) = \psi(t)$. ∎

Damit gilt $\psi = \varphi$ f.ü., und aus dem Satz von Beppo Levi folgt

$$\sup \int_a^b \varphi_n(t)\,dt = \int_a^b \varphi(t)\,dt = \int_a^b \psi(t)\,dt = \inf \int_a^b \psi_n(t)\,dt$$

$$\Rightarrow \quad \int_a^b (\psi_n(t) - \varphi_n(t))\,dt \to 0,$$

und das ist nach dem Riemannschen Integrabilitätskriterium genau die Riemann-Integrierbarkeit von f.

1.3 Beispiel Es gibt eine Riemann-integrierbare Funktion auf $[0,1]$, die nicht Borel-meßbar ist.

Beweis: Die Cantormenge C enthält eine nichtborelsche Teilmenge M (vgl. S. 49). Daher ist 1_M nicht Borel-meßbar. Andererseits ist 1_M auf der offenen Menge $[0,1] \setminus C$ identisch $= 0$, daher sind die Unstetigkeitspunkte in der Nullmenge C enthalten. Nach 1.2 ist 1_M Riemann-integrierbar.

Für uneigentlich Riemann-integrierbare Funktionen ist die Sache etwas verwickelter. So eine Funktion f ist genau dann L-integrierbar, wenn $|f|$ uneigentlich R-integrierbar ist, und L- und R-Integral stimmen dann überein. (Beweis zur Übung!)

IV.2 Der Hauptsatz der Differential- und Integralrechnung

Dieser Satz kann im Rahmen der Analysis I z.B. in folgender Form ausgesprochen werden:

- Ist $f:[a,b]\to\mathbb{R}$ R-integrierbar und stetig bei x_0, so ist $F: x \mapsto \int_a^x f(t)\,dt$ differenzierbar bei x_0 mit $F'(x_0)=f(x_0)$.
- Ist $F:[a,b]\to\mathbb{R}$ stetig differenzierbar [etwas schwächere Voraussetzungen würden genügen, z.B. F' existiert überall, ist beschränkt und R-integrierbar], so gilt $F(x)-F(a)=\int_a^x F'(t)\,dt$.

Wir wollen in diesem Abschnitt Versionen des Hauptsatzes im Rahmen der Lebesgueschen Theorie kennenlernen.

2.1 Satz $f:[a,b]\to\mathbb{R}$ sei λ-integrierbar und $F(x):=\int_a^x f\,d\lambda$. Ist x_0 ein Lebesguepunkt von f [Def. III.4.6]*), so existiert $F'(x_0)$, und es gilt $F'(x_0)=f(x_0)$. Also gilt $F'=f$ fast überall.

Beweis: Gelte $h_n\to 0$, und sei $E_n=[x,x+h_n[$ für $h_n>0$ bzw. $E_n=]x+h_n,x]$ für $h_n<0$. Dann gilt

*) Um allen Formalitäten Genüge zu tun, müßte man eigentlich f durch $f(x)=0$ für $x\notin[a,b]$ zu einer auf $\mathbb{R}$ definierten integrierbaren Funktion fortsetzen.

$$\left|\frac{F(x_0+h_n)-F(x_0)}{h_n}-f(x_0)\right| \le \frac{1}{\lambda(E_n)}\int_{E_n}|f-f(x_0)|\,d\lambda$$

$$\le \frac{1}{2}\cdot\frac{1}{\lambda(B(x_0,|h_n|))}\int_{B(x_0,|h_n|)}|f-f(x_0)|\,d\lambda$$

[denn $E_n\subset B(x_0,|h_n|):=\{y: |x_0-y|<|h_n|\}$ und $\lambda(E_n)=\frac12\lambda(B(x_0,|h_n|))\ (=|h_n|)$]

$\longrightarrow 0$ für Lebesguepunkte x_0.

$F'=f$ f.ü. wg. III.4.7 !

Wenden wir uns nun dem 2. Teil des Hauptsatzes zu: Welche F sind Integrale ihrer Ableitungen? Zur Beantwortung dieser Frage werden folgende Begriffe benötigt.

2.2 Definition $F:[a,b]\to\mathbb{R}$ sei eine Funktion.

a) F heißt von beschränkter Variation, falls

$$\sup \sum_{n=1}^{N} |F(t_{n-1})-F(t_n)| < \infty,$$

wo das Supremum über alle Zerlegungen

$$a=t_0<t_1<\dots<t_N=b \qquad (N\in\mathbb{N})$$

zu erstrecken ist.

b) F heißt absolutstetig, falls es zu jedem $\varepsilon>0$ ein $\delta>0$ gibt, so daß für alle disjunkten Teilintervalle $]a_1,b_1[,\ ]a_2,b_2[,\dots,\ ]a_N,b_N[$

mit $\sum_{n=1}^{N} (b_n - a_n) \le \delta$ gilt:

$$\sum_{n=1}^{N} |F(b_n) - F(a_n)| \le \varepsilon .$$

Über den Zusammenhang der Begriffe „Absolutstetigkeit von Funktionen" und „Absolutstetigkeit von Maßen" berichtet Korollar 2.6 unten.

2.3 Beispiele a) Offenbar sind absolutstetige Funktionen stetig (wähle $N=1$), aber nicht umgekehrt (s.u., Bsp. c)). Hingegen sind Lipschitz-stetige Funktionen (d.h. F mit $\exists L \ \forall s,t \ |F(s) - F(t)| \le L \cdot |s-t|$) natürlich absolutstetig. $t \mapsto \sqrt{t}$ ist eine absolutstetige Funktion auf $[0,1]$, die nicht Lipschitz-stetig ist.

b) Eine monotone Funktion auf $[a,b]$ ist gewiß von beschränkter Variation. Es ist leicht zu sehen, daß die Menge aller Funktionen von beschränkter Variation einen Vektorraum bildet; daher sind auch Differenzen monotoner Funktionen von beschränkter Variation. (Zur Umkehrung dieses Sachverhalts siehe 2.4. b) unten.)

$t \mapsto t \cdot \sin \frac{1}{t}$ bzw. $0 \mapsto 0$ ist nicht von beschränkter Variation auf $[0,1]$.

c) Die in Beispiel I: 4.10 f) konstruierte „Cantor-Funktion" F ist nicht absolutstetig (aber stetig!). Das ist der Skizze von S. 52 zu entnehmen: $[0,1] \setminus \bigcup_{i=1}^{2^m-1} O_i$ (O_i wie dort) besteht aus $2^m =: N$ abgeschlossenen Intervallen, sagen wir $[a_1, b_1], \ldots, [a_N, b_N]$. Für ihre Gesamtlänge gilt $\sum_{i=1}^{N} (b_i - a_i) \longrightarrow 0$ (vgl. S. 50), aber

stets ist

$$\sum_{i=1}^{N} |F(b_i) - F(a_i)| = 1.$$

(Bemerke, daß hier sogar die abgeschlossenen Teilintervalle $[a_1, b_1], \ldots, [a_N, b_N]$ disjunkt waren.)

d) Sei F wie in Satz 2.1. Dann ist F absolutstetig.

Um das einzusehen, betrachte das (endliche!) Maß ν auf $Bor([a,b])$, das durch

$$\nu(A) = \int_A |f| \, d\lambda$$

definiert ist. Wähle gemäß Lemma III.1.3 zu $\varepsilon > 0$ ein $\delta > 0$ mit

$$\lambda(A) \le \delta \Rightarrow \nu(A) \le \varepsilon .$$

Seien $a_1, b_1, \ldots, a_N, b_N$ wie in Def. 2.2 b) mit

$$\sum_{n=1}^{N} (b_n - a_n) \le \delta .$$

Setzt man $A =]a_1, b_1[\cup \ldots \cup]a_N, b_N[$, so folgt daß $\lambda(A) \le \delta$. Es folgt

$$\begin{aligned} \sum_{n=1}^{N} |F(b_n) - F(a_n)| &\le \sum_{n=1}^{N} \int_{a_n}^{b_n} |f| \, d\lambda \\ &= \int_A |f| \, d\lambda \\ &= \nu(A) \le \varepsilon . \end{aligned}$$

Damit ist die Absolutstetigkeit von F als notwendig dafür erkannt, daß F Integral seiner (ihrer?) Ableitung ist. Im Hauptergebnis dieses Abschnitts, Satz 2.5, wird die Umkehrung gezeigt. Dazu ist folgendes Lemma nützlich.

2.4 Lemma $F : [a,b] \to \mathbb{R}$ sei eine Funktion.

a) Ist F absolutstetig, so auch von beschränkter Variation.

b) Ist F von beschränkter Variation, so existieren monoton wachsende $F_1, F_2 : [a,b] \to \mathbb{R}$ mit $F = F_1 - F_2$.

c) Ist F zusätzlich absolutstetig, so können F_1 und F_2 in b) ebenfalls als absolutstetig gewählt werden.

Beweis: a) Wähle $\delta > 0$ zu $\varepsilon = 1$ (gemäß Def. 2.2). Seien $a = t_0 < t_1 < \dots < t_N = b$. Wir werden (*) $\sum_{i=1}^{N} |F(t_i) - F(t_{i-1})| \le M$ zeigen, wo $\frac{b-a}{M} < \delta$ und $M \in \mathbb{N}$ ist. Dazu bemerken wir zuerst, daß o.E. die Punkte $a + \frac{b-a}{M}$, $a + 2\frac{b-a}{M}$, $\dots$, $a + (M-1)\frac{b-a}{M}$ zu den t_i gehören (sonst nehmen wir sie einfach dazu und machen dadurch [nach der Dreiecksungleichung] die linke Seite von (*) größer). Für jedes Teilintervall $]a+(k-1)\frac{b-a}{M}, a + k\frac{b-a}{M}[$ mit den Teilpunkten

$$a + (k-1)\frac{b-a}{M} = t_{i_{k-1}} < t_{i_{k-1}+1} < \dots < t_{i_k} = a + k\frac{b-a}{M}$$

gilt dann nach Wahl von δ (beachte $\sum_{j=i_{k-1}}^{i_k-1} (t_{j+1} - t_j) = \frac{b-a}{M} < \delta$)

$$\sum_{j=i_{k-1}}^{i_k-1} |F(t_{j+1}) - F(t_j)| \le 1$$

Summierung $\sum_{k=1}^{M}$ liefert dann (*).

b) Zu $x \in [a,b]$ setze

$$V_F(x) = \sup \sum_{i=1}^{N} |F(t_i) - F(t_{i-1})| ,$$

wo sich das Supremum über alle $N \in \mathbb{N}$ und

$$a = t_0 < t_1 < \dots < t_N = x$$

erstreckt. Sind solche t_i gegeben und ist $y > x$, so folgt

$$V_F(y) \ge \sum_{i=1}^{N} |F(t_i) - F(t_{i-1})| + |F(y) - F(x)| ,$$

daher

$$V_F(y) \ge V_F(x) + |F(y) - F(x)|$$
$$\ge V_F(x) ,$$

so daß V_F monoton wächst. Ferner ergibt sich daraus

$$V_F(y) \ge V_F(x) + F(y) - F(x)$$

sowie

$$V_F(y) \ge V_F(x) + F(x) - F(y) .$$

Diese Ungleichungen zeigen, daß $F_1 := \frac{1}{2}(V_F + F)$ und $F_2 := \frac{1}{2}(V_F - F)$ wachsen; und natürlich ist $F = F_1 - F_2$.

c) Da die absolutstetigen Funktionen einen Vektorraum bilden (Beweis zur Übung), ist mit den obigen Bezeichnungen nur die Absolutstetigkeit von V_F zu zeigen. Wähle zu $\varepsilon > 0$ ein $\delta > 0$ gemäß der Absolutstetigkeit von F. Alsdann beobachten wir, daß für $\alpha < \beta$

$$V_F(\beta) - V_F(\alpha) = \sup_{\substack{N \in \mathbb{N} \\ \alpha = t_0 < \dots < t_N = \beta}} \sum_{i=1}^{N} |F(t_i) - F(t_{i-1})| \qquad (**)$$

gilt. Sind nun $]a_1, b_1[, \dots,]a_r, b_r[$ paarweise disjunkte Teilintervalle von $[a,b]$ mit Gesamtlänge $\sum_{j=1}^{r} (b_j - a_j) < \delta$, so ergibt sich durch Anwendung von (**) auf jedes Teilintervall $]\alpha,\beta[=]a_j, b_j[$ leicht

$$\sum_{j=1}^{r} |V_F(b_j) - V_F(a_j)| < \varepsilon$$

gemäß der Wahl von δ.

2.5 Satz Für eine Funktion $F : [a,b] \to \mathbb{R}$ sind äquivalent:

(i) F ist absolutstetig.

(ii) F' existiert fast überall, F' ist integrierbar, und für alle $x \in [a,b]$ gilt
$$F(x) - F(a) = \int_a^x F' \, d\lambda .$$

Beweis: Zuerst eine Vorbemerkung: Sollte F' überall existieren, so ist F' als punktweiser Limes stetiger Funktionen Borel-meßbar:
$$F'(x) = \lim_{n\to\infty} n \left(F(x+\tfrac{1}{n}) - F(x)\right).$$
Existiert F' fast überall, ist daher die Funktion
$$\varphi(x) = \begin{cases} F'(x) & \text{falls existent} \\ 0 & \text{sonst} \end{cases}$$
Lebesgue-meßbar. Es ist klar, daß mit dem Integral $\int^x F' \, d\lambda$ über die nur lückenhaft definierte Funktion F' eigentlich $\int^x \varphi \, d\lambda$ gemeint ist – vgl. die Bemerkung nach Lemma II. 3.3 auf S. 84. Nun zu

(ii) $\Rightarrow$ (i): Setzt man $G(x) = \int_a^x F' \, d\lambda$, so ist G absolutstetig nach Beispiel 2.3. d). Daher ist $F = G + F(a)$ ebenfalls absolutstetig.

(i) $\Rightarrow$ (ii): O.E. dürfen wir folgende vereinfachende zusätzliche Annahmen machen:

– F ist monoton wachsend (wg. Lemma 2.4 c))

– F ist streng monoton wachsend und bildet daher $[a,b]$ bijektiv auf $[F(a), F(b)]$ ab (sonst gehe vom monotonen F zu $\tilde{F}(x) = x + F(x)$ über).

Die Umkehrfunktion $G : [F(a), F(b)] \to [a,b]$ ist dann ebenfalls stetig, also Borel-meßbar. Es folgt, daß F Borelmengen auf Borelmengen abbildet

(denn $F(A) = G^{-1}(A)$ für $A \subset [a,b]$). Daher ist die Abbildung
$$\mu : \mathrm{Bor}([a,b]) \to [0,\infty]$$
$$\mu(A) = \lambda(F(A)) \quad (= (\lambda \circ G^{-1})(A))$$
wohldefiniert, und μ ist ein Maß (es handelt sich um das Bildmaß (Def. II. 5.1) des Lebesguemaßes auf $[F(a), F(b)]$ unter G).

Wir behaupten nun $\mu \ll \lambda$.

⌈ Sei $A \in \mathrm{Bor}([a,b])$ mit $\lambda(A) = 0$. O.B.d.A. ist $A \subset \,]a,b[$.

Zu $\varepsilon > 0$ wähle $\delta > 0$ gemäß Def. 2.2 b). Anschließend wähle eine offene Menge $U \subset \,]a,b[$ mit
$$A \subset U \quad \text{und} \quad \lambda(U) < \delta ;$$
die Regularität des Lebesguemaßes macht das möglich (I.4.5).

U besteht aus den paarweise disjunkten offenen Intervallen
$$]a_1, b_1[, \ ]a_2, b_2[, \ \dots .$$
Es folgt daher $\sum_{i=1}^{\infty} (b_i - a_i) < \delta$, so daß für alle N nach Wahl von δ
$$\sum_{i=1}^{N} (F(b_i) - F(a_i)) \le \varepsilon$$
gilt und deshalb auch
$$(*) \qquad \sum_{i=1}^{\infty} (F(b_i) - F(a_i)) \le \varepsilon .$$
Nun ist $F(A) \subset F(U) = \bigcup_{i=1}^{\infty} \,]F(a_i), F(b_i)[$, und $(*)$ liefert jetzt
$$\mu(A) = \lambda(F(A)) \le \lambda(F(U)) = \sum_{i=1}^{\infty} (F(b_i) - F(a_i)) \le \varepsilon .$$
Es folgt $\mu(A) = 0$. ⌋

Nach dem Satz von Radon-Nikodym existiert eine integrierbare Funktion $h: [a,b] \to \mathbb{R}$ mit $(A \in \mathrm{Bor}([a,b]))$

$$\mu(A) = \int_A h \, d\lambda$$

Speziell folgt für $A = [a,x]$

$$\mu(A) = \lambda([F(a), F(x)]) = F(x) - F(a) = \int_a^x h \, d\lambda$$

Satz 2.1 zeigt nunmehr, daß $F' = h$ f.ü. gelten muß, was zu zeigen war.

