

Reinhard Höpfner

Vorlesung Stochastik

*Kapitel III:

Lebesgue-Zerlegung von Massen

SoSe 16 – WS 16/17

Institut für Mathematik, Johannes Gutenberg Universität Mainz

October 31, 2016

Übersicht zu Kapitel III :

A. Absolutstetigkeit, Singularität, Lebesgue-Zerlegung

Definition der Absolutstetigkeit, der Äquivalenz, der Singularität von Massen 3.1–3.2

Beispiele 3.3

Formulierung: Hauptsatz über Lebesgue-Zerlegung σ -endlicher Masse 3.4

Beispiel 3.4'

Korollar: Satz von Radon-Nikodym 3.5

B. Beweise und Ergänzungen

Beispiel Cantorverteilung 3.6

Existenz der Lebesgue-Zerlegung 3.7

Eindeutigkeit der Lebesgue-Zerlegung 3.8

Beweis von 3.4 und 3.5 3.9

Konkrete Berechnung von Lebesgue-Zerlegungen 3.10

Beispiel 3.10'

Dichtequotienten und ihre Eindeutigkeitseigenschaft 3.11

Beispiel 3.11'

Transformationsformel für Dichten 3.12

In diesem Kapitel ist (Ω, \mathcal{A}) ein beliebiger messbarer Raum und μ ein Mass auf (Ω, \mathcal{A}) . Wie in 2.1 bezeichnet \mathcal{F}^* die Klasse aller nichtnegativen messbaren numerischen Funktionen (MNF), d.h. die Klasse aller \mathcal{A} -messbaren Abbildungen $f : \Omega \rightarrow [0, \infty]$.

A. Absolutstetigkeit, Singularität, Lebesgue-Zerlegung

3.1 Satz: a) Für festes $f \in \mathcal{F}^+$ wird durch

$$\nu(A) := \int_A f d\mu, \quad A \in \mathcal{A}$$

ein Mass ν auf \mathcal{A} definiert. Für Ereignisse $A \in \mathcal{A}$ gilt dabei

$$\mu(A) = 0 \implies \nu(A) = 0.$$

b) Ist das Ereignis $\{f = 0\}$ eine μ -Nullmenge, so verschärft sich die letzte Aussage zu

$$\text{für beliebiges } A \in \mathcal{A} : \mu(A) = 0 \iff \nu(A) = 0.$$

c) Ist μ ein σ -endliches Mass und ist $\{f = +\infty\}$ eine μ -Nullmenge, so ist auch ν ein σ -endliches Mass.

Beweis: a) Wegen Additivität (2.6) des Integrals auf \mathcal{F}^+ ist zunächst

$$A \longrightarrow \nu(A) = \int 1_A f d\mu, \quad A \in \mathcal{A}$$

ein Inhalt auf \mathcal{A} . Für aufsteigende Mengenfolgen $A_n \uparrow A$ in \mathcal{A} gilt mit Satz 2.5 und mit monotoner Konvergenz unter μ

$$\nu(A) = \int \sup_n (1_{A_n} f) d\mu = \sup_n \int 1_{A_n} f d\mu = \lim_n \nu(A_n),$$

also ist der Inhalt ν aufsteigend stetig und damit nach 1.17 ein Mass auf \mathcal{A} . Aus $\mu(A) = 0$ folgt mit 2.11 b) sofort $\int_A f d\mu = \nu(A) = 0$.

b) Sei nun zusätzlich $\{f = 0\}$ eine μ -Nullmenge. Dann kann man für beliebiges $A \in \mathcal{A}$ abschätzen

$$\nu(A) = \int 1_A f d\mu = \int 1_A (f 1_{\{f > 0\}}) d\mu \geq \int 1_A (f 1_{\{f > \frac{1}{n}\}}) d\mu \geq \frac{1}{n} \mu(A \cap \{f > \frac{1}{n}\}).$$

Man erhält

$$\nu(A) = 0 \implies \mu\left(A \cap \left\{f > \frac{1}{n}\right\}\right) = 0 \quad \text{für alle } n \in \mathbb{N}$$

und aufsteigende Stetigkeit von μ macht daraus

$$\mu(A) = \mu(A \cap \{f > 0\}) = 0.$$

c) Ist μ ein σ -endliches Mass auf (Ω, \mathcal{A}) , so existiert eine Mengenfølge $(C_n)_n$ in \mathcal{A} mit

$$C_n \uparrow \Omega, \quad \mu(C_n) < \infty \quad \text{für jedes } n.$$

Ist $\{f = +\infty\}$ eine μ -Nullmenge in \mathcal{A} , so sind die MNF f und $\tilde{f} := f1_{\{f < \infty\}} : \Omega \rightarrow [0, \infty)$ μ -äquivalent, siehe 2.11', also kann das in a) definierte Mass ν auch in der Form

$$\nu(A) = \int_A \tilde{f} d\mu, \quad A \in \mathcal{A}$$

geschrieben werden. Weiter definiert $E_n := \{0 \leq \tilde{f} \leq n\}$ eine Mengenfølge in \mathcal{A} mit der Eigenschaft $E_n \uparrow \Omega$ für $n \rightarrow \infty$. Damit hat man in \mathcal{A} eine aufsteigende Mengenfølge

$$D_n := C_n \cap E_n \uparrow \Omega, \quad n \rightarrow \infty$$

gefunden mit der Eigenschaft

$$\nu(D_n) = \int (1_{\{0 \leq \tilde{f} \leq n\}} \tilde{f}) 1_{C_n} d\mu \leq n \mu(C_n) < \infty, \quad n \geq 1.$$

Damit ist auch ν ein σ -endliches Mass auf (Ω, \mathcal{A}) . □

3.2 Definition: Betrachte Masse μ, ν auf demselben messbaren Raum (Ω, \mathcal{A}) .

a) ν heisst *dominiert durch* μ (oder: *μ -absolutstetig*, Schreibweise $\nu \ll \mu$) falls gilt

$$\text{für alle } A \in \mathcal{A} : \quad \mu(A) = 0 \implies \nu(A) = 0.$$

b) Man nennt die Masse μ, ν *äquivalent* (Schreibweise $\nu \sim \mu$) falls

$$\nu \ll \mu \quad \text{und} \quad \mu \ll \nu.$$

c) Man nennt μ, ν *zueinander singulär* (Schreibweise $\nu \perp \mu$) falls gilt

$$\text{es gibt ein } M \in \mathcal{A} \text{ mit } \mu(M) = 0 \text{ und } \nu(M^c) = 0.$$

d) Jede \mathcal{A} -Menge von vollem μ -Mass nennen wir einen Träger von μ .

3.3 Beispiele: Betrachte auf $(\Omega, \mathcal{A}) := (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ das Lebesgue-Mass λ und das zu $\mathbb{N}_0 \subset \mathbb{R}$ assoziierte Zählmass π , definiert durch

$$\pi(\mathbb{R} \setminus \mathbb{N}_0) = 0, \quad \pi(\{k\}) := 1 \text{ für alle } k \in \mathbb{N}_0.$$

Nach 3.1 und 3.2 gilt dann:

a) λ und π sind zueinander singulär: \mathbb{N}_0 ist Träger von π , zugleich aber als abzählbare Teilmenge von \mathbb{R} eine λ -Nullmenge.

b) Jede messbare Funktion $f : \mathbb{R} \rightarrow [0, \infty)$ definiert ein σ -endliches Mass

$$\nu(A) := \int_A f d\lambda, \quad A \in \mathcal{B}(\mathbb{R})$$

auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ mit $\nu \ll \lambda$; falls f auf \mathbb{R} strikt positiv ist, gilt sogar $\nu \sim \lambda$.

c) Jede Folge $(a_k)_{k \in \mathbb{N}_0}$ in $[0, \infty)$ definiert auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ ein σ -endliches Mass mit Träger \mathbb{N}_0

$$\nu(\{k\}) := a_k, \quad k \in \mathbb{N}_0, \quad \nu(\mathbb{R} \setminus \mathbb{N}_0) := 0.$$

Es gilt $\nu \ll \pi$, und ν kann für geeignetes $g : \mathbb{R} \rightarrow [0, \infty)$ \mathcal{A} -messbar in der Form

$$\nu(A) = \int_A g d\pi, \quad A \in \mathcal{A}$$

geschrieben werden; eine mögliche Wahl von g ist

$$g(y) := \begin{cases} a_k & \text{falls } y = k, k \in \mathbb{N}_0 \\ 0 & \text{sonst} \end{cases}, \quad y \in \mathbb{R}.$$

Gilt $a_k > 0$ für alle k , so sind die Masse ν und π äquivalent auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

d) Sei nun $\nu_1 \ll \lambda$ ein nach b) und $\nu_2 \ll \pi$ ein nach c) konstruiertes Mass. Dann ist $\nu := \nu_1 + \nu_2$ ein σ -endliches Mass auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, welches sowohl einen λ -absolutstetigen als auch einen λ -singulären Anteil besitzt. □

Wir formulieren nun den Hauptsatz dieses Kapitels:

3.4 Hauptsatz (Lebesgue-Zerlegung): Seien μ, ν σ -endliche Masse auf (Ω, \mathcal{A}) .

a) Stets existiert eine \mathcal{A} -messbare Funktion $f : \Omega \rightarrow [0, \infty)$ und eine μ -Nullmenge $N \in \mathcal{A}$ so dass ν bezüglich μ in der Form

$$\nu(A) = \nu(A \cap N) + \int_A f d\mu, \quad A \in \mathcal{A}$$

dargestellt werden kann. Jedes solche Paar (f, N) heisst *Lebesgue-Zerlegung von ν bezüglich μ* .

b) Sind (f_1, N_1) und (f_2, N_2) Lebesgue-Zerlegungen von ν bezüglich μ , $i = 1, 2$, so gilt

$$\nu(N_1 \Delta N_2) = 0 \quad \text{und} \quad \mu(\{f_1 \neq f_2\}) = 0.$$

c) Insbesondere sind in jeder Zerlegung des Masses ν gemäss a) die Anteile

$$\begin{aligned}\nu^{\text{ac}}(A) &:= \int_A f d\mu, \quad A \in \mathcal{A} \quad (\mu\text{-absolutstetiger Anteil}) \\ \nu^{\text{s}}(A) &:= \nu(A \cap N), \quad A \in \mathcal{A} \quad (\mu\text{-singulärer Anteil})\end{aligned}$$

als Masse auf (Ω, \mathcal{A}) eindeutig bestimmt.

3.4' Beispiel: In Beispiel 3.3 d) sieht die Lebesgue-Zerlegung von $\nu := \nu_1 + \nu_2$ bezüglich λ so aus: unter den dort gemachten Voraussetzungen gilt mit $f = \frac{d\nu_1}{d\lambda}$ aus 3.3 b) als Dichte des Lebesgue-absolutstetigen Anteils von ν und mit $N = N_0$ als λ -Nullmenge

$$\nu(A) = \nu(A \cap N) + \int_A f d\lambda, \quad A \in \mathcal{A}.$$

In diesem Beispiel stimmt $\nu^{\text{s}} = \nu(\cdot \cap N)$ mit ν_2 und ν^{ac} mit ν_1 überein.

Der Beweis des Hauptsatzes 3.4 wird in Teilkapitel B gegeben werden. Zusammen mit einer expliziten Konstruktion für Lebesgue-Zerlegungen wird dort (siehe 3.11) auch der 'Dichtequotient' betrachtet, der eine besonderes griffige Formulierung der Eindeutigkeitsaussage erlaubt. Sofort aus 3.4 folgt

3.5 Korollar: (Satz von Radon-Nikodym) Für σ -endliche Masse μ, ν auf (Ω, \mathcal{A}) sind gleichwertig:

- i) ν ist μ -absolutstetig ($\nu \ll \mu$);
- ii) es existiert eine \mathcal{A} -messbare Funktion $f : \Omega \rightarrow [0, \infty)$ mit der Eigenschaft

$$\nu(A) = \int_A f d\mu \quad \text{für alle } A \in \mathcal{A}.$$

Man nennt f in ii) eine Festlegung der *Dichte von ν bezüglich μ* (Schreibweise $f = \frac{d\nu}{d\mu}$). Eine solche Festlegung ist μ -fast sicher (damit wegen i) auch $(\mu + \nu)$ -fast sicher) eindeutig bestimmt.