Eine andere Art, die Äquivalenz aus Satz 2.5 auszudrücken, ist im folgenden Koroller enthalten.

2.6 Koroller Für eine monoton wachsende Funktion $F: [a,b] \to \mathbb{R}$ sind äquivalent:

(i) F ist absolutstetig.

(ii) Das Lebesgue-Stieltjes-Maß μ_F (Satz I.4.9) ist absolutstetig bzgl. λ.

Formal müßten wir erst F durch $F(x) = \begin{cases} F(a) & x < a \\ F(b) & x > b \end{cases}$ auf $\mathbb{R}$ fortsetzen, bevor wir I.4.9 benutzen!

Beweis: μ_F stimmt mit dem im obigen Beweis konstruierten Maß μ überein, denn nach I.4.9 ist μ_F das einzige Maß, das einem Intervall $[\alpha, \beta]$ den Wert $F(\beta) - F(\alpha)$ zuordnet. Daher folgt (ii) ⇒ (i) wie oben auf dieser Seite.

Ist umgekehrt F absolutstetig, so folgt nach 2.5

$$F(x) - F(a) = \int_a^x F' \, d\lambda ,$$

woraus sich wieder mit der Eindeutigkeitsaussage aus I.4.9 ergibt, daß μ_F bzgl. λ die Dichte F' besitzt, d.h. $\mu_F \ll \lambda$.

Noch einige Bemerkungen:

1) Es gibt eine überall differenzierbare Funktion F auf $[0,1]$, deren Ableitung beschränkt, aber nicht Riemann-integrierbar ist. (Volterra 1881, siehe Benedetto, S. 20) Der „klassische" Hauptsatz ist auf so eine Funktion natürlich nicht anwendbar. Hingegen kann man zeigen:
Falls F überall differenzierbar ist und F' Lebesgue-integrierbar ist, so muß F absolutstetig sein und daher $F(x) - F(a) = \int_a^x F' \, d\lambda$ gelten (siehe Rudin, 2. Aufl., S. 179).

2) Ein berühmtes Resultat von Lebesgue besagt, daß für jede Funktion F beschränkter Variation F' fast überall existiert; jedoch ist F i.a. nicht mehr Integral seiner Ableitung. (Die Cantor-Funktion F aus Bsp. 2.2 c) ist monoton wachsend und stetig*), und es gilt $F' = 0$ f.ü.!) Den Beweis (und viele mehr) findet man in den Büchern von z.B. Rudin, Royden, Hewitt-Stromberg, Benedetto ….

*) aber nicht konstant

IV. 3 Die Substitutionsregel für Integrale im $\mathbb{R}^k$

Diese lautet:

3.1 Satz Es sei $U \subset \mathbb{R}^k$ offen, $g : U \to \mathbb{R}^k$ stetig differenzierbar und injektiv; ferner sei die Jacobideterminante J_g überall $\neq 0$. Dann gilt für eine λ^k-integrierbare Funktion $f : g(U) \to \mathbb{R}$

$$(*) \qquad \int_{g(U)} f \, d\lambda^k = \int_U f \circ g \cdot |J_g| \, d\lambda^k .$$

Beweis: Zunächst sollte eine Ähnlichkeit zwischen dieser Substitutionsregel und dem Transformationssatz für Bildmaße II.5.3 ins Auge fallen. In der Tat werden wir 3.1 darauf zurückführen.

Wir setzen $V = g(U)$ und definieren $h : V \to U$ als Umkehrfunktion von g. Nach dem Satz über die inverse Funktion der Analysis III ist V ebenfalls offen und h stetig differenzierbar. Setzen wir noch $\varphi = f \circ g$, so ist (*) äquivalent zu

$$\int_V \varphi \circ h \, d\lambda^k = \int_{h(V)} \varphi \cdot |J_g| \, d\lambda^k .$$

Daher folgt (*) aus II.5.3, wenn wir zeigen können, daß das Bildmaß von λ^k unter h das λ^k-absolutstetige Maß auf U mit der Dichte $|J_g|$ ist.[*)]

Sei μ dieses Bildmaß, m.a.W., für $A \in \mathrm{Bor}(U)$ ist

$$\mu(A) = \lambda^k(h^{-1}(A)) = \lambda^k(g(A)) .$$

(Da h stetig und g bijektiv ist, bildet g Borelmengen auf Borelmengen ab.)

[*)] In Bsp. II.5.2.c) ist das für lineare h bewiesen. Da differenzierbare Funktionen durch lineare approximiert werden, erscheint diese Aussage natürlich.

Wir zeigen nun nacheinander

a) $\mu \ll \lambda^k$, d.h. g bildet λ^k-Nullmengen auf λ^k-Nullmengen ab.

b) $(D\mu)(x) = |J_g(x)|$ für $x \in U$; Korollar III.4.8 zeigt dann, daß $|J_g|$ die gewünschte Radon-Nikodym-Dichte ist. ($D\mu$ wurde in III.4.1 definiert)[*)]

Zu a): Es bezeichne $(\partial g)(x)$ die Ableitung von g an der Stelle x. ($(\partial g)(x)$ ist also eine $k \times k$-Matrix.) Wir werden zuerst zusätzlich annehmen, daß $\{(\partial g)(x) : x \in U\} \subset \mathbb{R}^{k^2}$ beschränkt und U eine Kugel ist. Dann existiert $C > 0$ mit $\|g(x) - g(y)\| \le C \cdot \|x - y\| \quad \forall x, y \in U$ (Analysis III!).

Sei nun $\varepsilon > 0$, und sei $N \subset U$ eine Nullmenge. Wir werden

$$\lambda^k(g(N)) \le \varepsilon$$

zeigen. Wegen der Regularität des Lebesguemaßes (I.4.5) ist nur zu zeigen:

$$\tilde{K} \subset g(N) \text{ kompakt} \Rightarrow \lambda^k(\tilde{K}) \le \varepsilon .$$

Nun ist $K = g^{-1}(\tilde{K}) = h(\tilde{K}) \subset N$ als stetiges Bild eines Kompaktums kompakt, und es ist natürlich $\lambda^k(K) = 0$. Nochmalige Anwendung der Regularität zeigt: Es existiert offenes O, $K \subset O \subset U$ mit $\lambda^k(O) \le \delta := \frac{\varepsilon}{(3C)^k}$. O kann als abzählbare Vereinigung offener Kugeln geschrieben werden, etwa $O = \bigcup_{i=1}^{\infty} B(x_i, r_i)$ (Beweis wie auf S. 14 oben). Da K kompakt ist, gilt schon

$$K \subset \bigcup_{i=1}^{n} B(x_i, r_i)$$

für ein geeignetes $n \in \mathbb{N}$. Auf diese endliche Familie wende nun Lemma III.4.3 an. Mit I wie dort folgt

$$\lambda^k(\tilde{K}) = \lambda^k(g(K))$$

$$\le \lambda^k\Big(g\big(\bigcup_I B(x_i, 3r_i)\big)\Big)$$

[*)] siehe S. 177 !

$$\le \lambda^k \Big(\bigcup_I B(g(x_i), 3r_i \cdot C) \Big) \qquad \text{(nach Wahl von } C\text{)}$$

$$\le \sum_I \lambda^k \big(B(g(x_i), 3r_i C) \big)$$

$$\le (3C)^k \sum_I \lambda^k \big(B(g(x_i), r_i) \big) \qquad (\mathrm{I}.4.4)$$

$$= (3C)^k \sum_I \lambda^k \big(B(x_i, r_i) \big) \qquad \text{(Translationsinvarianz)}$$

$$= (3C)^k \lambda^k \Big(\bigcup_I B(x_i, r_i) \Big) \qquad (\mathrm{III}.4.3\,a))$$

$$\le (3C)^k \lambda^k (D)$$

$$\le (3C)^k \delta$$

$$= \varepsilon .$$

(Die Abschätzung fällt deswegen etwas umständlich aus, weil die $B(x_i, r_i)$ nicht disjunkt sind; im Fall $k = 1$ läßt sich das relativ leicht erreichen, für $k \geq 2$ jedoch nicht so unmittelbar. Daher hier der Rückgriff auf III.4.3.)

Es bleibt, sich von den einschränkenden Voraussetzungen an U und g zu befreien. Dazu schreiben wir – wie gehabt – eine beliebige offene Menge U als abzählbare Vereinigung von Kugeln U_n, für die noch $\overline{U_n} \subset U$ ist (Beweis 2.)

Ist $N \subset U$ eine Nullmenge, so sind nach dem 1. Beweisteil alle $g(U_n \cap N)$ Nullmengen, denn als stetige Abbildung ist ∂g auf dem Kompaktum $\overline{U_n}$ beschränkt. Daher ist

$$g(N) = \bigcup_n g(U_n \cap N)$$

eine Nullmenge.

Zu b): Wir machen wieder zuerst eine vereinfachende Annahme, nämlich

$$\partial g(x) = \mathrm{Id} \qquad (= \text{Einheitsmatrix auf } \mathbb{R}^k)$$

für einen Punkt $x \in U$. Nach Definition von D_μ haben wir nun

$$(*) \qquad \lim_{r \to 0} \frac{\lambda^k(g(B(x,r)))}{\lambda^k(B(x,r))} = 1$$

zu zeigen. O.E. dürfen wir $x = 0$ und $g(x) = 0$ annehmen, was den Schreibaufwand etwas reduziert.

Sei $\varepsilon > 0$. Wähle zuerst $\eta > 0$ mit

$$1 - \varepsilon \le (1-\eta)^k \quad , \quad 1 + \varepsilon \geq (1+\eta)^k .$$

Dann wähle aufgrund der Differenzierbarkeit von g $\delta > 0$ mit

$$\|y\| < \delta \implies \|g(y) - y\| \le \eta \cdot \|y\| .$$

Sei $r < \delta$. Setzen wir zur Abkürzung

$$B = B(0,r) , \qquad B_1 = B(0, r(1-\eta)) , \quad B_2 = B(0, r(1+\eta)) .$$

Wir werden

$$(**) \qquad \left| \frac{\lambda^k(g(B))}{\lambda^k(B)} - 1 \right| \le \varepsilon$$

und damit $(*)$ beweisen. Dazu benötigen wir

$$B_1 \subset g(B) \subset B_2 .$$

Die zweite Inklusion folgt leicht aus der Dreiecksungleichung:

$$\|y\| < r \implies \|g(y)\| \le \|g(y) - y\| + \|y\|$$
$$\le (1+\eta) \|y\|$$
$$< (1+\eta) \cdot r .$$

Nun zur ersten Inklusion. Wir setzen

$$E_1 = B_1 \cap g(B) \quad , \quad E_2 = B_1 \setminus g(B) .$$

Dann gelten:

- $E_1 \cap E_2 = \emptyset, \quad E_1 \cup E_2 = B$
- E_1 ist offen (stetig diff'bare Funktionen bilden offene Mengen auf offene Mengen ab, [illegible])
- $E_1 \neq \emptyset \quad (0 \in E_1)$
- E_2 ist offen

⌈ denn $E_2 = B_1 \setminus g(B^-)$, und $g(B^-)$ ist kompakt (also abgeschlossen)

Dann gelte $\|y\| = r$. Ziel: $g(y) \notin B_1$

$$\|g(y)\| = \|y - (g(y) - y)\|$$

$$\geq \|y\| - \|g(y) - y\| \quad \text{(umgekehrte Dreiecksungleichung)}$$

$$\geq r - \eta r$$

$$= r(1-\eta)$$ ⌋

Eine offene Kugel ist jedoch „zusammenhängend", d.h. nie die nicht disjunkte Vereinigung zweier nichtleerer offener Teilmengen.

Daraus folgt $E_2 = \emptyset$, d.h. $B_1 \subset g(B)$.

⌈ Beweisskizze für die Zusammenhangseigenschaft einer Kugel B:

Falls E_1, E_2 offen und disjunkt mit $E_1 \cup E_2 = B$ und $\exists\, y_0 \in E_1,\ y_1 \in E_2$, betrachte zu $0 \leq t \leq 1$ $y_t = t y_1 + (1-t) y_0 \ (\in B)$.

Sei $\tau = \sup\{t : y_t \in E_1\}$. Man führt dann sowohl $y_\tau \in E_1$ als auch $y_\tau \in E_2$ leicht zu einem Widerspruch. ⌋ ⌋

Aus diesen Inklusionen ergibt sich

$$1 - \varepsilon \leq (1-\eta)^k = \frac{\lambda^k(B_1)}{\lambda^k(B)} \leq \frac{\lambda^k(g(B))}{\lambda^k(B)} \leq \frac{\lambda^k(B_2)}{\lambda^k(B)} = (1+\eta)^k \leq [illegible]$$

folglich (**).

Abschließend ist der Fall, daß $(\partial g)(x) =: S$ beliebig ist, zu behandeln. Auf jeden Fall ist in 3.1 vorausgesetzt, daß S invertierbar ist und daher $T := S^{-1}$ existiert. Setzen wir $\tilde{g} = T \circ g$ mit zugehörigem Maß

$$\tilde{\mu}(A) = \lambda^k(\tilde{g}(A)) = \lambda^k(T(g(A)) \underset{\text{I 4.4}}{=} |\det T|\ \lambda^k(g(A)) = |\det T| \cdot \mu(A),$$

so zeigt der erste Beweisteil wegen

$$(\partial\tilde{g})(x) = T \circ (\partial g)(x) = Id$$

$$(D\tilde{\mu})(x) = 1.$$

Da

$$(D\mu)(x) = |\det T|^{-1}\ D\tilde{\mu}(x)$$

$$= |\det T^{-1}|$$

$$= |J_g(x)|,$$

ist jetzt alles gezeigt.

Nachbemerkung zum Beweis.

- Präzisierung zu *), S. 173: $D\mu$ ist eigentlich nur für endliche Maße definiert.

  In die Berechnung von $D\mu(x)$ (wie von $J_g(x)$) gehen jedoch nur die Werte von g in einer Umgebung von x ein. Nach Einschränkung auf eine relativkompakte Umgebung darf man aber hier o.E. g als beschränkt und folglich μ als endlich voraussetzen.

Als **Beispiel** betrachte die Substitution auf Polarkoordinaten:

$$g: \left]0,\infty\right[\times \left]0,2\pi\right[\to \mathbb{R}^2$$

$$g(r,\varphi) = (r\cos\varphi,\ r\sin\varphi)$$

Hier ist $g(U) = \mathbb{R}^2 \setminus \{(x,0) : x \geq 0\}$ und $J_g(r,\varphi) = r$. In diesem Fall kann man die Substitutionsregel in der Form

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

schreiben, wobei wir über Nullmengen mit der angebrachten Großzügigkeit hinwegsehen.

Offensichtlich dürfen die Voraussetzungen von 3.1 auf Nullmengen verletzt werden.

Billingsley und Floret leiten die Substitutionsregel aus wesentlich allgemeineren Prinzipien ab; auch bei Rudin finden sich etwas allgemeinere Voraussetzungen.

Eine nichttriviale (theoretische) Anwendung der Substitutionsregel findet man z.B. in C.A. Rogers' Beweis des Brouwerschen Fixpunktsatzes in American Mathematical Monthly 87 (1980), S. 525 - 527.

V. Ausgewählte Fragen der Maß- und Integrationstheorie

V.1. Der Satz von Daniell - Stone

Wir haben in dem hier vorgestellten Aufbau der Integrationstheorie, ausgehend von einem positiven Maß μ, das Integral auf dem Funktionenraum $\mathcal{L}^1(\mu)$ erklärt. Die entscheidenden Eigenschaften des Integrationsprozesses sind dabei

- die Linearität,
- die Positivität ($f \geq 0 \Rightarrow \int_S f\,d\mu \geq 0$),
- die im Satz von Beppo Levi ausgedrückte „Stetigkeit".

Der Hauptsatz dieses Abschnitts besagt, daß man umgekehrt jedem solchen Funktional auf einem Funktionenraum ein Maß zuordnen kann, so daß das Funktional ein Integral „ist". Dabei geht man von folgenden abstrakten Begriffen aus:

1.1 Definition a) Es sei S eine Menge und $\mathcal{F}$ ein Vektorraum von Funktionen von S nach $\mathbb{R}$. $\mathcal{F}$ heißt

- Funktionenraum, falls $f \in \mathcal{F} \Rightarrow |f| \in \mathcal{F}$
- Stonescher Funktionenraum, falls zusätzlich $f \in \mathcal{F} \Rightarrow \min\{1, f\} \in \mathcal{F}$.

b) Sei $\mathcal{F}$ ein Funktionenraum und $\ell: \mathcal{F} \to \mathbb{R}$ linear. ℓ heißt

- positiv, falls $f \in \mathcal{F},\ f \geq 0 \Rightarrow \ell(f) \geq 0$
- Daniell-stetig, falls zusätzlich

$$f_1, f_2, \ldots \in \mathcal{F},\ f_1 \geq f_2 \geq \cdots \geq 0,\ \inf_n f_n = 0 \Rightarrow \ell(f_n) \xrightarrow[n\to\infty]{} 0.$$

12 Bemerkungen und Beispiele a) $|f|$, $\min\{1,f\}$, $f \geq 0$, $\inf f_n$ etc. sind natürlich punktweise erklärt.

b) Man setzt für zwei Funktionen f und g

$$f \wedge g := \min\{f,g\} \quad \left(= \frac{f+g}{2} - \left|\frac{f-g}{2}\right| \right)$$

$$f \vee g := \max\{f,g\} \quad \left(= \frac{f+g}{2} + \left|\frac{f-g}{2}\right| \right).$$

In einem Funktionenraum gilt also

$$f,g \in \mathcal{F} \Rightarrow f \wedge g \in \mathcal{F}, \quad f \vee g \in \mathcal{F}.$$

Daher ist die Stonesche Bedingung $1 \wedge f \in \mathcal{F}$ automatisch erfüllt, falls sogar $1 \in \mathcal{F}$ gilt.

c) Sei $(S, \mathfrak{A}, \mu)$ ein Maßraum, $\mathcal{F} = \mathcal{L}^1(\mu)$, $\ell(f) = \int_S f \, d\mu$. Dann ist $\mathcal{F}$ ein Stonescher Funktionenraum, und ℓ ist ein positives lineares Daniell-stetiges Funktional. (Letzteres ergibt sich unmittelbar aus dem Satz von Beppo Levi.) Beachte $1 \in \mathcal{F} \Leftrightarrow \mu(S) < \infty$.

d) $\mathcal{F} = \{f : [0,1] \to \mathbb{R} : f \text{ stetig differenzierbar}\}$ bildet keinen Funktionenraum

e) Sei $\mathcal{F} = \{f : [0,1] \to \mathbb{R} : f \text{ ist stetig}, f(0) = 0, \lim\limits_{t \to 0} \frac{f(t)}{t} \text{ existiert}\}$,

$$\ell : \mathcal{F} \to \mathbb{R}, \quad \ell(f) = \lim_{t \to 0} \frac{f(t)}{t} \quad (= f'(0)).$$

$\mathcal{F}$ ist ein Stonescher Funktionenraum, und ℓ ist linear und positiv, jedoch nicht Daniell-stetig: betrachte $f_n(t) = \min\{t, \frac{1}{n}\}$.

f) Sei $f_0(t) = t$ $(t \in [0,1])$. $\mathcal{F} := \{\alpha f_0 : \alpha \in \mathbb{R}\}$ ist ein Funktionenraum, jedoch die Stonesche Bedingung ist verletzt.

(Ein interessanteres Beispiel folgt unter 14.)

13 Theorem (Satz von Daniell-Stone)

Es seien $\mathcal{F}$ ein Stonescher Funktionenraum auf einer Menge S sowie ℓ ein positives lineares Daniell-stetiges Funktional auf $\mathcal{F}$.

Dann existieren eine σ-Algebra $\mathfrak{A} \subset \mathcal{P}(S)$, für die alle $f \in \mathcal{F}$ $\mathfrak{A}$-meßbar sind, und ein Maß $\mu : \mathfrak{A} \to [0,\infty]$ mit $\mathcal{F} \subset \mathcal{L}^1(\mu)$, so daß

$$\ell(f) = \int_S f \, d\mu \qquad \forall f \in \mathcal{F}.$$

Ist $\mathfrak{A}$ minimal gewählt und existieren $f_n \in \mathcal{F}$ mit $\sup_n f_n = 1$, ist μ eindeutig bestimmt. In diesem Fall ist μ σ-endlich.

Beweis: 1. Schritt Für zwei Funktionen $f, g : S \to \mathbb{R}$ definiere das „Intervall"

$$]f,g] = \{(s,t) \in S \times \mathbb{R} : f(s) < t \leq g(s)\}.$$

Es sei $\mathcal{K} = \{]f,g] : f,g \in \mathcal{F}, f \leq g\}$

Dann ist $\mathcal{K}$ $\cap$-stabil

$$]f_1,g_1] \cap]f_2,g_2] =]f_1 \vee f_2, f_1 \vee f_2 \vee (g_1 \wedge g_2)],$$

und die Differenz zweier Intervalle in $\mathcal{K}$ ist als disjunkte Vereinigung von Intervallen darstellbar

$$]f_1,g_1] \setminus]f_2,g_2] =]f_1, f_1 \vee (g_1 \wedge f_2)] \cup]g_1 \wedge (g_2 \vee f_1), g_1]$$

(genau wie bei Intervallen in $\mathbb{R}$!!!). Es folgt, daß der von $\mathcal{K}$ erzeugte Ring („die Figuren") genau aus den endlichen disjunkten Vereinigungen von Intervallen in $\mathcal{K}$ besteht. (Beweis?)

2. Schritt Definiere $\tau : \mathcal{K} \to [0, \infty[$ durch

$$\tau(\,]f,g]\,) = \ell(g-f)$$

τ ist wohldefiniert, $\tau(\emptyset) = 0$ (denn $\emptyset = \,]0,0]$), und τ ist σ-additiv.

⌈ Seien $]f,g]$, $]f_1,g_1]$, $]f_2,g_2]$, … $\in \mathcal{K}$, und sei

$$]f,g] = \bigcup_{i=1}^{\infty} \,]f_i, g_i] \qquad \text{(disjunkte Vereinigung).}$$

Es folgt für alle $s \in S$, $t \in \mathbb{R}$:

$$f(s) < t \le g(s) \iff \text{Es ex. genau ein } i \in \mathbb{N} \text{ mit } f_i(s) < t \le g_i(s)$$

Das heißt

$$]f(s), g(s)] = \bigcup_{i=1}^{\infty} \,]f_i(s), g_i(s)] \qquad \forall s \in S$$

als disjunkte Vereinigung. Da das Lebesguemaß σ-additiv ist, ergibt sich daraus

$$g(s) - f(s) = \sum_{i=1}^{\infty} (g_i(s) - f_i(s)) \qquad \forall s \in S$$

oder

$$g - f = \sum_{i=1}^{\infty} (g_i - f_i).$$

Anders gesagt, konvergiert $\left(g - f - \sum_{i=1}^{n} (g_i - f_i)\right)_{n \in \mathbb{N}}$ monoton fallend gegen 0 (beachte, daß diese Funktionen wirklich in $\mathcal{F}$ liegen!), so daß Linearität und Daniell-Stetigkeit von ℓ

$$\tau(\,]f,g]) = \ell(g-f)$$
$$= \sum_{i=1}^{\infty} \ell(g_i - f_i) = \sum_{i=1}^{\infty} \tau(\,]f_i, g_i])$$

liefern. ⌋

Es folgt nun leicht aus 1., daß die kanonische Fortsetzung von τ auf den von $\mathcal{K}$ erzeugten Ring wohldefiniert und ebenfalls σ-additiv ist. Dieses Prämaß setze mit Carathéodory auf die erzeugte σ-Algebra

$\Sigma := \sigma(\mathcal{K})$ fort. Die Fortsetzung (genauer: „eine solche Fortsetzung" [mangels Eindeutigkeit]) $\tau : \Sigma \to [0, \infty]$ werde ebenfalls mit τ bezeichnet.

3. Schritt Sei

$$\mathcal{R} = \{A \subset S : A \times \,]0,1] \in \Sigma\}$$

sowie

$$\mathfrak{A} = \sigma(\mathcal{R}).$$

Es ist klar, daß $\mathcal{R}$ ein Ring (sogar ein σ-Ring, Def. I.1.1.b)) ist. Ferner setze

$$\mu : \mathcal{R} \to [0, \infty], \quad \mu(A) = \tau(A \times \,]0,1]),$$

so daß μ ein Prämaß ist. Ferner betrachte eine ebenfalls mit μ bezeichnete Fortsetzung auf $\mathfrak{A}$. (Carathéodory!)

4. Schritt Sei $f \in \mathcal{F}$, $f \ge 0$. Setze

$$g_n = (n(f - f \wedge 1)) \wedge 1.$$

Wegen der Stoneschen Eigenschaft von $\mathcal{F}$ liegen die $g_n \in \mathcal{F}$.

Nach Konstruktion gilt

(a) $0 \le g_n \nearrow 1_{\{f > 1\}}$

sowie (b) $]0, c\,g_n] \nearrow \{f > 1\} \times \,]0, c]$ für $c > 0$.

$]0, g_n] \in \mathcal{K}$ zeigt dann $\{f > 1\} \in \mathcal{R}$. Ferner gilt

(c) $$\mu(\{f>1\}) = \tau(\{f>1\} \times \,]0,1])$$
$$\le \tau(\,]0,f]) \qquad (\text{da } \{f>1\} \times \,]0,1] \subset \,]0,f])$$
$$= \ell(f - 0) = \ell(f) < \infty$$

sowie (d) $$\tau(\{f>1\} \times \,]0,c]) = c \cdot \mu(\{f>1\}) \qquad (c > 0)$$

⌈ $$\tau(\{f>1\} \times \,]0,c]) \underset{\substack{(b)\\ \tau\ \sigma\text{-additiv}}}{=} \lim_{n \to \infty} \tau(\,]0, c \cdot g_n])$$
$$= \lim_{n \to \infty} \ell(c\, g_n)$$

$$= c \lim_{n\to\infty} \ell(g_n)$$

$$= c \cdot \lim_{n\to\infty} T(\,]0,g_n]\,)$$

$$= c \cdot T(\{f>1\} \times\,]0,1]\,)$$

$$\overset{(L)}{=} c\ \mu(\{f>1\}).$$

Durch Übergang zu $\frac{f}{r}$ in (c) und (d) und Differenzbildung (hier ist dann (c) wesentlich) erhält man für $\infty \geq R > r > 0$ schließlich

$$(e) \quad T(\{R \geq f > r\} \times\,]0,c]\,) = c \cdot \mu(\{R \geq f > r\})$$

5. Schritt Weisen wir nun nach, daß $\mathfrak{A}$ und μ das Gewünschte leisten. Sei $f \in \mathcal{F}$. Da $f^+ = f \vee 0 \in \mathcal{F}$ ist, folgt aus 4b, daß

$$\{f^+ > r\} = \{\tfrac{f^+}{r} > 1\} \in \mathcal{R} \qquad \forall\, r > 0$$

ist. Daher auch

$$\{f^+ > 0\} = \bigcup_{n=1}^{\infty} \{f^+ > \tfrac{1}{n}\} \in \mathcal{R}$$

und wegen $f^+ \geq 0$ selbstverständlich für $r < 0$

$$\{f^+ > r\} = S \in \mathfrak{A}.$$

Daher ist f^+ $\mathfrak{A}$-meßbar, analog f^-, so daß auch $f = f^+ - f^-$ $\mathfrak{A}$-meßbar ist.

Sei nun $f \in \mathcal{F}$, $f \geq 0$. Wir zeigen

$$\ell(f) = \int_S f \, d\mu.$$

Da f $\mathfrak{A}$-meßbar ist, existiert eine Folge von $\mathfrak{A}$-Treppenfunktionen f_n mit $f_n \nearrow f$. (Leider ist i.a. $f_n \notin \mathcal{F}$.) Wie der Beweis von II.18 zeigt, können die f_n in der Form

$$f_n = \sum_{i=1}^{N_n} r_{in}\, 1_{\{r_{i+1,n} \geq f > r_{in}\}} \qquad (\text{mit } r_{N_n+1,n} = \infty,\ r_{in} > 0\,)$$

angesetzt werden. Betrachte

$$A_{in} := \{r_{i+1,n} \geq f > r_{in}\} \in \mathcal{R}$$

und $\Sigma \ni\,]0,f_n] \nearrow\,]0,f]$. Daher

$$\int_S f\, d\mu = \lim_{n\to\infty} \int_S f_n \, d\mu \qquad (\text{Beppo Levi})$$

$$= \lim_{n\to\infty} \sum_i r_{in}\, \mu(A_{in}) \qquad (\text{Def. } \textstyle\int_S \cdot\, d\mu)$$

$$= \lim_{n\to\infty} \sum_i T(A_{in} \times\,]0,r_{in}]\,) \qquad (4.(e))$$

$$= \lim_{n\to\infty} T(\bigcup_i A_{in} \times\,]0,r_{in}]\,) \qquad (\text{da die } A_{in} \text{ paarw. disj. sind})$$

$$= \lim_{n\to\infty} T(\,]0,f_n]\,) \qquad (\text{Def. }]0,f_n])$$

$$= T(\,]0,f]\,) \qquad (\,]0,f_n] \nearrow\,]0,f]\,)$$

$$= \ell(f) \qquad (\text{Def. } T)$$

Daher $\mathcal{F} \subset \mathcal{L}^1(\mu)$ $\quad (f \in \mathcal{F} \Rightarrow |f| \in \mathcal{F}$ und $|f| \geq 0 \underset{\text{s.o.}}{\Rightarrow} \int_S |f|\, d\mu = \ell(|f|) < \infty)$ und

$$\int_S f\, d\mu = \int_S (f^+ - f^-)\, d\mu = \int_S f^+ d\mu - \int_S f^- d\mu$$

$$\underset{f^+,\,f^- \in \mathcal{F}}{=} \ell(f^+) - \ell(f^-) \underset{\ell \text{ linear.}}{=} \ell(f).$$

6. Schritt: Eindeutigkeit: Sei

$$\mathcal{E} = \{\, \{f>1\} : f \in \mathcal{F},\ f \geq 0\,\},$$

$$\mathfrak{A}_{\mathcal{F}} = \sigma(\mathcal{E}).$$

Ist $\mathcal{D}$ eine σ-Algebra auf S, bezüglich derer alle $f \in \mathcal{F}$ meßbar sind, so gilt gewiß $\mathcal{E} \subset \mathcal{D}$ und deshalb $\mathfrak{A}_{\mathcal{F}} \subset \mathcal{D}$. Umgekehrt ist im 5. Schritt gezeigt worden, daß alle $f \in \mathcal{F}$ $\mathfrak{A}_{\mathcal{F}}$-meßbar sind. Daher ist $\mathfrak{A}_{\mathcal{F}}$ die in der Eindeutigkeitsaussage beschriebene minimale σ-Algebra.
Sei neben dem oben konstruierten μ (genauer: $\mu|_{\mathfrak{A}_{\mathcal{F}}}$, beachte $\mathfrak{A}_{\mathcal{F}} \subset \mathfrak{A}$, wo $\mathfrak{A}$ wie oben) ein weiteres Maß $\nu : \mathfrak{A}_{\mathcal{F}} \to [0, \infty]$ mit

$$\int_S f \, d\nu = \ell(f) \qquad \forall\, f \in \mathcal{F}$$

vorgelegt. Zu $f \in \mathcal{F}$, $f \geq 0$, definiere g_n wie unter 4. Es folgt dann

$$\begin{aligned} \mu(\{f>1\}) &= \int_S 1_{\{f>1\}} \, d\mu \\ &= \lim \int_S g_n \, d\mu \\ &= \lim \ell(g_n) \\ &= \lim \int_S g_n \, d\nu \\ &= \int_S 1_{\{f>1\}} \, d\nu \\ &= \nu(\{f>1\}), \end{aligned}$$

d.h.
$$\mu|_{\mathcal{E}} = \nu|_{\mathcal{E}}.$$

Um $\mu = \nu$ zu erhalten, verwenden wir den Eindeutigkeitssatz I.3.11: $\mathcal{E}$ ist $\cap$-stabil, da $\{f>1\} \cap \{g>1\} = \{f \wedge g > 1\}$; ist (f_n) eine Folge in $\mathcal{F}$ mit $\sup f_n = 1$ (die wir nach evtl. Übergang zu $(g_n) = (f_1 \vee \dots \vee f_n)$ als monoton wachsend annehmen dürfen), so gilt
$$S = \bigcup_{n=1}^{\infty} \{2f_n > 1\},$$
und auf diesen Mengen sind ja μ und ν endlich. (Das zeigt gleichzeitig die σ-Endlichkeit von μ in diesem Fall.)