B. Beweise und Ergänzungen

Beispiel 3.3 d) darf nicht dahingehend missverstanden werden, dass λ -singuläre Wahrscheinlichkeitsmasse auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ stets eine Verteilungsfunktion besitzen, die Sprünge aufweist: Stetigkeit der Verteilungsfunktion auf \mathbb{R} ist eine notwendige, aber nicht hinreichende Bedingung für λ -Absolutstetigkeit, wie das folgende Beispiel zeigt:

3.6 Beispiel: Die auf $[0, 1]$ konzentrierte *Cantorverteilung* ist ein Beispiel für eine Wahrscheinlichkeitsverteilung auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, die Lebesgue-singulär ist und gleichzeitig eine auf ganz \mathbb{R} stetige Verteilungsfunktion besitzt. Man definiert sie wie folgt.

1) Betrachte eine fallende Folge $(C_n)_n$ von Kompakta in \mathbb{R}

$$\begin{aligned} C_0 &:= [0, 1] \\ C_1 &:= [0, \frac{1}{3}] \cup [\frac{2}{3}, 1] \\ C_2 &:= [0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{1}{3}] \cup [\frac{2}{3}, \frac{7}{9}] \cup [\frac{8}{9}, 1] \\ &\dots \end{aligned}$$

usw., wobei jeweils C_{n+1} aus C_n durch 'Auswischen' des mittleren Drittels aller Segmente entsteht. Folglich besteht C_k aus 2^k Intervallen der Länge 3^{-k} , also

$$(+)$$

$$\lambda(C_k) = \left(\frac{2}{3}\right)^k, \quad k \in \mathbb{N}_0.$$

Betrachte nun

$$C := \bigcap_{n=0}^{\infty} C_n \in \mathcal{B}(\mathbb{R}).$$

Als Limes einer fallenden Folge von Kompakta ist C kompakt, und absteigende Stetigkeit von λ (genauer: der Restriktion von λ auf $[0, 1]$) zusammen mit (+) zeigt: C ist eine λ -Nullmenge.

2) Sicher ist C nach dem Cantorschen Durchschnittssatz nichtleer. Dabei kann jeder Punkt von C durch eine 0-1-Folge dargestellt werden, die in der Art eines random walk ('Kopf oder Zahl') vorschreibt, ob ein zur Zeit k einem Segment von C_k zugeordneter Wanderer beim Auswischen des mittleren Drittels dieses Segmentes sich für das verbleibende erste oder dritte Drittel des Segments entscheiden wird. Damit ist C insbesondere überabzählbar. Mit dyadischer Entwicklung kann die Menge aller 0-1-Folgen aber auch bijektiv auf die Menge aller Punkte des Intervals $[0, 1]$ abgebildet werden. Damit hat C dieselbe Mächtigkeit wie $[0, 1]$. C ist darüberhinaus 'perfekt' (vgl. Hewitt-Stromberg S. 70-71): nach Konstruktion der fallenden Folge $(C_k)_k$ findet man in jeder noch so kleinen Umgebung eines Punktes $x \in C$ unendlich viele weitere Punkte von C . Damit ist C eine abgeschlossene Menge, welche keine isolierten Punkte besitzt. Zusammenfassend gilt: C ist eine perfekte abgeschlossene Menge vom Lebesgue-Mass 0, und besitzt die Mächtigkeit des Kontinuums.

3) Betrachte nun für jedes $n \geq 0$ die Gleichverteilung ν_n auf C_n

$$\nu_n := \frac{1}{\lambda(C_n)} \lambda(\cdot \cap C_n)$$

und die zugehörige Verteilungsfunktion F_n . Notwendig ist F_n konstant auf C_n^c und hat auf den Segmenten von C_n konstante Steigung $(\frac{3}{2})^n$ wegen (+). Betrachtet man nun auf einem beliebigen Segment von C_n den Übergang von F_n zu F_{n+1} , so zeigt die Konstruktion aus Schritt 1):

$$\sup_{t \in [0,1]} |F_{n+1}(t) - F_n(t)| = \frac{1}{6} 2^{-n}, \quad n \geq 0.$$

Damit ist $(F_n)_n$ eine Cauchyfolge in $C([0,1])$. Wegen der Vollständigkeit von $C([0,1])$ gibt es ein $F \in C([0,1])$ so dass $(F_n)_n$ gleichmässig auf $[0,1]$ gegen F konvergiert. Notwendig ist F nichtfallend, und es gilt $F(0) = 0$ und $F(1) = 1$. Da F stetig, ist F nach 1.19 eine Verteilungsfunktion auf \mathbb{R} .

4) Der stetigen Verteilungsfunktion F ist in eindeutiger Weise ein Wahrscheinlichkeitsmass ν auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ zugeordnet: ν heisst Cantorverteilung auf $[0,1]$. Nach Konstruktion ist ν konzentriert auf die Lebesgue-Nullmenge C . Damit ist $\nu \perp \mathfrak{M}$ bewiesen. \square

Wir führen den Beweis des Hauptsatzes 3.4 (siehe 3.9) über zwei Hilfssätze.

3.7 Hilfssatz: (Existenz) Seien μ, ν zwei σ -endliche Masse auf (Ω, \mathcal{A}) . Dann gibt es eine \mathcal{A} -messbare Funktion $f : \mathbb{R} \rightarrow [0, \infty)$ und eine μ -Nullmenge $N \in \mathcal{A}$ so dass

$$\nu(A) = \nu(A \cap N) + \int_A f d\mu, \quad A \in \mathcal{A}.$$

Damit ist *eine* Zerlegung des Masses ν in einen μ -absolutstetigen

$$\nu^{ac}(A) := \int_A f d\mu, \quad A \in \mathcal{A}$$

und einen μ -singulären Anteil

$$\nu^s(A) := \nu(A \cap N), \quad A \in \mathcal{A}$$

gefunden.