Eine wichtige Anwendung werden wir mit dem Rieszschen Darstellungssatz im nächsten Abschnitt kennenlernen (2.12).
Grundsätzlich ist es auch möglich, die gesamte Integrationstheorie – ohne vorherigen Aufbau der Maßtheorie – direkt aus den Daniell-stetigen Funktionalen zu gewinnen; vgl. dazu die zum Daniellschen bzw. Bourbakischen Zugang in der Einleitung genannten Bücher.
Abschließend eine Bemerkung zu den im Satz von Daniell-Stone gemachten Voraussetzungen:
Daß ohne die Bedingung $1 = \sup_n f_n$ keine Eindeutigkeit für das darstellende Maß erwartet werden kann, zeigt das triviale Beispiel $S = [0,1]$, $\mathcal{F} = \{0\}$, $\ell = 0$. (Ein anderer Beweis von 1.3 zeigt jedoch, daß unter zusätzlichen „Regularitätsforderungen" stets Eindeutigkeit (und Existenz!) erzielt werden kann; vgl. etwa Bauer, § 39.)
Die Daniell-Stetigkeit ist im Hinblick auf den Satz von Beppo Levi gewiß notwendig. Betrachtet man in der Situation des Beispiels 1.2 f) das Daniell-stetige positive lineare Funktional $\ell(\alpha f_0) = \alpha$, so gibt es, obwohl $\mathcal{F}$ nicht Stonesch ist, ein darstellendes Maß für ℓ, nämlich z.B. das Dirac-Maß δ_1. (Auch $\mu = 2\lambda$ würde es tun!!!) Das folgende Beispiel zeigt jedoch, daß die Stonesche Eigenschaft für die Existenz eines darstellenden Maßes μ wesentlich ist.

1.4 Beispiel Dieses Beispiel fußt auf dem Begriff der Baireschen Kategorie. Sei T ein metrischer (oder auch nur topologischer) Raum. Eine Teilmenge M heißt

- nirgends dicht, wenn $\mathrm{int}(M^-) = \emptyset$ ist [*],
- von 1. Kategorie, wenn es nirgends dichte $M_1, M_2, \ldots$ mit $M = \bigcup_{n=1}^{\infty} M_n$ gibt,
- von 2. Kategorie, wenn M nicht von 1. Kategorie ist.

(? $\mathbb{Q}$ ist $\mathbb{Q}$ von 1. Kategorie in $\mathbb{R}$ Baires ?)

Der Bairesche Kategoriensatz besagt, daß vollständige metrische Räume von 2. Kategorie sind.

Nun zum gewünschten Beispiel. Es sei $S = [1,2]$ und

$$F = \{ f : [1,2] \to \mathbb{R} : \text{ Es ex. } M \subset [1,2] \text{ von 1. Kategorie und } \alpha \in \mathbb{R} \text{ mit } f(t) = \alpha t \quad \forall t \notin M \}.$$

Da die Vereinigung von 2 (sogar von abzählbar vielen) Mengen 1. Kategorie wieder von 1. Kategorie ist, ist F ein Funktionenraum.

Bemerke nun, daß das α aus der Definition von F durch $f \in F$ eindeutig bestimmt ist: Sei nämlich

$$f(t) = \alpha_1 t \quad \text{für } t \notin M_1,$$
$$f(t) = \alpha_2 t \quad \text{für } t \notin M_2.$$

Nach dem Baireschen Kategoriensatz ist $M_1 \cup M_2 \neq [1,2]$, daher existiert t_0 mit $\alpha_1 t_0 = f(t_0) = \alpha_2 t_0$; folglich $\alpha_1 = \alpha_2$.

Also ist die Definition $\ell(f) = \alpha$ sinnvoll. Es ist nicht schwer, ℓ als linear und positiv zu erkennen. ℓ ist auch Daniell-stetig: Gelte $f_n \downarrow 0$, wo etwa $f_n(t) = \alpha_n t$ für $t \notin M_n$ ist. Da $M = \bigcup_n M_n$ von 1. Kategorie, $[1,2]$ aber nach Baire von 2. Kategorie ist, existiert $t_0 \in [1,2] \setminus M$. Für so ein t_0 gilt $\alpha_n t_0 \to 0$, folglich $\ell(f_n) = \alpha_n \to 0$.

Wäre ℓ als $\int_S \cdot \, d\mu$ darstellbar, so müßte wegen

[*] $\mathrm{int}(A)$ bezeichnet das Innere einer Menge A

$$\mu(S) = \int_S 1 \, d\mu < \int_S t \, d\mu = \ell(\mathrm{id}) < \infty$$

μ endlich sein. Da für alle $M \subset [1,2]$ von 1. Kategorie $1_M \in F$ und $\ell(1_M) = 0$ ist, müßte für μ

$$(*) \qquad \mu(M) = 0 \qquad \forall M \subset [1,2] \text{ von 1. Kategorie}$$

gelten. Sei $F(t) = \mu([1,t])$ die Verteilungsfunktion von μ. Ist t_0 ein Punkt, wo F nicht stetig ist, so gilt $\mu(\{t_0\}) \neq 0$ (vgl. S. 51) im Widerspruch zu $(*)$, da $\{t_0\}$ nirgends dicht ist.

Ist jedoch F stetig, und nicht konstant, so kann man mit einer Cantor-artigen Konstruktion eine kompakte nirgends dichte Teilmenge M mit $\mu(M) > 0$ gewinnen (Widerspruch zu $(*)$!); die Einzelheiten bleiben der Fantasie der Leser vorbehalten! Resultat: $F = \text{const.}$, also $\mu = 0$, aber $\mu = 0$ stellt ℓ gar nicht dar: Widerspruch!

I 2 Maße auf metrischen Räumen

In diesem Abschnitt wird der Zusammenhang von Maß und Topologie angesprochen; für $\mathbb{R}^k$ war ein solcher Zusammenhang z. B. in den Sätzen I.4.6, II.1.9, II.4.9 aufgestellt worden.

Es soll hier nicht größtmögliche Allgemeinheit angestrebt werden. Insbesondere soll auf die Untersuchung allgemeiner lokalkompakter topologischer Räume zugunsten lokalkompakter metrischer Räume, die wir zusätzlich als σ-kompakt voraussetzen werden, verzichtet werden. Viele der hier vorgestellten Sätze behalten jedoch – evtl. leicht modifiziert – ihre Gültigkeit in allgemeineren Rahmen. (Manche jedoch nicht!)

Der zweite Teil dieses Abschnitts behandelt eine in der Wahrscheinlichkeitstheorie bedeutende Klasse metrischer Räume (die übrigens kein Analogon unter den allgemeinen topologischen Räumen hat).

Wir beginnen mit einigen topologischen Bezeichnungen und Vorüberlegungen.

2.1 Definition Es sei S [präziser: (S,d)] ein metrischer Raum.

a) S heißt lokalkompakt, wenn jeder Punkt $s \in S$ eine kompakte Umgebung (d.h. eine kompakte Teilmenge U mit $s \in \mathrm{int}(U)$) besitzt.

b) S heißt σ-kompakt, falls eine Folge kompakter Teilmengen (K_n) mit $S = \bigcup_{n=1}^{\infty} K_n$ existiert.

c) $C^b(S) = \{ f : S \to \mathbb{R} : f \text{ stetig und beschränkt} \}$

$\mathcal{K}(S) = \{ f : S \to \mathbb{R} : f \text{ stetig und } \mathrm{supp}(f) := \{f \neq 0\}^- \text{ kompakt} \}$

(Man sagt, die $f \in \mathcal{K}(S)$ haben kompakten Träger.)

Das kanonische Beispiel eines lokalkompakten σ-kompakten Raums ist $\mathbb{R}^k$.

$\mathbb{Q}$ ist σ-kompakt, aber nicht lokalkompakt.

Sei S eine Menge, versehen mit der diskreten Metrik

$$d(s,t) = 1 \quad \text{falls } s \neq t, \quad d(s,s) = 0.$$

(S,d) ist dann stets lokalkompakt und genau dann σ-kompakt, wenn S höchstens abzählbar ist.

Hier noch ein interessantes Beispiel für einen kompakten metrischen Raum:

Sei W die Menge aller Wahrscheinlichkeitsmaße auf $\mathrm{Bor}([0,1])$.

Für μ und $n \geq 0$ sei $\quad p_n(\mu) = \int_0^1 t^n \, d\mu(t)$, und es sei

ein spezielles Maß

$$d(\mu_1, \mu_2) = \sum_{n=0}^{\infty} 2^{-n} \frac{|p_n(\mu_1 - \mu_2)|}{1 + |p_n(\mu_1 - \mu_2)|}$$

für $\mu_1, \mu_2 \in W$. Man kann zeigen, daß d eine Metrik ist und (W,d) kompakt ist. (Der Beweis kann leicht mit funktionalanalytischen Methoden geführt werden.) Übrigens gilt $d(\mu_n, \mu) \xrightarrow[n\to\infty]{} 0$ (also $\mu_n \to \mu$) genau dann, wenn

$$\int_0^1 f \, d\mu_n \to \int_0^1 f \, d\mu \quad \forall f \in C[0,1] \quad \text{gilt.}$$

2.2 Lemma (Lemma von Urysohn)

Sei (S,d) ein metrischer Raum, $A, B \subset S$ abgeschlossen und disjunkt. Dann existiert $f \in C^b(S)$ mit $\{f = 0\} = A$, $\{f = 1\} = B$.

(Es kann sogar $0 \leq f \leq 1$ erreicht werden.)

Bew.: Setze

$$f(s) = \frac{d(s,A)}{d(s,A) + d(s,B)}.$$

($d(s,A) = \inf_{a \in A} d(s,a)$ definiert eine stetige Funktion!)

2.3 Lemma Sei (S,d) ein lokalkompakter metrischer Raum. Sei $K \subset U$ mit K kompakt, U offen. Dann existiert offenes O mit $K \subset O \subset U$ und O^- kompakt.

Bew.: Zu jedem $s \in K$ existiert offenes U_s, für das $s \in U_s$ und U_s^- kompakt gelten, da S lokalkompakt ist. O.E. ist $U_s \subset U$ (sonst gehe zu $U_s \cap U$ über). Natürlich überdecken die U_s K:

$$K \subset \bigcup_{s \in K} U_s$$

Also reicht bereits eine endliche Teilüberdeckung

$$K \subset \bigcup_{i=1}^{n} U_{s_i} \cdots O ,$$

und $O \subset U$, $\quad O^- = \bigcup_{i=1}^{n} U^-_{s_i}$ ist als endliche Vereinigung kompakter Mengen kompakt.

Nun zurück zur Maßtheorie.

<u>2.4 Definition</u> (S,d) sei ein metrischer Raum.

a) Die von den offenen Mengen erzeugte σ-Algebra heißt die Borel-σ-Algebra $Bor(S)$.

b) Ein Maß $\mu : Bor(S) \to [0,\infty]$ heißt Borelmaß auf S.

c) Ein Borelmaß μ heißt regulär, falls

(i) $\mu(C) < \infty \quad \forall\, C \subset S,\ C$ kompakt

(ii) $\mu(A) = \inf\{\mu(O) : A \subset O,\ O \text{ offen}\}$
$= \sup\{\mu(C) : C \subset A,\ C \text{ kompakt}\} \quad \forall A \in Bor(S)$

Eine wichtige Warnung an alle Bücherleser: Diese Begriffe werden in der Literatur sehr uneinheitlich verwandt (z.B. haben Behrends und Rudin verschiedene Regularitätsbegriffe), im Fall kompakter metrischer Räume fallen jedoch alle in der Literatur auftauchenden Versionen zusammen.

In Def. I.1.8 wurde für $S \subset \mathbb{R}^k$ schon einmal die Borel-σ-Algebra $Bor(S)$ definiert. Lemma I.1.9 (mit $\mathcal{E} = \{O \subset \mathbb{R}^k : O \text{ offen}\}$) zeigt aber, daß die alte Definition mit der neuen übereinstimmt.

Natürlich werden die Borelmengen auch von den abgeschlossenen Mengen erzeugt, das System der kompakten Mengen erzeugt jedoch i.a. eine kleinere σ-Algebra: Sei etwa $S = \mathbb{R}$, d = diskrete Metrik, so daß $Bor(S) = P(S)$. Da jedoch genau die endlichen Mengen kompakt sind, erzeugen diese die σ-Algebra der abzählbaren und coabzählbaren Teilmengen (Bsp. I.1.2.e)). Dagegen gilt:

<u>2.5 Lemma</u> Ist S σ-kompakt, so wird $Bor(S)$ von den kompakten Teilmengen erzeugt.

Beweis: Schreibe $S = \bigcup K_n$ mit kompakten K_n. Jede abgeschlossene $A \subset S$ ist daher abzählbare Vereinigung kompakter Mengen: $\bigcup_n (A \cap K_n) = A$, was die Behauptung zeigt.

<u>2.6 Satz</u> Sei S ein metrischer Raum. Dann ist $Bor(S)$ die kleinste σ-Algebra, für die alle $f \in C^b(S)$ meßbar sind. Ist S zusätzlich lokalkompakt und σ-kompakt, so kann $C^b(S)$ durch $\mathcal{K}(S)$ ersetzt werden.

Beweis: Zunächst einmal ist klar, daß eine stetige Abbildung $f : S_1 \to S_2$ zwischen metrischen Räumen $Bor(S_1)$-$Bor(S_2)$-meßbar ist. (Folgt sofort aus I.1.2 a), da $f^{-1}(O)$ für offene $O \subset S_2$ offen, also borelsch ist.) Damit ist $Bor(S)$ eine σ-Algebra mit der im Satz genannten Eigenschaft. Sei $\mathfrak{A}$ eine weitere. Um $Bor(S) \subset \mathfrak{A}$ zu zeigen, reicht es,

$A \in \mathfrak{A}$ für alle abgeschlossenen A zu zeigen. In der Tat: wähle $f \in C^b(S)$ mit $A = \{f = 0\}$ nach Lemma 2.2. Ist f $\mathfrak{A}$-meßbar, folgt sofort $A \in \mathfrak{A}$.
Zum Zusatz: Hier ist wegen 2.5 $C \in \mathfrak{A}$ für alle kompakten C zu zeigen. Wende 2.3 mit $U = S$ an, erhalte O, und wende dann 2.2 an, um stetiges f mit $\{f = 0\} = O^c$, $\{f = 1\} = C$ zu erhalten (beachte $C \subset O \Rightarrow C \cap O^c = \emptyset$). Wegen $\overline{\{f \neq 0\}} = \overline{O}$ ist $f \in \mathcal{K}(S)$, und die $\mathfrak{A}$-Meßbarkeit von f erzwingt $C \in \mathfrak{A}$.

2.7 Satz Sei S ein lokalkompakter σ-kompakter metrischer Raum. Dann ist jedes Borelmaß, das auf kompakten Mengen endlich ist, regulär.

Beweis: wörtlich wie bei I 4.6 !

2.8 Satz (Satz von Lusin)
Sei S ein kompakter metrischer Raum, μ ein endliches Borelmaß auf S, $\varepsilon > 0$, $f: S \to \mathbb{R}$ eine meßbare Funktion. Dann existiert kompaktes $K \subset S$ mit $\mu(S \setminus K) \leq \varepsilon$, so daß die Restriktion $f|_K : K \to \mathbb{R}$ eine stetige Funktion auf K ist.

Beweis: wie bei II.1.9 ! (Beachte, daß μ nach 2.7 regulär ist.)

2.9 Satz Sei S lokalkompakt und μ ein reguläres Borelmaß auf S. Für $1 \leq p < \infty$ ist dann $\mathcal{K}(S)$ dicht in $L^p(\mu)$ (bzgl. $\|\cdot\|_p$).

Beweis: Aus II.4.7 (Dichtheit der Treppenfunktionen in $L^p(\mu)$) ergibt sich, daß nur folgende Aussage gezeigt zu werden braucht:
$A \in \mathrm{Bor}(S)$, $\mu(A) < \infty$ $(\Leftrightarrow 1_A \in L^p(\mu))$, $\varepsilon > 0$ $\Rightarrow \exists f \in \mathcal{K}(S)$ $\|f - 1_A\|_{L^p} \leq \varepsilon$.

Beweis hierfür: μ ist regulär nach Voraussetzung. Wähle also kompaktes $C \subset A$ und offenes $U \supset A$ mit $\mu(U \setminus C) \leq \varepsilon^p$.
Wähle O gemäß 2.3 und stetiges f mit $\{f = 1\} = C$ und $\{f = 0\} = O^c$ (2.2), so daß also $f \in \mathcal{K}(S)$ ist.
Da $|f - 1_A| \leq 1_{U \setminus C}$ nach Konstruktion,
folgt $\|f - 1_A\|_{L^p} \leq \mu(U \setminus C)^{1/p} \leq \varepsilon$.