Beweis: 1) Wir zeigen, dass man sich von σ -endlichen auf endliche Masse μ, ν zurückziehen kann: für μ, ν gibt es aufsteigende Folgen $M_n \uparrow \Omega, F_n \uparrow \Omega$ in \mathcal{A} so dass $\mu(M_n) < \infty, \nu(F_n) < \infty$ für alle $n \in \mathbb{N}$. Jede dieser Folgen kann durch $H_n := M_n \cap F_n, n \geq 1$, ersetzt werden. Setze $H_0 := \emptyset$. Dann ist $E_n := H_n \setminus H_{n-1}, n \geq 1$, eine Folge paarweise disjunkter Mengen in \mathcal{A} mit

$$\Omega = \bigcup_{n=1}^{\infty} E_n, \quad \mu(E_n) < \infty, \quad \nu(E_n) < \infty, \quad n \geq 1.$$

Kann man nun den Satz beweisen für $\mu_n := \mu(\cdot \cap E_n)$, $\nu_n := \nu(\cdot \cap E_n)$ anstelle der ursprünglichen μ , ν , mit \mathcal{A} -messbaren Funktionen $f_n : \Omega \rightarrow [0, \infty)$ und μ_n -Nullmengen $N_n \in \mathcal{A}$

$$\nu_n(A) = \nu_n(A \cap N_n) + \int_A f_n d\mu_n, \quad A \in \mathcal{A}, \quad n \in \mathbb{N},$$

so liefert 'Ausschneiden und Zusammenkleben'

$$f := \sum_{n=1}^{\infty} f_n 1_{E_n}, \quad N := \bigcup_{n=1}^{\infty} (N_n \cap E_n)$$

wegen der Disjunktheit der $(E_n)_n$ die Behauptung des Hilfssatzes für ν und μ .

2) Sei also ab jetzt $\mu(\Omega) < \infty$ und $\nu(\Omega) < \infty$ vorausgesetzt. Definiere

$$\mathcal{H} := \{g \in \mathcal{F}^+ : \text{es gilt } \int_A g d\mu \leq \nu(A) \text{ für alle } A \in \mathcal{A}\},$$

dann gilt: i) \mathcal{H} ist abgeschlossen unter Maximumsbildung: aus $g_1, g_2 \in \mathcal{H}$ folgt $g_1 \vee g_2 \in \mathcal{H}$, wegen

$$\begin{aligned} \int_A g_1 \vee g_2 d\mu &= \int_{A \cap \{g_1 < g_2\}} g_2 d\mu + \int_{A \cap \{g_1 \geq g_2\}} g_1 d\mu \\ &\leq \nu(A \cap \{g_1 < g_2\}) + \nu(A \cap \{g_1 \geq g_2\}) = \nu(A), \quad A \in \mathcal{A}; \end{aligned}$$

ii) \mathcal{H} ist abgeschlossen unter aufsteigender Konvergenz: für jede aufsteigende Folge $(g_n)_n \subset \mathcal{H}$ gilt $\sup_n g_n \in \mathcal{H}$, wegen monotoner Konvergenz 2.7 und nach Definition von \mathcal{H} .

Die Eigenschaften i) und ii) erlauben die Konstruktion eines maximalen Elements in der Klasse \mathcal{H} , d.h. eines $f \in \mathcal{H}$ mit der Eigenschaft

$$(+)$$

$$g \in \mathcal{H}, \quad g \geq f \quad \implies \quad \mu(\{g > f\}) = 0.$$

Dazu geht man aus von

$$\alpha := \sup_{g \in \mathcal{H}} \int_{\Omega} g d\mu \leq \nu(\Omega) < \infty.$$

Wegen i) und ii) können Folgen $(g_n)_n \subset \mathcal{H}$ mit $\lim_{n \rightarrow \infty} \int_{\Omega} g_n d\mu = \alpha$ stets aufsteigend gewählt werden, und für verschiedene Wahlen solcher Folgen müssen die jeweiligen oberen Einhüllenden $\bar{g} := \sup_{n \geq 1} g_n$ bis auf μ -Äquivalenz übereinstimmen. Damit kann ein maximales Element von \mathcal{H} durch

$$(*) \quad f := \bar{g} = \sup_{n \geq 1} g_n \quad \text{für eine beliebige aufsteigende Folge } (g_n)_n \in \mathcal{H} \text{ mit } \int_{\Omega} g_n d\mu \uparrow \alpha$$

festgelegt werden: verschiedene Festlegungen einer messbaren Funktion f wie in (*) stimmen μ -fast sicher überein, und jede Festlegung hat die Eigenschaft (+). Man nennt f eine Festlegung des

μ -wesentlichen Supremums von \mathcal{H} . Damit kann man ' μ -absolutstetige Anteile' von ν in optimaler Weise ausschöpfen, denn wegen (+) leistet jede Festlegung f des μ -wesentlichen Supremums:

$$(++) \quad \int_A f d\mu \leq \nu(A) \quad \text{und} \quad \int_A g d\mu \leq \int_A f d\mu \quad \text{für alle } g \in \mathcal{H}, \text{ alle } A \in \mathcal{A}.$$

Da ν ein endliches Mass ist, muss $\{f = +\infty\}$ eine μ -Nullmenge in \mathcal{A} sein. Insbesondere kann f dann stets durch $\tilde{f} := f1_{\{f < \infty\}}$ ersetzt werden: folglich kann f in (+) und (++) gleich in Form

$$f : \Omega \longrightarrow [0, \infty) \quad \mathcal{A}\text{-messbar}$$

angesetzt werden. Ab jetzt bezeichnet f eine solche Festlegung des μ -wesentlichen Supremums von \mathcal{H}

3) Mit f aus Schritt 2) definiere Mengenfunktionen

$$(**) \quad \nu^{ac}(A) := \int_A f d\mu, \quad A \in \mathcal{A}$$

$$(***) \quad \nu^s(A) := \nu(A) - \int_A f d\mu, \quad A \in \mathcal{A}.$$

Dies liefert eine Zerlegung $\nu = \nu^{ac} + \nu^s$. Sicher ist ν^{ac} nach 3.1 ein Mass auf (Ω, \mathcal{A}) . Damit ist ν^s einerseits Differenz zweier endlicher Masse (d.h. wohldefinierte Differenz zweier σ -additiver Mengenfunktionen) und andererseits *nichtnegative* Mengenfunktion auf (Ω, \mathcal{A}) , nach Definition der Klasse \mathcal{H} . Also ist ν^s ein endliches Mass auf (Ω, \mathcal{A}) .

Dagegen ist noch nicht gezeigt, dass die Masse ν^s und ν^{ac} zueinander singulär sind. In 4) bis 6) unten werden wir daher eine μ -Nullmenge $N \in \mathcal{A}$ konstruieren, die ein Träger des Masses ν^s ist: dann gilt

$$\nu^s(A) = \nu^s(A \cap N) \quad \text{und} \quad \nu^{ac}(A \cap N) = \int_{A \cap N} f d\mu \leq \int_N f d\mu = 0$$

zusammen mit

$$\nu^s(A) = \nu^s(A \cap N) = [\nu - \nu^{ac}](A \cap N) = \nu(A \cap N)$$

wegen (**) und (***) für alle $A \in \mathcal{A}$. Damit ist das Mass ν^s in (***) bestimmt zu

$$\nu^s = \nu(\cdot \cap N) \quad \text{auf } (\Omega, \mathcal{A})$$

und der Beweis des Hilfssatzes ist abgeschlossen.

4) Für $\varepsilon > 0$ und $M \in \mathcal{A}$ setze

$$\mathcal{G}_{\varepsilon, M} := \{ B \in \mathcal{A} : B \subset M, \nu^s(B) < \varepsilon \mu(B) \}.$$

Wir beweisen zuerst

$$(\times) \quad \mu(M) > 0 \implies \mathcal{G}_{\varepsilon, M} \neq \emptyset \quad \text{für jedes } \varepsilon > 0.$$

Indirekt: wäre dies nicht so, könnte man ein $M \in \mathcal{A}$ mit $\mu(M) > 0$ und ein $\varepsilon > 0$ finden, so dass die Abschätzung $\nu^s(B) \geq \varepsilon \mu(B)$ für *alle* Teilmengen $B \subset M$, $B \in \mathcal{A}$ gültig wäre, also

$$\int_A (\varepsilon \cdot 1_M) d\mu = \varepsilon \mu(A \cap M) \leq \nu^s(A \cap M) \leq \nu(A) - \int_{A \cap M} f d\mu, \quad A \in \mathcal{A}$$

nach (**). Mit dieser Abschätzung wäre aber auch $(f + \varepsilon)1_M$ in der Klasse \mathcal{H} , im Widerspruch zur Wahl von f als μ -wesentliches Supremum in (*). Also ist (\times) bewiesen.

5) Wir zeigen als nächstes: für jedes $n \geq 1$ gibt es ein Ereignis $C_n \in \mathcal{A}$ mit der Eigenschaft

$$(\times \times) \quad \mu(C_n^c) = 0 \quad \text{und} \quad \nu^s(C_n) < 2^{-n} \mu(C_n).$$

Dies sieht man so. Nutze in einem ersten Schritt (\times) mit $M = \Omega$ und $\varepsilon = 2^{-n}$, um eine 'möglichst grosse' Menge $K_{n,1} \in \mathcal{G}_{2^{-n}, \Omega}$ auszuwählen, die der Ungleichung

$$\nu^s(K_{n,1}) < 2^{-n} \mu(K_{n,1})$$

genügt. Ist $\mu(K_{n,1}^c)$ noch strikt positiv, nutze (\times) in einem zweiten Schritt mit $M = K_{n,1}^c$ und $\varepsilon = 2^{-n}$, um eine 'möglichst grosse' Teilmenge $K_{n,2} \in \mathcal{G}_{2^{-n}, K_{n,1}^c}$ auszuwählen, die der Ungleichung

$$\nu^s(K_{n,2}) < 2^{-n} \mu(K_{n,2})$$

genügt. Insbesondere ist $K_{n,2}$ eine Teilmenge von $K_{n,1}^c$, also gilt

$$\nu^s(K_{n,1} \dot{\cup} K_{n,2}) < 2^{-n} \mu(K_{n,1} \dot{\cup} K_{n,2}).$$

Ist $\mu((K_{n,1} \dot{\cup} K_{n,2})^c)$ noch strikt positiv, nutze (\times) in einem dritten Schritt mit $M = (K_{n,1} \dot{\cup} K_{n,2})^c$, und so weiter: solange verbleibende Komplemente $(K_{n,1} \dot{\cup} \dots \dot{\cup} K_{n,i-1})^c$ noch strikt positive μ -Masse aufweisen, kann man mit einem i -ten Schritt das Verfahren fortsetzen und ein 'möglichst grosses' $K_{n,i} \in \mathcal{G}_{2^{-n}, (K_{n,1} \dot{\cup} \dots \dot{\cup} K_{n,i-1})^c}$ auswählen so dass

$$\nu^s(K_{n,1} \dot{\cup} \dots \dot{\cup} K_{n,i}) < 2^{-n} \mu(K_{n,1} \dot{\cup} \dots \dot{\cup} K_{n,i})$$

gilt. Im Ergebnis erhält man eine Familie disjunkter $K_{n,i}$ mit der Eigenschaft

$$C_n := \bigcup_i K_{n,i} \quad \text{hat die Eigenschaft} \quad \nu^s(C_n) < 2^{-n} \mu(C_n).$$

Zum vollständigen Nachweis von ($\times \times$) müssen wir noch nachweisen, dass C_n^c eine μ -Nullmenge ist.

Dazu präzisieren wir die oben verwendete Formulierung 'möglichst gross' im i -ten Schritt zu

$$(\diamond) \quad M_i := (K_{n,1} \dot{\cup} \dots \dot{\cup} K_{n,i-1})^c, \quad \gamma_i := \sup \{ \mu(B) : B \in \mathcal{G}_{2^{-n}, M_i} \}, \quad \mu(K_{n,i}) > 0.9 \cdot \gamma_i$$

und vergleichen unter Beachtung von $M_{i+1} \dot{\cup} K_{n,i} = M_i$ die Klassen $\mathcal{G}_{2^{-n}, M_{i+1}}$ und $\mathcal{G}_{2^{-n}, M_i}$.