Diese Aussage war im Beweis von III 4.7 benötigt worden; dort war auch ein Beweis für $S = \mathbb{R}^k$, $\mu = \lambda^k$ skizziert worden, der auf der Gruppenstruktur von $\mathbb{R}^k$ basiert.
Als nächstes wollen wir den Satz von Daniell-Stone auf den Stoneschen Funktionenraum $\mathcal{K}(S)$ anwenden. Es wird sich herausstellen, daß in diesem Fall jedes positive lineare Funktional Daniell-stetig ist. Das folgt aus dem folgenden Satz von Dini:

2.10 Satz S sei kompakt, $f_1 \geq f_2 \geq \cdots \geq 0$, $f_n \in C(S)$,[*] und (f_n) konvergiere punktweise gegen 0. Dann ist die Konvergenz bereits gleichmäßig.

[*] $C(S) = \{f : S \to \mathbb{R} : f \text{ stetig}\}$. Bemerke $C(S) = C^b(S)$ für kompakte S.

Beweis: Für die offenen Mengen $\{f_n < \varepsilon\}$ ($\varepsilon > 0$ beliebig vorgegeben) gilt $\{f_n < \varepsilon\} \subset \{f_{n+1} < \varepsilon\}$ wegen der Monotonie und $\bigcup_{n\in\mathbb{N}} \{f_n < \varepsilon\} = S$ wegen der punktweisen Konvergenz. Da S kompakt ist, ist $S = \{f_N < \varepsilon\}$ für passendes $N \in \mathbb{N}$, folglich

$$f_n(s) < \varepsilon \qquad \forall s \in S, \ \forall n \geq N.$$

Das war zu zeigen.

<u>2.11 Korollar</u> Jedes positive lineare Funktional ℓ auf $\mathcal{K}(S)$ (S ein metrischer Raum) ist Daniell-stetig.

Beweis: Gelte $f_n \downarrow 0$. Dann gilt auch $\sqrt{f_n} \downarrow 0$, und nach dem Satz von Dini ist die Konvergenz gleichmäßig auf der kompakten Menge $S_1 := \operatorname{supp}(f_1) = \{f_1 \neq 0\}^-$. Da alle f_n auf S_1^c verschwinden, konvergiert $(\sqrt{f_n})$ gleichmäßig auf S gegen 0. Zu $\varepsilon > 0$ gibt es also n_0 mit

$$\sqrt{f_n(s)} \leq \varepsilon \qquad \forall s \in S, \ \forall n \geq n_0.$$

Daher

$$\begin{aligned} \ell(f_n) &= \ell(\sqrt{f_n}\,\sqrt{f_n}) \\ &\leq \ell(\sqrt{f_n}\,\sqrt{f_1}) && \text{da } f_n \leq f_1 \text{ und da } \ell \text{ positiv ist} \\ &\leq \ell(\varepsilon\,\sqrt{f_1}) && \text{für } n \geq n_0 \quad (\text{da } \ell \text{ positiv, linear}) \\ &= \varepsilon\,\ell(\sqrt{f_1}). \end{aligned}$$

Das zeigt $\ell(f_n) \xrightarrow[n\to\infty]{} 0$.

Nun haben wir alle Hilfsmittel zum Beweis des Hauptsatzes.

<u>2.12 Theorem</u> (Rieszscher Darstellungssatz)

Sei S ein lokalkompakter σ-kompakter metrischer Raum und $\ell: \mathcal{K}(S) \to \mathbb{R}$ ein positives lineares Funktional. Dann existiert ein eindeutig bestimmtes reguläres Borelmaß μ mit

$$\ell(f) = \int_S f \, d\mu \qquad \forall f \in \mathcal{K}(S).$$

Beweis: ℓ ist Daniell-stetig nach 2.11. Der Satz von Daniell-Stone (1.3) liefert ein darstellendes Maß μ für ℓ auf der kleinsten σ-Algebra, für die die $f \in \mathcal{K}(S)$ meßbar sind (beachte, daß $\mathcal{K}(S)$ ein Stonescher Funktionenraum ist); diese σ-Algebra ist $\mathrm{Bor}(S)$ nach 2.6.

μ ist auf kompakten Mengen endlich, denn eine Kombination von 2.2 und 2.3 (wie schon einmal angewandt, vgl. S. 145 oben) zeigt, daß jedes kompakte $C \subset S$ in der Form $C = \{f = 1\}$ für ein $f \in \mathcal{K}(S)$, $0 \leq f \leq 1$, geschrieben werden kann. Daher

$$\mu(C) = \int_S 1_C \, d\mu \leq \int_S f \, d\mu = \ell(f) < \infty.$$

Mit 2.7 muß μ regulär sein.

Zur Eindeutigkeit: Es reicht (siehe 1.3), 1 als $\sup f_n$ für geeignete $f_n \in \mathcal{K}(S)$ darzustellen. Dazu sei $S = \bigcup K_n$ (K_n kompakt), dann wähle $f_n \in \mathcal{K}(S)$ mit $0 \leq f_n \leq 1$, $K_n = \{f_n = 1\}$ (Beweis s.o.).

Es sei betont, daß der Rieszsche Darstellungssatz unter der allgemeinen Voraussetzung, daß S ein lokalkompakter topologischer Raum ist, gilt. Konsultiere z.B. die Bücher von Rudin oder Behrends! (Der Beweis des allgemeinen

Falls unterscheidet sich stark von dem hiesigen, denn einige hier verwendete Ingredienzien gelten nur für metrische Räume (so ist z.B. nicht jedes endliche Maß auf einem kompakten topologischen Raum regulär; für ein Gegenbeispiel siehe Halmos, S. 231). Radon und Behrends geben übrigens wesentlich verschiedene Beweise!)

Der Rieszsche Darstellungssatz ist der Ausgangspunkt des Aufbaus der Maß- und Integrationstheorie à la Bourbaki. Bourbaki definiert ein positives lineares Funktional auf $\mathcal{K}(S)$ als „Radonmaß" und entwickelt daraus die Integrationstheorie. Am Schluß dieser Entwicklung steht dann der Begriff des Maßes auf der Borel-σ-Algebra (Einzelheiten entnehme man den in der Einleitung genannten Büchern.)

Zu diesen beiden Ansätzen schreibt L. Schwartz in der Einleitung seines Buches "Radon Measures on Arbitrary Topological Spaces and Cylindrical Measures":

> There are essentially two methods of presenting the theory of measure: the abstract theory and the theory of Radon measures on locally compact spaces. The justification for the abstract theory is essentially that the topology has, *a priori*, nothing to do with the problem. The justification for the Radon measures is that, in fact, every set in analysis occurs with a topology or several topologies.
>
> The defects of the abstract theory are (i) the catastrophe of image measures (see § 5), (ii) the product of Borel σ-algebras of Hausdorff spaces is not in general the Borel σ-algebra of the product, (iii) an abstract measure even on the Borel σ-algebra has in general no support, etc. The fundamental defect of the usual theory of Radon measures is that the spaces of functions which occur in the theory of probability are not locally compact.

Was den letzten Kritikpunkt angeht, ist die folgende Klasse topologischer Räume von großer Bedeutung in der Wahrscheinlichkeitstheorie:

2.13 Definition Ein topologischer Raum S heißt polnischer Raum, wenn S separabel*) ist und die Topologie von einer Metrik d, die (S,d) zu einem vollständigen metrischen Raum macht, erzeugt wird.

(Kürzer gesagt: Ein polnischer Raum ist ein vollständig metrisierbarer separabler topologischer Raum.)

N. Bourbaki prägte diesen Begriff, um an die Leistungen der polnischen Topologen (Borsuk, Kuratowski, Sierpinski, Ulam, ...) in den 20er & 30er Jahren zu erinnern.

Es ist zu dieser Definition zu bemerken, daß nicht jede Metrik, die die Topologie eines polnischen Raumes erzeugt, vollständig ist. Betrachte z.B. den polnischen Raum $\mathbb{R}$ (die übliche Metrik $d(s,t) = |s-t|$ ist vollständig). Aber auch $\tilde{d}(s,t) = |\arctan s - \arctan t|$ erzeugt die euklidische Topologie, und (n) ist eine nichtkonvergente $\tilde{d}$-Cauchy-Folge! Andererseits ist die Wahl einer erzeugenden vollständigen Metrik nicht immer naheliegend. Hier ein harmloses Beispiel: In der üblichen Metrik $d(s,t) = |s-t|$ ist das offene Intervall $]-\frac{\pi}{2}, \frac{\pi}{2}[$ nicht vollständig. $d'(s,t) = |\tan s - \tan t|$ definiert aber eine vollständige Metrik, die die übliche Topologie erzeugt. – Entscheidend ist die Existenz einer vollständigen Metrik, die konkrete Gestalt ist relativ unerheblich.

Daß es mehr polnische Räume gibt, als man im ersten Moment erwartet, zeigt der nächste Satz.

*) S heißt separabel, wenn es eine abzählbare dichte Teilmenge gibt.

<u>2.14 Satz</u> a) Eine abgeschlossene Teilmenge eines polnischen Raums ist polnisch [bzgl. der Relativtopologie].

b) Eine offene Teilmenge eines polnischen Raums ist polnisch [bzgl. der Relativtopologie].

c) Ein endliches oder abzählbares Produkt polnischer Räume ist polnisch [bzgl. der Produkttopologie].

d) Eine G_δ - Teilmenge (das ist eine Teilmenge der Form $\bigcap_{n\in\mathbb{N}} U_n$ mit offenen U_n [G wie Gebiet - δ wie Durchschnitt]) eines polnischen Raums ist polnisch [bzgl. der Relativtopologie].

Beweis: Wir übernehmen aus der Topologie die Tatsache, daß eine Teilmenge eines separablen metrischen Raums selbst separabel ist. Auch das höchstens abzählbare Produkt separabler topologischer Räume ist separabel.

Sei (S,d) ein vollständiger metrischer Raum.

a) Ist $F \subset S$ abgeschlossen, so ist $d|_{F\times F}$ eine vollständige Metrik für F.

b) Sei $U \subset S$ offen. Definiere für $s,t \in U$

$$d_0(s,t) = d(s,t) + \left| \frac{1}{d(s,U^c)} - \frac{1}{d(t,U^c)} \right| .$$

Es ist klar, daß d_0 eine Metrik auf U definiert. Ferner gilt

$$d_0(s_n,s) \to 0 \iff d(s_n,s) \to 0 ,$$

so daß d_0 die Relativtopologie erzeugt.

(„$\Rightarrow$" wegen $d \le d_0$

„$\Leftarrow$" da $d(\cdot,U^c)$ stetig [bzgl. d] ist.)

Sei nun (s_n) eine d_0 - Cauchyfolge in U. Wegen $d \le d_0$ ist (s_n) auch eine d - Cauchyfolge, so daß $\lim_{n\to\infty} s_n = s \in S$ existiert. Wäre $s \notin U$, folgte $d(s_n,U^c) \to 0$, und $\left(\frac{1}{d(s_n,U^c)}\right)$ wäre keine Cauchyfolge reeller Zahlen im Widerspruch zu

$$\left| \frac{1}{d(s_n,U^c)} - \frac{1}{d(s_m,U^c)} \right| \le d_0(s_n,s_m).$$

Akzeptieren wir c) für einen Moment und kommen wir zum Beweis von d): Betrachte offene $U_n \subset S$ und $G = \bigcap_{n\in\mathbb{N}} U_n$.

Nach b) & c) ist $\prod_{n=1}^{\infty} U_n$ polnisch. Betrachte darin nun die „Diagonale"

$$\Delta = \{ (u_n) \in \prod U_n : u_n = u_1 \ \forall n \ge 2 \}.$$

Mit geringen Kenntnissen über die Produkttopologie kann man einsehen, daß Δ abgeschlossen in $\prod U_n$ ist. Daher ist auch Δ polnisch (a)).

Da schließlich $\varphi : G \to \Delta$, $\varphi(s) = (s,s,s,\dots)$ eine Bijektion ist, die mitsamt ihrer Umkehrabbildung stetig ist, ist G polnisch.

(Die Stetigkeit folgt wieder aus elementaren Eigenschaften der Produkttopologie.)

c) Seien (S_n, d_n) $(n \in \mathbb{N})$ vollständige metrische Räume. Es ist nicht schwer zu verifizieren, daß

$$d\left((s_n)_{n\in\mathbb{N}}, (t_n)_{n\in\mathbb{N}}\right) = \sum_{n=1}^{\infty} 2^{-n} \frac{d_n(s_n,t_n)}{1+d_n(s_n,t_n)}$$

eine vollständige Metrik ist, die die Produkttopologie erzeugt.

Im Falle endlich vieler Faktoren $S_1, \dots, S_N$ reicht es übrigens, mit

$$d\left((s_1,\dots,s_N),(t_1,\dots,t_N)\right) = \sum_{n=1}^{N} d_n(s_n,t_n)$$

zu arbeiten.

2.15 Beispiele a) $\mathbb{R}^k$ ist der Archetyp des polnischen Raums. Allgemeiner ist jeder separable Banachraum (z. B. $C([0,1])$, $L^p[0,1]$ $(1 \le p < \infty)$) polnisch. Übrigens besagt ein Satz von Riesz aus der Funktionalanalysis, daß kein unendlichdimensionaler Banachraum lokalkompakt ist!

Das Archetypische an $\mathbb{R}$ wird durch den folgenden Isomorphiesatz von Kuratowski ausgedrückt:

> Ist S ein polnischer Raum und $A \in \mathrm{Bor}(S)$ überabzählbar, so existiert eine Bijektion $\varphi: A \to \mathbb{R}$, so daß φ und φ^{-1} Borel-meßbar sind.

(Zum Beweis siehe etwa Cohn, S. 275.)

b) $\mathbb{Q}$ ist nicht polnisch. Das folgt aus dem Satz von Baire (vgl. S. 1??), wonach ein polnischer Raum von 2. Kategorie ist. (Gibt es ein direktes Argument?) Hingegen bilden die Irrationalzahlen $\mathbb{R} \setminus \mathbb{Q}$ einen polnischen Raum, wie aus 2.14 d) folgt (setze $U_n = \mathbb{R} \setminus \{r_n\}$, wo $\mathbb{Q} = \{r_1, r_2, \dots\}$). Dieses Beispiel ist fundamental in der Theorie der „analytischen Mengen", vgl. Cohn, Chap. VIII.

c) Aus Resultaten der allgemeinen Topologie ergibt sich, daß ein lokalkompakter σ-kompakter metrischer Raum S polnisch ist. (Unter diesen Voraussetzungen ist nämlich die Ein-Punkt-Kompaktifizierung $\alpha(S)$ metrisierbar, und kompakte metrische Räume sind separabel und vollständig. S ist schließlich offen in $\alpha(S)$.)

d) Sei (S,d) ein vollständiger metrischer separabler (also polnischer) Raum. Auf $W(S)$, der Menge der Wahrscheinlichkeitsborelmaße auf S, sei

$$d_0(\mu,\nu) = \inf\{\varepsilon > 0 : \mu(A) \le \nu(A^\varepsilon) + \varepsilon \quad \forall A \in \mathrm{Bor}(S)\},$$

wo $A^\varepsilon = \{s \in S : d(s,A) < \varepsilon\}$. Man kann zeigen, daß d_0 eine Metrik ist (sog. Prohorov-Metrik), die $W(S)$ zu einem polnischen Raum macht. Ferner gilt

$$d_0(\mu_n,\mu) \to 0 \iff \int_S f\,d\mu_n \to \int_S f\,d\mu \quad \forall f \in C^b(S).$$

Die Bedeutung der polnischen Räume für die Maßtheorie resultiert aus dem nächsten Satz sowie dem im nächsten Abschnitt zu besprechenden Satz von Kolmogoroff (3.5).

2.16 Satz Auf einem polnischen Raum S ist jedes endliche Borelmaß μ regulär.

Beweis: Wir verwenden die Strategie des Beweises von I 4.6. Die Beweisteile a), b), c) übertragen sich. Es bleibt d) zu zeigen, d. h.

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

für alle abgeschlossenen $C \subset S$.

Da abgeschlossene Teilmengen polnischer Räume polnisch sind (2.14 a)), reicht es, dies für $C = S$ zu zeigen.

Sei dazu d eine vollständige, die Topologie von S erzeugende Metrik, und sei (s_n) eine dichte Folge in S. Für $n, m \in \mathbb{N}$ definiere

$$B_{nm} = \{s \in S : d(s, s_n) \le \tfrac{1}{m}\}.$$

Für fixes m ist dann $S = \bigcup_{n=1}^{\infty} B_{nm}$, daher existiert zu $\varepsilon > 0$ $N(m) \in \mathbb{N}$ mit

$$\mu\Big(S \setminus \bigcup_{n=1}^{N(m)} B_{nm}\Big) < \frac{\varepsilon}{2^m}$$

(benutze I.24). Setze

$$K = \bigcap_{m=1}^{\infty} \bigcup_{n=1}^{N(m)} B_{nm}.$$

Es folgt

$$\mu(S \setminus K) = \mu\Big(\bigcup_{m=1}^{\infty}\big(S \setminus \bigcup_{n=1}^{N(m)} B_{nm}\big)\Big)$$

$$\leq \sum_{m=1}^{\infty} \mu\big(S \setminus \bigcup_{n=1}^{N(m)} B_{nm}\big)$$

$$\leq \varepsilon,$$

und es bleibt, K als kompakt zu entlarven.