Sei $\tilde{B} \in \mathcal{G}_{2^{-n}, M_{i+1}}$ beliebig. Dann sind \tilde{B} und $K_{n,i}$ notwendig disjunkt, es gilt

$$\nu^s(K_{n,i}) < 2^{-n} \mu(K_{n,i}) \quad \text{und} \quad \nu^s(\tilde{B}) < 2^{-n} \mu(\tilde{B})$$

und damit

$$\nu^s(\tilde{B} \dot{\cup} K_{n,i}) < 2^{-n} \mu(\tilde{B} \dot{\cup} K_{n,i}),$$

folglich gehört $\tilde{B} \dot{\cup} K_{n,i}$ zum System $\mathcal{G}_{2^{-n}, M_i}$. Die 'möglichst grosse' Wahl von $K_{n,i}$ im Sinne von (\diamond) stellt sicher, dass die μ -Masse von \tilde{B} höchstens noch $0.1 \cdot \gamma_i$ betragen kann. Da \tilde{B} in $\mathcal{G}_{2^{-n}, M_{i+1}}$ beliebig war, zeigt dies $\gamma_{i+1} < 0.1 \cdot \gamma_i$ für alle i . Insbesondere ist die Folge $(\gamma_i)_i$ eine Nullfolge.

Angenommen, die Menge $C_n^c = \bigcap_i M_i$ wäre keine μ -Nullmenge. Dann wäre wegen (\times) das System $\mathcal{G}_{2^{-n}, C_n^c}$ nichtleer, es existierte also ein $\tilde{B} \in \mathcal{A}$ mit den Eigenschaften

$$\tilde{B} \subset C_n^c, \quad \nu^s(\tilde{B}) < 2^{-n} \mu(\tilde{B}), \quad \text{insbesondere} \quad \mu(\tilde{B}) \neq 0.$$

Die Mengenfolge $(M_i)_i$ ist absteigend, also müsste dieses \tilde{B} mit $\mu(\tilde{B}) \neq 0$ in jedem der Systeme $\mathcal{G}_{2^{-n}, M_i}$ enthalten sein, im Widerspruch dazu, dass $(\gamma_i)_i$ aus (\diamond) bereits als Nullfolge nachgewiesen wurde: also war die Annahme absurd. Folglich ist C_n^c eine μ -Nullmenge, und $(\times \times)$ ist vollständig bewiesen.

6) Mit $(\times \times)$ ist insbesondere eine Folge $(C_n)_n \subset \mathcal{A}$ gefunden mit den Eigenschaften

$$(\diamond \diamond) \quad \mu(C_n^c) = 0, \quad \nu^s(C_n) < 2^{-n} \mu(\Omega), \quad n \geq 1.$$

Für beliebiges m kann $\bigcup_{n \geq m} C_n$ wegen der ersten Aussage in $(\diamond \diamond)$ als Träger von μ aufgefasst werden. Diese Mengenfolge ist absteigend in m , wobei die zweite Aussage in $(\diamond \diamond)$ zeigt:

$$\nu^s\left(\bigcup_{n \geq m} C_n\right) \longrightarrow 0 \quad \text{für} \quad m \rightarrow \infty.$$

Mit absteigender Stetigkeit unter beiden Massen ν^s und μ sieht man nun: die Menge

$$C := \limsup_n C_n = \bigcap_m \bigcup_{n \geq m} C_n$$

ist einerseits eine Menge von vollem μ -Mass, und andererseits eine ν^s -Nullmenge. Mit $N := C^c$ ist nun die am Ende von Schritt 3) behauptete μ -Nullmenge konstruiert. \square

Nenne unter den Voraussetzungen des Hilfssatzes 3.7 jedes dort konstruierte Paar (f, N) :

$$(L) \quad \begin{cases} f : \Omega \rightarrow [0, \infty) \text{ } \mathcal{A}\text{-messbar} \\ N \text{ eine } \mu\text{-Nullmenge in } \mathcal{A} \\ \nu(A) = \nu(A \cap N) + \int_A f d\mu, \quad A \in \mathcal{A} \end{cases}$$

eine Lebesgue-Zerlegung von ν bezüglich μ . Es gilt die folgende Eindeutigkeitseigenschaft:

3.8 Hilfssatz: Für je zwei Lebesgue-Zerlegungen $(f_1, N_1), (f_2, N_2)$ von ν bezüglich μ in 3.7 gilt

$$\nu(N_1 \Delta N_2) = 0 \quad \text{und} \quad \mu(\{f_1 \neq f_2\}) = 0.$$

Beweis: Gelte (L) sowohl für (f_1, N_1) als auch für (f_2, N_2) . Aus der dritten Eigenschaft in (L) folgt

$$\int_{A \cap N_1^c \cap N_2^c} f_1 d\mu = \nu(A \cap N_1^c \cap N_2^c) = \int_{A \cap N_1^c \cap N_2^c} f_2 d\mu, \quad A \in \mathcal{A} :$$

damit sind nach 2.11' die Funktionen $f_i 1_{N_1^c \cap N_2^c}, i = 1, 2$, μ -äquivalent. Weiter

$$\{f_1 \neq f_2\} \subset \{f_1 1_{N_1^c \cap N_2^c} \neq f_2 1_{N_1^c \cap N_2^c}\} \cup N_1 \cup N_2$$

wobei $N_1 \cup N_2$ nach (L) eine μ -Nullmenge ist: also hat man $\mu(\{f_1 \neq f_2\}) = 0$. Sind aber f_1, f_2 μ -äquivalent, so schreibt sich die dritte Eigenschaft in (L) für $A = N_2^c$ in der Form

$$\begin{aligned} \nu(N_2^c \cap N_1) &= \nu(N_2^c) - \int_{N_2^c} f_1 d\mu \\ &= \nu(N_2^c) - \int_{N_2^c} f_2 d\mu = \nu(N_2^c \cap N_2) = 0 : \end{aligned}$$

also ist $N_2^c \cap N_1$ eine ν -Nullmenge. Mit vertauschten Rollen der Indices $i = 1, 2$ ist auch $N_1^c \cap N_2$ eine ν -Nullmenge. Damit ist die symmetrische Differenz $N_1 \Delta N_2$ eine ν -Nullmenge. \square

3.9 Beweis des Hauptsatzes 3.4 und des Korollars 3.5: Mit den Hilfssätzen 3.7 und 3.8 sind die Aussagen a) und b) des Hauptsatzes 3.4 bewiesen; die Aussage c) in 3.4 folgt aus a) und b). Damit ist 3.4 nun vollständig bewiesen, und 3.5 ergibt sich also sofortige Folgerung: Wird ν durch μ dominiert, so ist die μ -Nullmenge N in (L) auch eine ν -Nullmenge. \square

Wie berechnet man aber Lebesgue-Zerlegungen explizit? Wir formulieren dies hier für Wahrscheinlichkeitsmasse, wegen der Wichtigkeit von Dichten oder Dichtequotienten in der Statistik; analoge

Aussagen gelten aber auch für σ -endliche Masse.

3.10 Satz: Betrachte zwei Wahrscheinlichkeitsmasse P und Q auf einem messbaren Raum (Ω, \mathcal{A}) . Wähle ein σ -endliches Mass μ auf (Ω, \mathcal{A}) so dass

$$P \ll \mu, \quad Q \ll \mu$$

(ein solches existiert stets, eine mögliche Wahl ist $\mu := P + Q$). Wähle nach 3.5 μ -Dichten

$$p = \frac{dP}{d\mu}, \quad q = \frac{dQ}{d\mu}.$$

Dann liefert

$$(f, N) : \quad f := \frac{p}{q} 1_{\{q>0\}}, \quad N := \{q = 0\}$$

eine Lebesgue-Zerlegung von P bezüglich Q .

Beweis: Für beliebiges $A \in \mathcal{A}$ schreibt man

$$\begin{aligned} P(A) &= P(A \cap \{q = 0\}) + P(A \cap \{q > 0\}) \\ &= P(A \cap \{q = 0\}) + \int 1_A 1_{\{q>0\}} \frac{p}{q} \cdot q \, d\mu \\ &= P(A \cap N) + \int_A f \, dQ, \end{aligned}$$

und es gilt

$$Q(N) = \int 1_N q \, d\mu = \int 1_{\{q=0\}} q \, d\mu = 0. \quad \square$$

3.10' Beispiel: Betrachte mit den Notationen aus 3.10 zwei Gleichverteilungen

$$P := \mathcal{R}(0, 2), \quad Q := \mathcal{R}(-1, 1) \quad \text{auf } (\mathbb{R}, \mathcal{B}(\mathbb{R}))$$

und wähle das Lebesguemass $\lambda := \mu$ als dominierendes Mass. Dann hat man μ -Dichten

$$p := \frac{dP}{d\lambda} = \frac{1}{2} \cdot 1_{(0,2)} \quad , \quad q := \frac{dQ}{d\lambda} = \frac{1}{2} \cdot 1_{(-1,1)}$$

und schreibt nach 3.10 eine Lebesgue Zerlegung von P bezüglich Q in der Form

$$(f, N) : \quad N := \mathbb{R} \setminus (-1, 1) = \{q = 0\} \quad , \quad f := 0 \cdot 1_{(-1,0]} + 1 \cdot 1_{(0,1)} = \frac{p}{q} 1_{\{q>0\}}.$$

Im Rahmen der Eindeutigkeitsaussagen aus 3.4 sind ausserhalb des Trägers von $P + Q$ andere Festlegungen von N und ausserhalb des Trägers von Q andere Festlegungen von f möglich: wie man sofort nachprüft, ist bei festem $c > 0$ auch

$$(\tilde{f}, \tilde{N}) : \quad \tilde{N} := [1, 2) \quad , \quad \tilde{f} := 0 \cdot 1_{(-1,0]} + 1 \cdot 1_{(0,1)} + c \cdot 1_{\mathbb{R} \setminus (-1,1)}$$

eine Lebesgue Zerlegung von P bezüglich Q . □

Die griffigste Formulierung der Eindeutigkeitsaussage erhält man, wenn man die Lebesgue-Zerlegung in Form eines Dichtequotienten ansetzt. Dies ist konstruktiv im Sinne von 3.10.

3.11 Satz: Seien P und Q Wahrscheinlichkeitsmasse auf (Ω, \mathcal{A}) .

a) *Dichtequotient (Likelihood Ratio) von P bez. Q* heisst jede \mathcal{A} -messbare Abbildung $L : \Omega \rightarrow [0, \infty]$ mit folgender Eigenschaft: die Definition

$$(DQ1) \quad f := L 1_{\{L < \infty\}} \quad , \quad N := \{L = +\infty\}$$

liefert eine Lebesgue-Zerlegung

$$(DQ2) \quad Q(\{L = \infty\}) = 0 \quad , \quad P(A) = P(A \cap \{L = \infty\}) + \int_A L dQ \quad , \quad A \in \mathcal{A}$$

von P bezüglich Q .

b) Abbildungen $L : \Omega \rightarrow [0, \infty]$ mit den Eigenschaften (DQ1)+(DQ2) existieren stets.

c) Verschiedene Festlegungen des Dichtequotienten stimmen $(P + Q)$ -fast sicher überein.

Beweis: 1) Zum Beweis der Existenz wähle nach 3.10 ein dominierendes Mass μ und μ -Dichten $p = \frac{dP}{d\mu}$, $q = \frac{dQ}{d\mu}$, und definiere

$$L : \Omega \rightarrow [0, \infty] \quad , \quad L := \frac{p}{q} 1_{\{q > 0\}} + \infty 1_{\{q = 0\}} .$$

Nach 3.5 sind Dichten p, q Funktionen $\Omega \rightarrow [0, \infty)$, also ist die MNF L auf $\{q > 0\}$ endlich. Dies identifiziert $\{0 \leq L < \infty\}$ mit $\{q > 0\}$ und $\{L = \infty\}$ mit $\{q = 0\}$. Satz 3.10 zeigt (DQ1) und (DQ2).