Auf jeden Fall ist K nach Konstruktion abgeschlossen.
Sei nun $t_1, t_2, \ldots$ eine Folge in K. Wir werden zeigen, daß (t_n) eine Cauchyteilfolge besitzt, die wegen der Vollständigkeit von S und Abgeschlossenheit von K in K konvergieren muß.

1. Schritt dazu: Wegen $K \subset \bigcup_{n=1}^{N(1)} B_{n1}$ existiert n_1, so daß $B_{n_1 1}$ unendlich viele der t_n enthält, sagen wir

$$u_1 := t_{k_1}, \quad u_2 := t_{k_2}, \quad u_3 = t_{k_3}, \text{ etc.}$$

2. Schritt: Wegen $K \subset \bigcup_{n=1}^{N(2)} B_{n2}$ existiert n_2, so daß $B_{n_2 2}$ unendlich viele der u_k enthält, sagen wir

$$v_1 := u_{l_1}, \quad v_2 := u_{l_2}, \quad v_3 := u_{l_3}, \text{ etc.}$$

3. Schritt (und die übrigen): sollte jetzt klar sein.

Nach Konstruktion ist die Diagonalfolge

$$t_1, u_2, v_3, v_4, \ldots$$

eine Cauchyteilfolge von (t_n).

V 3. Maße auf unendlichen Produkten

In II.6 wurde das Produkt zweier σ-endlicher Maßräume eingeführt und studiert; die Ausweitung auf endlich viele Faktoren ist kanonisch (vgl. S. 125). Hier sollen nun Produkte mit unendlich vielen Faktoren untersucht werden, was von fundamentaler Bedeutung für die Wahrscheinlichkeitstheorie ist (s.u., S. 214 ff.)

Sei $I \neq \emptyset$ eine Indexmenge, und für alle $i \in I$ sei eine Menge S_i gegeben. Man setzt

$$\prod_{i\in I} S_i = \{ s : I \to \bigcup_{i \in I} S_i : \; s_i := s(i) \in S_i \},$$

nach dem Auswahlaxiom ist $\prod_{i\in I} S_i \neq \emptyset$, sobald alle $S_i \neq \emptyset$ sind.
Eine vertraute (und naheliegende) Schreibweise für die $s \in \prod_{i \in I} S_i$ ist $(s_i)_{i\in I}$. (Teste dein Verständnis dieses Begriffs im Fall endlicher I!)
Stimmen alle S_i überein, etwa $S_i = S \; \forall i \in I$, schreibt man

$$S^I := \prod_{i\in I} S_i ;$$

S^I besteht also aus allen Abbildungen von I nach S.
In Anwendungen ist häufig $I = \mathbb{N}$ oder $I \subset \mathbb{R}$ anzutreffen.

Seien nun meßbare Räume $(S_i, \mathfrak{A}_i)$, $i \in I$, vorgelegt.

3.1 Definition a) Ein meßbares Rechteck ist eine Menge $A \subset \prod_{i \in I} S_i$ der Form

$$A = A_{i_1} \times \cdots \times A_{i_r} \times \prod_{i \neq i_1, \ldots, i_r} S_i$$

wo $r \in \mathbb{N}$, $A_{i_k} \in \mathfrak{A}_{i_k}$

b) Die von den meßbaren Rechtecken erzeugte σ-Algebra heißt Produkt-σ-Algebra. Schreibweise: $\bigotimes_{i\in I}\mathfrak{A}_i$ bzw., falls alle $\mathfrak{A}_i = \mathfrak{A}$, $\mathfrak{A}^I$.

c) Die von den meßbaren Rechtecken erzeugte Algebra heißt Algebra $\mathcal{Z}$ der Zylindermengen.

Offensichtlich ist A genau dann eine Zylindermenge, wenn es eine endliche Menge $F \subset I$ und $A_F \in \bigotimes_{i\in F}\mathfrak{A}_i$ mit

$$A = A_F \times \prod_{i\in I\setminus F} S_i$$

gibt (Beweis mit good-sets-principle!)

Man kann sich daher Zylindermengen als Mengen von Funktionen vorstellen, deren Funktionswerte an endlich vielen Stellen (nämlich bei den $i_k\in F$) meßbaren Einschränkungen unterworfen sind (nämlich $(s_{i_1},\dots,s_{i_r})\in A_F$, falls $F=\{i_1,\dots,i_r\}$).

Analog hängt eine Menge $A\in\bigotimes_{i\in I}\mathfrak{A}_i$ nur „von abzählbar vielen Koordinaten" ab, was präzise folgendes bedeutet:

Es existiert eine höchstens abzählbare Teilmenge $J\subset I$ mit

(*) $(s_i)_{i\in I}\in A,\quad s_j=t_j\ \forall j\in J \Rightarrow (t_i)_{i\in I}\in A.$

Der Beweis hierfür verläuft ebenfalls mit dem good-sets-principle.

Diese Bemerkung deutet an, daß die Produkt-σ-Algebra relativ klein ist, siehe Beispiel 33 b).

Das folgende Lemma ist analog zu II.6.2.

32 Lemma Seien $(S_i,\mathfrak{A}_i)$ meßbare Räume, $\pi_j: \prod_{i\in I} S_i\to S_j$ sei durch $\pi_j((s_i)_{i\in I}) = s_j$ definiert.

a) Die π_j sind $\bigotimes_{i\in I}\mathfrak{A}_i$-$\mathfrak{A}_j$-meßbar.

b) Ist $(S,\mathfrak{A})$ ein weiterer meßbarer Raum, so ist eine Abbildung $T: S\to\prod_{i\in I}S_i$ genau dann $\mathfrak{A}$-$\bigotimes_{i\in I}\mathfrak{A}_i$-meßbar, wenn alle $\pi_j\circ T: S\to S_j$ $\mathfrak{A}$-$\mathfrak{A}_j$-meßbar sind.

Beweis: a) Für $A_j\in\mathfrak{A}_j$ ist $\pi_j^{-1}(A_j) = A_j\times\prod_{i\neq j}S_i$ ein meßbares Rechteck.

b) Mit a) ist klar, daß die $\pi_j\circ T$ meßbar sind, wenn T es ist. Zur Umkehrung reicht es zu zeigen, daß $T^{-1}(R)\in\mathfrak{A}$ ist, wenn R ein meßbares Rechteck ist; denn diese erzeugen die Produkt-σ-Algebra. Ist nun $R = A_{i_1}\times\dots\times A_{i_r}\times\prod_{i\neq i_k} S_i$, so ist in der Tat

$$\begin{aligned} T^{-1}(R) &= T^{-1}\Big(\bigcap_{k=1}^r \pi_{i_k}^{-1}(A_{i_k})\Big) \\ &= \bigcap_{k=1}^r T^{-1}\big(\pi_{i_k}^{-1}(A_{i_k})\big) \\ &= \bigcap_{k=1}^r (\pi_{i_k}\circ T)^{-1}(A_{i_k}) \in \mathfrak{A}. \end{aligned}$$

33 Beispiele a) Sei $I=\mathbb{N}$ und $S_i=\mathbb{R}\ \forall i$. Dann ist

$$A = \{(s_i)_{i\in\mathbb{N}} : |s_i|\le 1\ \forall i\in\mathbb{N}\}\in \mathrm{Bor}(\mathbb{R})^{\mathbb{N}},$$

da

$$A = \bigcap_{i\in\mathbb{N}}\pi_i^{-1}([-1,1]).$$ Genauso ist klar, daß

$$\sup: (s_i)_{i\in\mathbb{N}}\mapsto \sup s_i$$

eine $\overline{\mathbb{R}}$-wertige, $\mathrm{Bor}(\mathbb{R})^{\mathbb{N}}$-meßbare Funktion definiert.

b) Sei nun $I = [0,1]$ sowie $S_i = \mathbb{R}$ und $\mathfrak{A}_i = \mathrm{Bor}(\mathbb{R})$ $\forall i$. Da $C[0,1] \subset \mathbb{R}^{[0,1]}$ nicht von abzählbar vielen Koordinaten im Sinn von (*), S. 206 abhängt, gehört $C[0,1]$ nicht zur Produkt-σ-Algebra. Genausowenig trifft das für

$$\{(s_i)_{i\in[0,1]} : |s_i| < 1 \quad \forall i \in [0,1]\}$$

$$\left(= \{f : [0,1] \to \mathbb{R} : -1 \le f \le 1\}\right)$$

zu.

Dasselbe Argument zeigt, daß einpunktige Mengen nicht zu $\mathrm{Bor}(\mathbb{R})^{[0,1]}$ gehören!

Es sollen nun Wahrscheinlichkeitsmaße auf unendlichen Produkträumen konstruiert werden. Nehmen wir an, ein Wahrscheinlichkeitsmaß $\mu : \bigotimes_{i\in I} \mathfrak{A}_i \to [0,1]$ sei gegeben. Bezeichnet man für $F \subset I$ mit $\pi_F : \prod_{i\in I} S_i \to \prod_{i\in F} S_i$ die kanonische Projektion (die nach 32 bzgl der Produkt-σ-Algebra meßbar ist), so erhält man eine Familie von Bildmaßen $\mu_F := \mu \circ \pi_F^{-1}$. Diese sind folgendermaßen untereinander verknüpft: Gilt $F \subset G \subset I$ und setzt man $\pi_{FG} : \prod_{i\in G} S_i \to \prod_{i\in F} S_i$, $(s_i)_{i\in G} \mapsto (s_i)_{i\in F}$, so ist

$$(*) \qquad \mu_F = \mu_G \circ \pi_{FG}^{-1} \qquad (\text{für } F \subset G)$$

(da $\pi_{FG} \circ \pi_G = \pi_F$.) In Anwendungen stellt sich jedoch das umgekehrte Problem: Gegeben ist eine Familie von Wahrscheinlichkeitsmaßen μ_F auf $\bigotimes_{i\in F} \mathfrak{A}_i$, wo F das System $\mathcal{F}(I)$ aller endlichen Teilmengen von I durchläuft, so daß die Konsistenzbedingung (*) erfüllt ist (So eine Familie heißt projektive Familie von Wahrscheinlichkeitsmaßen.) Existiert dann ein Maß μ auf $\bigotimes_{i\in I} \mathfrak{A}_i$ mit

$$\mu_F = \mu \circ \pi_F^{-1} \qquad \forall F \in \mathcal{F}(I) \;?$$

(So ein Maß heißt projektiver Limes der μ_F.) Der Satz von Kolmogoroff (3.5) gibt darauf die Antwort.

34 Beispiele projektiver Familien:

a) $(S_i, \mathfrak{A}_i, \mu_i)$ $(i \in I)$ seien Wahrscheinlichkeitsräume. Für $F \in \mathcal{F}(I)$ sei

$$\mu_F = \bigotimes_{i\in F} \mu_i$$

(ein Produkt von endlich vielen Wahrscheinlichkeitsmaßen). Diese Familie ist projektiv

[Sei etwa $F = \{i_1, \dots, i_r\}$, $G = \{i_1, \dots, i_r, i_{r+1}, \dots, i_s\}$.

Für $A \in \bigotimes_{i\in F} \mathfrak{A}_i$ gilt

$$\begin{aligned}(\mu_G \circ \pi_{FG}^{-1})(A) &= \mu_G(\pi_{FG}^{-1}(A)) \\ &= \mu_G(A \times S_{i_{r+1}} \times \dots \times S_{i_s}) \\ &= \mu_F(A) \cdot \mu_{i_{r+1}}(S_{i_{r+1}}) \cdots \mu_{i_s}(S_{i_s}) \\ &= \mu_F(A).\end{aligned}$$

(In der vorletzten Zeile wurde benutzt, daß die Produktmaßbildung [mit endlich vielen Faktoren] assoziativ ist:

$$\mu_{i_1} \otimes \dots \otimes \mu_{i_s} = (\mu_{i_1} \otimes \dots \otimes \mu_{i_r}) \otimes (\mu_{i_{r+1}} \otimes \dots \otimes \mu_{i_s})$$

)]

b) Sei $I = [0,1]$ und $S_i = \mathbb{R}$ $\forall i$. Wir betrachten die Borel-σ-Algebra $\mathfrak{A}_i = \mathrm{Bor}(\mathbb{R})$ (so daß $\bigotimes_{i\in F} \mathfrak{A}_i = \mathrm{Bor}(\mathbb{R}^k)$ für $F = \{i_1, \dots, i_k\}$, II 6.3. a)). Ist $F = \{t_1, \dots, t_k\}$ mit $0 < t_1 < \dots < t_k$, so sei μ_F das λ^k-absolutstetige Maß mit der Dichte

$$\Phi_F(x_1, \dots, x_k) = (2\pi)^{-k/2} \left[t_1 (t_2 - t_1) \cdots (t_k - t_{k-1})\right]^{-\frac{1}{2}} \cdot \exp\left(-\frac{1}{2} \sum_{i=1}^{k} \frac{(x_i - x_{i-1})^2}{t_i - t_{i-1}}\right)$$

(hier ist $t_0 = x_0 = 0$ zu setzen), es sei $\mu_{\{0\}\cup F} = \delta_0 \otimes \mu_F$.
(δ_0 = Dirac-Maß bei 0)

Mit dem Satz von Fubini zeigt man leicht, daß es sich wirklich um Wahrscheinlichkeitsmaße handelt (beachte $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx = 1$). Außerdem ist es elementar, wenn auch etwas langwierig, zu zeigen, daß die μ_F eine projektive Familie bilden. (Wir kommen später noch einmal auf dieses Beispiel zurück.)

3.5 Theorem (Satz von Kolmogoroff)

Es seien S_i polnische Räume, versehen mit ihren Borel-σ-Algebren $\mathfrak{A}_i = Bor(S_i)$ $(i \in I)$. Dann besitzt jede projektive Familie von Wahrscheinlichkeitsmaßen einen eindeutig bestimmten projektiven Limes.
[Diese Begriffe wurden auf S. 208 eingeführt.]

Dem Beweis schicken wir ein Lemma voraus:

3.6 Lemma Für polnische Räume $S_1, \dots, S_n$ gilt

$$Bor(S_1 \times \dots \times S_n) = Bor(S_1) \otimes \dots \otimes Bor(S_n).$$

Beweis: Da alle Projektionen $\pi_i : S_1 \times \dots \times S_n \to S_i$ stetig, also Borel-meßbar sind, ist nach 3.2 die identische Abbildung auf $S_1 \times \dots \times S_n$ $Bor(S_1 \times \dots \times S_n)$-$Bor(S_1) \otimes \dots \otimes Bor(S_n)$-meßbar, d.h. es gilt „$\supset$". Für die umgekehrte Inklusion benötigen wir die Aussage, daß es in einem separablen metrischen Raum eine Folge offener Mengen $U_1, U_2, \dots$ gibt, die die Topologie erzeugt, d.h. jede offene Menge ist (abzählbare!) Vereinigung gewisser dieser U_i. (Man nennt das das 2. Abzählbarkeitsaxiom

in der Topologie.) (Für $\mathbb{R}$ siehe S. 11 oben, allgemein s.u., Fußnote *).)
Sind nun U_{im}, $m \in \mathbb{N}$, „Basen der Topologie" (im obigen Sinn) für S_i, so bilden die $U_{1m_1} \times \dots \times U_{nm_n}$ $(m_1, \dots, m_n \in \mathbb{N})$ eine Basis der Topologie von $S_1 \times \dots \times S_n$. Jede dieser Mengen liegt offenbar in $\bigotimes_{i=1}^{n} Bor(S_i)$, also auch alle abzählbaren Vereinigungen, d.h. jede offene Teilmenge von $S_1 \times \dots \times S_n$. Es folgt „$\subset$".
[Die Vollständigkeit war in diesem Lemma unerheblich.]

Beweis von 3.5: $(\mu_F)_{F \in \mathfrak{F}(I)}$ sei eine projektive Familie.
Ist

$$A = A_F \times \prod_{i \notin F} S_i \qquad (F \subset I \text{ endlich})$$

eine Zylindermenge, so setze

$$\mu_0(A) = \mu_F(A_F).$$

Wegen der Konsistenzbedingung (*) von S. 208 wird so eine wohldefinierte Mengenfunktion μ_0 auf der Algebra $\mathcal{Z}$ der Zylindermengen erklärt.
Da alle μ_F additiv sind, ist auch μ_0 additiv; und natürlich ist $\mu_0(\emptyset) = 0$.
Es bleibt, μ_0 als σ-additiv zu erkennen. Denn dann kann nach Carathéodory μ_0 zu einem Maß μ auf $\bigotimes_{i \in I} \mathfrak{A}_i = \sigma(\mathcal{Z})$ fortgesetzt werden, für das nach Konstruktion $\mu_F = \mu \circ \pi_F^{-1}$ $\forall F \in \mathfrak{F}(I)$ gilt. Ferner muß jeder projektive Limes der μ_F auf $\mathcal{Z}$ mit μ_0 übereinstimmen; I.3.8 zeigt daher die Eindeutigkeitsaussage.
Zum Beweis der σ-Additivität von μ_0 verwenden wir das Kriterium aus

*) Ordne etwa die $\{s : d(s, s_m) < \frac{1}{k}\}$, $k, m \in \mathbb{N}$, in einer Folge an, wo $\{s_1, s_2, \dots\}$ eine abzählbare dichte Menge bildet.

I.2.4. Seien also $A_n \in \mathcal{Z}$ mit $A_n \searrow \emptyset$. Die A_n haben dann die Form

$$A_n = \tilde{A}_n \times \prod_{i \notin F_n} S_i$$

für geeignete F_n und $\tilde{A}_n \in \mathcal{B}or\big(\prod_{i\in F_n} S_i\big) \underset{3.6}{=} \bigotimes_{i\in F_n} \mathcal{B}or(S_i)$. Da die μ_{F_n} Borelwahrscheinlichkeitsmaße auf dem polnischen Raum $\prod_{i\in F_n} S_i$ sind (2.11 nebst 3.6), sind sie regulär (2.16). Folglich existieren zu $\varepsilon > 0$ kompakte $\tilde{K}_n \subset \tilde{A}_n$ mit

$$\mu_{F_n}(\tilde{A}_n \setminus \tilde{K}_n) \le \varepsilon \cdot 2^{-n},$$

d.h., für

$$K_n := \tilde{K}_n \times \prod_{i\notin F_n} S_i$$

gilt

$$\mu_0(A_n \setminus K_n) \le \varepsilon \cdot 2^{-n}.$$

Nun ist erst recht $\bigcap_{n=1}^{\infty} K_n = \emptyset$ (da ja $K_n \subset A_n$), und wir behaupten nun (Beweis folgt):

(*) Es existiert n_0 mit $\bigcap_{n=1}^{n_0} K_n = \emptyset$.

Akzeptiert man (*) einen Moment, geht es nun weiter wie im Beweis von I.2.5: Für $C_n = \bigcap_{i=1}^{n} K_i$ folgt wie auf S. 26 induktiv

$$\mu_0(A_n \setminus C_n) \le \varepsilon \cdot (1 - 2^{-n}) \qquad (n \ge 1)$$

und daraus, da $C_n = \emptyset$ für $n \ge n_0$,

$$\mu_0(A_n) = \mu_0(A_n \setminus C_n) \le \varepsilon, \qquad \text{für } n \ge n_0$$

also in der Tat $\mu_0(A_n) \to 0$, was zu zeigen war.

Es bleibt, (*) zu beweisen. Das wäre klar im Fall kompakter S_i, da in diesem Fall nach dem Satz von Tychonoff aus der Topologie die K_n in

der Produkttopologie kompakt wären. Für den allgemeinen Fall machen wir zunächst die Notation etwas handlicher: Zuerst dürfen wir o.E. $F_n \subset F_{n+1}$ $\forall n$ annehmen (da man sonst in den $\tilde{A}_n$ gewisse S_i „herausmultipliziert"). Des weiteren sind für unser Problem nur abzählbar viele Koordinaten relevant (nämlich die in $\bigcup_n F_n$), daher sei o.E. $I = \mathbb{N}$ angenommen. Übergang zu $\prod_{i\in F_1} S_i$, $\prod_{i\in F_2\setminus F_1} S_i$, … statt $S_1, S_2, \dots$ gestattet schließlich, $F_n = \{1, \dots, n\}$ anzunehmen. Es ist jetzt also

$$K_n = \tilde{K}_n \times S_{n+1} \times S_{n+2} \times \cdots \quad \text{mit } \tilde{K}_n \subset S_1 \times \cdots \times S_n \text{ kompakt.}$$

Nehmen wir nun an, alle $C_m = \bigcap_{i=1}^{m} K_i$ wären $\neq \emptyset$. Wähle $s^{(m)} \in C_m$; so ist dann

$$s^{(m)} \in K_n \quad \text{für } n \le m \qquad (\text{da } C_m \subset K_n \text{ für } n \le m).$$

Bezeichnet $\pi_n : \prod_{i=1}^{\infty} S_i \to \prod_{i=1}^{n} S_i$ die kanonische Projektion, so liegt für jedes $n \in \mathbb{N}$ die Folge $(\pi_n(s^{(m)}))_{m\in\mathbb{N}}$ bis auf die ersten Glieder in der kompakten Menge $\tilde{K}_n$. Durch einen Diagonalfolgentrick erreicht man nun, daß es eine mit $\xi^{(m)}$ bezeichnete Teilfolge gibt, so daß für jedes $n \in \mathbb{N}$ $(\pi_n(\xi^{(m)}))_{m\in\mathbb{N}}$ konvergiert. Insbesondere konvergiert $(\xi^{(m)})$ punktweise, etwa gegen ξ; nach Konstruktion liegt dann ξ in allen K_n. Also erhält man den Widerspruch

$$\bigcap_{n=1}^{\infty} K_n \neq \emptyset. \qquad \text{(Ein Diagonalfolgenargument tauchte bereits auf S. 204 auf.)}$$

Wenden wir den Satz von Kolmogoroff auf die Beispiele aus 3.4 an: Sind in 3.4 a) die S_i zusätzlich polnisch, so folgt aus 3.5 die Existenz und Eindeutigkeit eines Produktmaßes $\bigotimes_{i\in I}\mu_i$, für das gilt:

$$\bigotimes_{i\in I}\mu_i\Big(A_{i_1} \times \cdots \times A_{i_r} \times \prod_{i \neq i_k} S_i\Big) = \mu_{i_1}(A_{i_1}) \cdot \ldots \cdot \mu_{i_r}(A_{i_r}).$$

Es ist wichtig zu bemerken, daß man Existenz und Eindeutigkeit des Produktmaßes $\bigotimes_{i\in I}\mu_i$ sogar für beliebige Wahrscheinlichkeitsräume beweisen kann (z.B. Behrends, S. 104)!

Der Satz von Kolmogoroff ist jedoch ganz ohne Voraussetzungen an die S_i falsch; Gegenbeispiele findet man z.B. bei Halmos, S. 214, oder Gänssler / Stute, Wahrscheinlichkeitstheorie, S. 261. Die „Kompaktheit" kann jedoch formal etwas abgeschwächt werden (Behrends, S. 115).

Das Produktmaß gestattet in der Wahrscheinlichkeitstheorie die mathematische Modellierung für Folgen unabhängiger Zufallsvariablen.

Auch das Beispiel 3.4 b) ist von großer Bedeutung in der Wahrscheinlichkeitstheorie. Zunächst liefert der Satz von Kolmogoroff den projektiven Limes der μ_F als ein Maß μ auf der Produkt-σ-Algebra von $\mathbb{R}^{[0,1]}$. Die Projektionen π_t ($t \in [0,1]$) bilden eine Familie von Zufallsvariablen (= meßbare Funktionen), einen sog. stochastischen Prozeß, auf dem Wahrscheinlichkeitsraum $(\mathbb{R}^{[0,1]}, (\mathrm{Bor}(\mathbb{R}))^{[0,1]}, \mu)$ mit folgenden Eigenschaften:

- $\pi_0 = 0$ μ-f.ü.

  $$\left[\ \mu(\{\pi_0 = 0\}) = \mu \circ \pi_0^{-1}(\{0\}) = \delta_0(\{0\}) = 1\ \right]$$

- für $t > s \geq 0$ ist $\pi_t - \pi_s$ normalverteilt mit Erwartungswert $\left(= \int (\pi_t - \pi_s)\, d\mu\right) = 0$ und Varianz $\left(= \int (\pi_t - \pi_s)^2\, d\mu\right) = t - s$.

- Ferner sind die Zufallsvariablen $\pi_{t_1}, \pi_{t_2} - \pi_{t_1}, \ldots, \pi_{t_k} - \pi_{t_{k-1}}$ ($0 < t_1 < \ldots < t_k$) unabhängig (d.h., vgl. II. 6.7 b), die gemeinsame Verteilung auf $\mathbb{R}^k$ ist das Produktmaß der einzelnen Verteilungen).

$$\Big[\ \mu(\{\pi_t - \pi_s \in A\}) = \mu_{\{s,t\}}(\{\pi_t - \pi_s \in A\}) \qquad (A \in \mathrm{Bor}(\mathbb{R}))$$

$$= \int_{\{(x,y):\ y - x \in A\}} \Phi_{\{s,t\}}\, d\lambda^2$$

$$\underset{\text{Fubini}}{=} \int_{\mathbb{R}} dx\, \frac{1}{(2\pi)^{1/2} s^{1/2}} \exp\left(\frac{-x^2}{2s}\right) \int_{\{y:\ y - x \in A\}} dy\, \frac{1}{(2\pi)^{1/2}(t-s)^{1/2}} \exp\left(\frac{-(y-x)^2}{2(t-s)}\right)$$

$$\underset{\text{Subst. } y - x = z}{=} \underbrace{\int_{\mathbb{R}} dx\, \frac{1}{(2\pi)^{1/2} s^{1/2}} \exp\left(\frac{-x^2}{2s}\right)}_{=1} \cdot \int_A dz\, \frac{1}{(2\pi)^{1/2}(t-s)^{1/2}} \exp\left(\frac{-z^2}{2(t-s)}\right)$$

$$= \int_A \varphi(z)\, dz$$

mit $\varphi(z) = \dfrac{1}{(2\pi)^{1/2}(t-s)^{1/2}} \exp\left(\dfrac{-z^2}{2(t-s)}\right)$, was wirklich die Dichte der fraglichen Normalverteilung ist *)

Zur Unabhängigkeit: Setze zur Abkürzung $\varphi_\lambda(z) = \dfrac{1}{(2\pi)^{1/2} \lambda^{1/2}} \exp\left(-\dfrac{z^2}{2\lambda}\right)$.

Es ist dann ($A_1, \ldots, A_k \in \mathrm{Bor}(\mathbb{R})$)

$$\mu(\{\pi_{t_1} \in A_1, \pi_{t_2} - \pi_{t_1} \in A_2, \ldots, \pi_{t_k} - \pi_{t_{k-1}} \in A_k\})$$

$$= \mu_{\{t_1, \ldots, t_k\}}(\{\pi_{t_1} \in A_1, \pi_{t_2} - \pi_{t_1} \in A_2, \ldots, \pi_{t_k} - \pi_{t_{k-1}} \in A_k\})$$

$$= \int_{\{(x_1, \ldots, x_k):\ x_1 \in A_1, x_2 - x_1 \in A_2, \ldots, x_k - x_{k-1} \in A_k\}} \Phi_{\{t_1, \ldots, t_k\}}\, d\lambda^k$$

$$\underset{\text{Fubini}}{=} \int_{A_1} dx_1\, \varphi_{t_1}(x_1) \int_{\{x_2:\ x_2 - x_1 \in A_2\}} dx_2\, \varphi_{t_2 - t_1}(x_2) \cdots \int_{\{x_k:\ x_k - x_{k-1} \in A_k\}} dx_k\, \varphi_{t_k - t_{k-1}}(x_k)$$

$$\underset{z_j = x_j - x_{j-1}}{=} \int_{A_1} dz_1\, \varphi_{t_1}(z_1) \int_{A_2} dz_2\, \varphi_{t_2 - t_1}(z_2) \cdots \int_{A_k} dz_k\, \varphi_{t_k - t_{k-1}}(z_k)$$

$$\underset{\text{s.o.}}{=} \mu(\{\pi_{t_1} \in A_1\}) \cdot \mu(\{\pi_{t_2} - \pi_{t_1} \in A_2\}) \cdots \mu(\{\pi_{t_k} - \pi_{t_{k-1}} \in A_k\}),$$

was zu zeigen war $\Big]$

---

*) Der Fall $s = 0$ ist entsprechend zu behandeln.

Pfad dieses stochastischen Prozesses heißt jede Abbildung
$$t \mapsto \pi_t(f) \qquad (f \in \mathbb{R}^{[0,1]}).$$
(Das ist f selbst!)

Stochastische Prozesse mit der auf S. 214 unten beschriebenen Eigenschaften treten z.B. bei der Beschreibung von - sagen wir - Gasmolekülen auf, die infolge von sehr vielen Zusammenstößen erratische Bewegungen ausführen. (Die erste Bedingung besagt, daß sich das Teilchen zum Zeitpunkt $t=0$ bei $x=0$ befindet.)

(Der obige Prozeß (π_t) würde z.B. die zeitliche Entwicklung der x-Koordinate beschreiben.)

Das hier gegebene Modell ist nun von daher höchst unbefriedigend, daß die Pfade des Prozesses, nämlich die $f \in \mathbb{R}^{[0,1]}$, nicht stetig zu sein brauchen - was die Physik offensichtlich verlangt. Es stellt sich also die Frage, ob das Maß μ sinnvoll auf $C[0,1]$ und die Spur-σ-Algebra $C[0,1] \cap (\mathrm{Bor}(\mathbb{R}))^{[0,1]}$ (Def. I.1.8) eingeschränkt werden kann (Beachte $C[0,1] \notin (\mathrm{Bor}(\mathbb{R}))^{[0,1]}$! Bsp. 3.3 b) !)

Das folgende Kriterium ist dafür hinreichend. Wir setzen $\tilde{\pi}_t = \pi_t|_{C[0,1]}$, $\tilde{\pi}_F = \pi_F|_{C[0,1]}$

3.7 Satz Sei $I = [0,1]$ und $(\mu_F)_{F \in \mathcal{F}(I)}$ eine projektive Familie von Wahrscheinlichkeitsmaßen sowie μ ihr projektiver Limes auf $\mathbb{R}^I$. Falls $p>0$, $\delta>0$ und $K>0$ mit
$$\underset{s,t \in I}{\forall} \int_{\mathbb{R}^I} |\pi_s - \pi_t|^p \, d\mu = \int_{\mathbb{R}^2} |x_1 - x_2|^p \, d\mu_{\{s,t\}}(x_1,x_2) \le K \cdot |s-t|^{1+\delta}$$
existieren, gibt es ein eindeutig bestimmtes reguläres Borelmaß $\tilde{\mu}$ auf $C[0,1]$ [*1] mit
$$\tilde{\mu} \circ \tilde{\pi}_F^{-1} = \mu_F \qquad \text{für alle } F \in \mathcal{F}(I).$$

Ein Teil des Beweises wird als Lemma formuliert.

3.8 Lemma Die Spur $\mathfrak{A}$ der Produkt-σ-Algebra $\mathrm{Bor}(\mathbb{R})^I$ auf $C(I)$ stimmt mit der Borel-σ-Algebra $\mathrm{Bor}(C(I))$ überein.

Beweis. $\mathfrak{A} \subset \mathrm{Bor}(C(I))$: Dazu ist zu zeigen, daß die identische Abbildung auf $C(I)$ $\mathrm{Bor}(C(I))$-$\mathfrak{A}$-meßbar ist, was definitionsgemäß genau dann der Fall ist, wenn die identische Einbettung von $C(I)$ in $\mathbb{R}^I$ $\mathrm{Bor}(C(I))$-$\mathrm{Bor}(\mathbb{R})^I$-meßbar ist. Das folgt nun aus 3.2, da die $\tilde{\pi}_t$ stetig auf $C(I)$ sind.
$\mathrm{Bor}(C(I)) \subset \mathfrak{A}$: Im Beweis von 3.6 wurde bemerkt, daß jede offene Menge von $C(I)$ abzählbare Vereinigung offener Kugeln ist. Da jede offene Kugel abzählbare Vereinigung abgeschlossener Kugeln ist, reicht es,

---

[*1] $C[0,1]$ wird mit der Supremumsnorm $\|f\|_\infty = \sup_t |f(t)|$ versehen und wird so zu einem polnischen Raum.

$$A = \{ g \in C(I) : \|g - f\|_\infty \le r \} \in \mathfrak{A}$$

zu zeigen. Das ist aber klar wegen

$$A = \bigcap_{t \in I \cap \mathbb{Q}} \{ g : |\tilde{\pi}_t (g - f)| \le r \}$$

$$= \bigcap_{t \in I \cap \mathbb{Q}} \left(\pi_t^{-1}([f(t) - r, f(t) + r]) \cap C[0,1] \right).$$

Beweis von 37. Es genügt, $\tilde{\mu}$ auf der Spur-σ-Algebra $\mathfrak{A}$ zu finden (3.8), die Regularität gilt automatisch (2.16). Die Eindeutigkeit von $\tilde{\mu}$ folgt wie 35. Dabei bemerkt man, daß die erste Gleichheit in der genannten Bedingung eine Konsequenz des Transformationssatzes für Bildmaße II.5.3 ist.