2) Zum Beweis der Eindeutigkeit seien \mathcal{A} -messbare Funktionen $L_1, L_2 : \Omega \rightarrow [0, \infty]$ mit den Eigenschaften (DQ1) und (DQ2) gegeben. Definiere

$$f_1 := L_1 1_{\{L_1 < \infty\}} \quad , \quad f_2 := L_2 1_{\{L_2 < \infty\}} .$$

Dann sind $f_1, f_2 : \Omega \rightarrow [0, \infty)$ Festlegungen der Dichte des Q absolutstetigen Anteils von P wie in 3.4, und die Eindeutigkeitsaussagen in 3.4 implizieren

$$\left\{ \begin{array}{l} \{L_1 = \infty\} , \{L_2 = \infty\} \text{ sind } Q\text{-Nullmengen} , \\ P(\{L_1 = \infty\} \Delta \{L_2 = \infty\}) = 0 , \\ f_1 , f_2 \text{ stimmen } Q\text{-fast sicher überein} . \end{array} \right.$$

Trivialerweise gilt dann aber auch

$$(P + Q) (\{L_1 = \infty\} \Delta \{L_2 = \infty\}) = 0.$$

Weiter kann der Q -absolutstetige Anteil P^{ac} von P

$$P^{ac}(A) = P(A \cap \{L_1 < \infty\}) = P(A \cap \{L_2 < \infty\})$$

geschrieben werden als

$$P^{ac}(A) = \int_A f_1 dQ = \int_A f_2 dQ, \quad A \in \mathcal{A}.$$

Nach 3.4 stimmen die Funktionen f_1 und f_2 Q -fast sicher überein. Wegen $P^{ac} \ll Q$ stimmen sie dann erst recht $(Q + P^{ac})$ -fast sicher überein. Es bleibt der Q -singuläre Anteil P^s von P zu betrachten, der auf $\{L_1 = \infty\} \cup \{L_2 = \infty\}$ konzentriert ist, aber für diesen sind $f_1 = L_1 1_{\{L_1 < \infty\}}$ und $f_2 = L_2 1_{\{L_2 < \infty\}}$ äquivalent zu 0. Also gilt

$$L_1 1_{\{L_1 < \infty\}} = L_2 1_{\{L_2 < \infty\}} \quad (Q + P)\text{-fast sicher}$$

und damit die Behauptung. □

3.11' Beispiel: Wir setzen Beispiel 3.10' fort und betrachten die Gleichverteilungen

$$P := \mathcal{R}(0, 2) \quad , \quad Q := \mathcal{R}(-1, 1) \quad \text{auf } (\mathbb{R}, \mathcal{B}(\mathbb{R})).$$

Mit den Notationen aus 3.10' kann der Dichtequotient von P bezüglich Q durch

$$L := 0 \cdot 1_{(-1, 0]} + 1 \cdot 1_{(0, 1)} + \infty \cdot 1_{\mathbb{R} \setminus (-1, 1)}$$

festgelegt werden; andere Festlegungen stimmt bis auf $(P + Q)$ -Nullmengen mit dieser überein. □

Am Schluss dieses Kapitels sei an die Transformationsformel für Dichten erinnert:

3.12 Transformationsformel: Betrachte $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \mathbb{N})$ für beliebiges $d \geq 1$. Seien U, V offen in \mathbb{R}^d , sei $g : \mathbb{R}^d \rightarrow \mathbb{R}^d$ ein Diffeomorphismus zwischen U und V (d.h.: g ist bijektiv, sowohl $g : U \rightarrow V$ als auch die Umkehrabbildung $g^{-1} : V \rightarrow U$ sind stetig differenzierbar), und sei

$$J(y) := \begin{pmatrix} D_1(g_1^{-1})(y) & \dots & D_d(g_1^{-1})(y) \\ \dots & \dots & \dots \\ D_1(g_d^{-1})(y) & \dots & D_d(g_d^{-1})(y) \end{pmatrix}, \quad y \in V$$

die Jacobimatrix zu g^{-1} (g_k^{-1} bezeichnet die k -te Komponente von g^{-1} , und D_j die partielle Ableitung nach dem j -ten Argument).

a) Sei P ein Wahrscheinlichkeitsmass auf $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ mit Lebesgue-Dichte $p = \frac{dP}{d\lambda}$, sei P konzentriert auf U . Das Bildmass $Q = P^g$ von P unter g ist dann konzentriert auf V , und hat die Lebesgue-Dichte

$$q(y) := 1_V(y) p(g^{-1}(y)) |\det J(y)|, \quad y \in \mathbb{R}^d.$$

b) Ist X eine \mathbb{R}^d -wertige Zufallsvariable auf einem Grundraum $(\Omega_0, \mathcal{A}_0, P_0)$ mit $\mathcal{L}(X|P_0) = P$ wie in a), so ist $q(\cdot)$ wie in a) die Dichte von $Q = \mathcal{L}(g(X) | P_0)$.

Beweis: Da die Aussagen a) und b) gleichwertig sind, reicht ein Beweis in der Form b): betrachte $W \subset \mathbb{R}^d$ offen mit $W \subset V$, dann

$$\begin{aligned} P_0(g(X) \in W) &= P_0(X \in g^{-1}(W)) = \int_{g^{-1}(W)} p(x) dx \\ &= \int_W p(g^{-1}(y)) |\det J(y)| dy = \int_W q(y) dy. \end{aligned}$$

Dies ist die wohlbekannte Transformationsformel der Analysis (z.B. Barner und Flohr, Analysis, Band 2, 1983, S. 320). Die offenen Mengen erzeugen $\mathcal{B}(\mathbb{R}^d)$, und es gilt $P_0(g(X) \in V^c) = 0$. Die Masse $Q = P^g = \mathcal{L}(g(X)|P_0)$ und $A \rightarrow \int_A q d\lambda$ stimmen also auf $\mathcal{B}(\mathbb{R}^d)$ überein. \square