Nun zum eigentlichen Beweis.

Einer Funktion $f \in \mathbb{R}^I$ ordne die stetige Funktion $\tilde{f}_n$ zu, die bei $0, 2^{-n}, 2 \cdot 2^{-n}, \ldots, 1$ mit f übereinstimmt und zwischen diesen Knotenpunkten linear ist. Die Abbildung

$$\Phi_n : \mathbb{R}^I \to C(I), \quad \Phi_n(f) = \tilde{f}_n$$

ist dann $\mathrm{Bor}(\mathbb{R})^I$ - $\mathfrak{A}$ - meßbar (denn die $\tilde{\pi}_t \circ \Phi_n$ sind meßbar).

Außerdem gilt

$$\|\tilde{f}_n - \tilde{f}_{n+1}\|_\infty \le \max_{1 \le i \le 2^n} \left| f\left(\tfrac{i}{2^n}\right) - f\left(\tfrac{i-1}{2^n}\right) \right| .$$

Wir zeigen nun, daß für μ-fast alle f die Folge $(\tilde{f}_n)$ gleichmäßig konvergiert.

[Da $C[0,1]$ einen Banachraum unter der Supremumsnorm bildet, ist nur

$$\|\tilde{f}_n - \tilde{f}_m\|_\infty \to 0 \qquad \text{für } n, m \to \infty$$

für μ-fast alle f zu zeigen. Nun gilt für $a > 0$

$$\mu(\{ f : \|\tilde{f}_n - \tilde{f}_{n+1}\|_\infty > a \}) \qquad \text{(die Meßbarkeit dieser Menge folgt aus 38)}$$

$$\underset{\text{s.o.}}{\le} \mu\left(\left\{ f : \max_{1 \le i \le 2^n} \left| f\left(\tfrac{i}{2^n}\right) - f\left(\tfrac{i-1}{2^n}\right) \right| > a \right\}\right)$$

$$= \mu\left(\bigcup_{i=1}^{2^n} \left\{ f : \left| f\left(\tfrac{i}{2^n}\right) - f\left(\tfrac{i-1}{2^n}\right) \right| > a \right\} \right)$$

$$\le \sum_{i=1}^{2^n} \mu\left(\left\{ |\pi_{i/2^n} - \pi_{(i-1)/2^n}| > a \right\} \right)$$

$$\le a^{-r} \sum_{i=1}^{2^n} \int_{\mathbb{R}^I} |\pi_{\frac{i}{2^n}} - \pi_{\frac{i-1}{2^n}}|^r \, d\mu$$

(Tschebyscheffsche Ungleichung II.3.8)

$$\le a^{-r} \sum_{i=1}^{2^n} K \, (2^{-n})^{1+\delta} \qquad \text{nach Vor.}$$

$$= K \, a^{-r} \, 2^{-n\delta}$$

Wähle ϑ mit $0 < \vartheta < \frac{\delta}{r}$ und wende die obige Abschätzung für $a = 2^{-n\vartheta}$ an. Es folgt

$$\mu(\{ f : \|\tilde{f}_n - \tilde{f}_{n+1}\|_\infty > 2^{-n\vartheta} \}) \le K \cdot 2^{-n(\delta - r\vartheta)}$$

und

$$\sum_{n=1}^{\infty} \mu(\{ f : \|\tilde{f}_n - \tilde{f}_{n+1}\|_\infty > 2^{-n\vartheta} \}) < \infty .$$

Das 1. Borel-Cantelli-Lemma (vgl. S. 90) liefert

$$\|\tilde{f}_n - \tilde{f}_{n+1}\|_\infty \le 2^{-n\vartheta} \qquad \text{für hinreichend große } n$$

μ-fast überall. Daher konvergiert

$$\sum_{n=1}^{\infty} \|\tilde{f}_n - \tilde{f}_{n+1}\|_\infty$$

für μ-fast alle f, und für diese f gilt

$$\|\tilde{f}_n - \tilde{f}_m\|_\infty \le \sum_{i=n+1}^{m} \|\tilde{f}_i - \tilde{f}_{i-1}\|_\infty \longrightarrow 0 \quad \text{mit } n, m \to \infty \,]$$

Es existiert eine meßbare Teilmenge $\Omega \subset \mathbb{R}^I$ mit $\mu(\Omega) = 1$, so daß

$$\lim \Phi_n(f) \quad \text{für } f \in \Omega \text{ existiert.}$$

Damit definiert man eine Abbildung

$$\Phi : \mathbb{R}^I \to C(I), \quad \Phi(f) = \lim_{n\to\infty} \Phi_n(f) \qquad (f \in \Omega)$$

$$\Phi(f) = 0 \qquad (f \notin \Omega).$$

Φ ist meßbar, denn alle $\tilde{\pi}_t \circ \Phi$ sind – als punktweise Limites von $1_\Omega \cdot \tilde{\pi}_t \circ \Phi_n$ – meßbar.

Das gesuchte Maß wird nun durch $\tilde{\mu} = \mu \circ \Phi^{-1}$ definiert.

Zum Beweis, daß $\tilde{\mu}$ die richtigen „Randverteilungen" μ_F besitzt, benötigen wir:

(*) Für fixes $t \in [0,1]$ gilt $\pi_t = \tilde{\pi}_t \circ \Phi$ μ-fast überall.[*)]

Das ist klar, wenn t von der Form $\frac{k}{2^n}$ ist, und gilt in diesem Fall sogar überall. Beliebiges $t \in [0,1]$ stelle man als $t = \lim t_j$ mit dyadischen Brüchen t_j dar. Aus der Voraussetzung von 3.7 folgt nun

$$\int_{\mathbb{R}^I} |\pi_{t_j} - \pi_t|^p \, d\mu \xrightarrow[j\to\infty]{} 0,$$

und wegen II 3.9 existiert eine f.ü. konvergente Teilfolge:

$$\pi_{t_{j_n}} \to \pi_t \qquad \mu\text{-f.ü.}$$

Da $\Phi(f)$ stets stetig ist, gilt auch

$$\tilde{\pi}_{t_{j_n}} \circ \Phi \to \tilde{\pi}_t \circ \Phi \qquad (\text{überall}).$$

Daraus folgt die Behauptung.

---

*) Das heißt nicht $f = \Phi(f)$ μ-fast überall, da die obige Nullmenge von t abhängt, aber $[0,1]$ ist überabzählbar. Es folgt nur $f = \Phi(f)$ μ-f.ü. auf fixen abzählbaren Teilmengen von $[0,1]$.

Der Beweis von 3.7 wird jetzt abgeschlossen durch

$$\tilde{\mu} \circ \tilde{\pi}_F^{-1}(A) = \tilde{\mu}(\{g \in C[0,1] : (g(t_1), \ldots, g(t_k)) \in A\})$$

(wo $F = \{t_1, \ldots, t_k\}$ und $A \in \mathrm{Bor}(\mathbb{R}^k)$)

$$= \mu(\{f \in \mathbb{R}^I : (\Phi f(t_1), \ldots, \Phi f(t_k)) \in A\})$$

(nach Def. von $\tilde{\mu}$)

$$= \mu(\{f \in \mathbb{R}^I : (f(t_1), \ldots, f(t_k)) \in A\})$$

(nach (*) von S. 218)

$$= \mu_F(A).$$

Kehren wir zum Abschluß zum Beispiel von S. 214 zurück. Da hier die $\pi_t - \pi_s$ normalverteilt sind, kann man ohne Mühe die Bedingung von Satz 3.7 verifizieren (z.B. für $p = 4$):

$$\int_{\mathbb{R}^I} |\pi_t - \pi_s|^4 \, d\mu = \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{t-s}} \int_{-\infty}^{\infty} x^4 \exp\left(-\frac{x^2}{2(t-s)}\right) dx$$

$$= 3(t-s)^2$$

(leichte Rechnung).

Das durch Satz 3.7 gelieferte Maß auf $C[0,1]$ heißt *Wienersches Maß* μ_W, der zugehörige Prozeß der $\tilde{\pi}_t$ mit den auf S. 214 unten genannten Eigenschaften und stetigen Pfaden *Brownsche Bewegung*.

Man kann zeigen, daß mit Wahrscheinlichkeit 1 (d.h. μ_W-fast sicher) die Pfade der Brownschen Bewegung an keiner Stelle differenzierbar sind! (Die Menge der differenzierbaren Funktionen ist übrigens keine Borelsche Teilmenge von $C[0,1]$!)

Literaturliste

H. Bauer: Wahrscheinlichkeitstheorie. 3. Auflage 1978.

E. Behrends: Maß- und Integrationstheorie. 1987.

D. Cohn: Measure Theory. 1980.

P. Halmos: Measure Theory. 1950.

A. Mukherjea, K. Pothoven: Real and Functional Analysis. 2. Auflage 1984.

Diese Bücher können alle als Begleitlektüre zur Vorlesung empfohlen werden. Am elementarsten scheint mir Behrends Buch zu sein, am anspruchsvollsten ist Mukherjea/Pothoven. Cohn ist sehr detailliert (beinahe zu detailliert). Bauer und Halmos sind schon Klassiker; die Vorlesung weicht von Halmos entscheidend ab, indem σ-Algebren und nicht σ-Ringe betrachtet werden.

W. Hackenbroch: Integrationstheorie. 1987.

E. Henze: Einführung in die Maßtheorie. 1971.

Zwei Taschenbücher, die das Wichtigste enthalten. Henzes Buch kommt mir etwas dröge vor.

E. Hewitt, K. Stromberg: Real and Abstract Analysis. 1965.

H. L. Royden: Real Analysis. 2. Auflage 1968.

W. Rudin: Real and Complex Analysis. 2. Auflage 1974, 3. Auflage 1986.

Teile dieser Bücher sind etwas schwierig zu lesen, aber umso lohnender. Ich empfehle – auch dem Leser [illegible] – dringend das Buch von Rudin.

R. Ash: Measure, Integration and Functional Analysis.

(Entstanden aus den ersten vier Kapiteln von "Real Analysis and Probability" (1972) desselben Autors.)

P. Billingsley: Probability and Measure. 2. Auflage 1986.

Besonders interessant für Anwendungen in der Wahrscheinlichkeitstheorie.

K. Floret: Maß- und Integrationstheorie. 1981.

A. J. Weir: Integration and Measure (2 Bde.) 1973/1974.

Diese Autoren bevorzugen einen konkurrierenden Aufbau, nämlich den sog. Daniellschen Zugang.

J. J. Benedetto: Real Variable and Integration. 1976.

I. N. Pesin: Classical and Modern Integration Theories. 1970.

Hier findet man einiges zur Geschichte der Maßtheorie.

K. Jacobs: Measure and Integral.

Ein absolut unersetzliches Buch, das fast alles enthält, was es an Wissenswertem in der Maßtheorie gibt.

R. M. Dudley: Real Analysis and Probability. 1989.

P. A. Meyer: Probabilités et Potentiel. 1966.

Beide Bücher mit Ausrichtung auf die [theoretische] Wahrscheinlichkeitstheorie.

G. Pedersen: Analysis Now. 1988.

J. Dieudonné: Grundzüge der modernen Analysis. Bd. 2. 1975 (dt.).

N. Bourbaki: Intégration. I – IX. 1965 ff.

Bourbakischer Aufbau der Maß- und Integrationstheorie.

L. Schwartz: Radon Measures on Arbitrary Topological Spaces. 1973.

Spezialmonographie zu Maßen auf topologischen Räumen.

K. R. Parthasarathy: Probability Measures on Metric Spaces. 1967.

dto.

K. P. S. Bhaskara-Rao, M. Bhaskara-Rao: Theory of Charges. 1984.

Alles über nicht σ-additive Inhalte.

N. Dunford, J. T. Schwartz: Linear Operators. I: General Theory. 1958.

Bibel des Funktionalanalytikers. Enthält einen großen Abschnitt über Integration [banachraumwertiger Funktionen].

D. Fremlin: Topological Riesz Spaces and Measure Theory. 1974.

Eine abstrakte Einführung auf der Grundlage der Theorie der geordneten Vektorräume.

bitte mit einheften

Korrigenda

S. 16: In Zeile 2 lies: $\{x_0 + a : a \in A_0\}$

S. 36: Zu Zeile −6: Ich habe inzwischen erfahren, daß der heute „Dynkinsystem" genannte Begriff bereits in den 20er Jahren bei Sierpiński auftaucht.

S. 28: In Zeile −5 ff ist α durch μ^* zu ersetzen.

S. 42: Ab Zeile 10 ist ein Index falsch, und schon geht alles schief. Hier die korrigierte Fassung:

$$T(Q) = \{ (\ldots) \quad (j \neq 1) \}$$

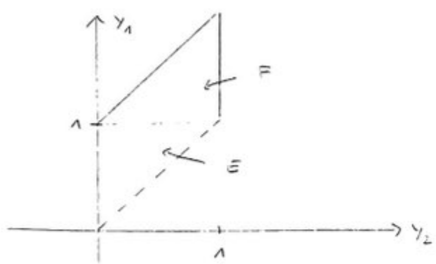
Sei

$$E = (\ldots) \quad y_1 \leq 1\},$$

$$F = (\ldots) \quad y_1 > 1\}.$$

Nun ist $F + (-1, 0, \ldots, 0) \cup E = Q$ (disjunkte Vereinigung), daher

$$\det T = 1 = \lambda^k(Q) = \lambda^k(F + (-1,0,\ldots,0)) + \lambda^k(E)$$
$$= \lambda^k(E) + \lambda^k(F) = \lambda^k(E \cup F) = \lambda^k(T(Q))$$



[etc.]

S.47 Der Beweis von e) kann wesentlich vereinfacht werden. Wähle nämlich $R>0$, so daß $K \subset U := \{x : \|x\| < R\}$ ist. Da U^- kompakt ist, gilt $\mu(U) \le \mu(U^-) < \infty$.

Der im Text gegebene Beweis ist nötig für die später diskutierte Verallgemeinerung I 2.7!

S.63 In Zeile 12 ist der Satz von Tietze gemeint.

S.65 In Zeile −4 lies $\varphi = \sum_{i=1}^{n} \alpha_i 1_{A_i}$.

S.70 Der Beweis von Korollar 2.6 benutzt Lemma 2.8! (Das ist kein Desaster, da 2.8 nicht auf 2.6 beruht.)

S.60 In Zeile −7 lies statt „siehe Kap I" : „siehe S. 147 ff."

S.114 Ergänze ein Komma am Ende von Zeile 12!

S.123 In Zeile 10 lies $\mathrm{Bor}([0,1]^2)$.

S.130 In Zeile −2 lies $A \in \sigma(Z_n)$.

S.163 In Zeile 3 ist der Faktor $\frac{1}{2}$ durch 2 zu ersetzen.

S.165 In Zeile 11 lies $\lambda(A) \le \delta$.

S.206 Definition 3.1 c) sollte so heißen:

c) Eine Zylindermenge ist eine Menge der Form

$$A_F \times \prod_{i \in I \setminus F} S_i ,$$

wo $F \subset I$ endlich und $A_F \in \bigotimes_{i \in F} \mathfrak{A}_i$ ist.

Danach sollte es heißen:

Offenbar bildet das System $\mathcal{Z}$ aller Zylindermengen eine Algebra. (Beweis zur Übung!)

S.221 letzte Zeile: Das dort erwähnte Resultat wurde von S. Mazurkiewicz, Fund. Math. 27 (1936), 244 - 249 bewiesen; siehe auch C.A. Rogers (ed.), Analytic Sets, S.91. Auch die nirgends differenzierbaren Funktionen bilden eine nichtborelsche Teilmenge von $C[0,1]$. Das ist (noch) schwieriger zu beweisen, siehe R.D. Mauldin, Pacific J. Math. 80 (1979), 199 - 205. Wenngleich die differenzierbaren bzw. nirgends differenzierbaren Funktionen keine Borelmengen bilden, so sind sie doch „universell meßbar", d.h. sie unterscheiden sich von einer Borelmenge nur um eine μ-Nullmenge (und zwar für jedes Wahrscheinlichkeitsmaß auf $C[0,1]$!).

Einen einfachen Beweis der Nichtdifferenzierbarkeit der Brownschen Pfade findet man bei R. Durrett, Brownian Motion and Martingales in Analysis, S. 6.

S.215 Zeile 1: lies: $\int_{\mathbb{R}} dx\,(\cdot) \int_{\{y:\, y-x \in A\}} dy\,(\cdots)$