HANDOUT #1
KONVERGENZ VON W-MASSEN, VERTEILUNGEN UND ZUFALLSVARIABLEN

FOLKMAR BORNEMANN

Die Vielfalt der Konvergenzberiffe für W-Maße, Verteilungen und Zufallsvariablen (wie vage, schwache, stochastische und fast sichere Konvergenz) kann schnell zu Verwirrung (das englische Wort „bewildering“ trifft es perfekt) führen, insbesondere wenn sie wie in unserer Vorlesung „gestapelt“ werden. Die folgende Zusammenstellung zeigt, wie ich für mich selbst Ordnung in die Angelegenheit bringe und welche Resultate ich grundsätzlich wichtig finde; sie dient mir selbst zum Nachschlagen.

1. Topologische Masse und vage Konvergenz

Jede sinnvolle Konvergenzbetrachtung von W-Maßen verlangt nach einer topologischen Struktur des zugrundeliegenden Wahrscheinlichkeitsraums. Dazu betrachtet man Hausdorff-Räume[1] $\mathcal{X}$, die von den offenen Mengen erzeugte σ-Algebra $\mathcal{B}(\mathcal{X})$ der borelschen Mengen und „reguläre“ Maße (also solche Maße, die in einem gewissen Sinne mit der topologischen Struktur verträglich sind). Ein Maß $\mu : \mathcal{B}(\mathcal{X}) \to [0, \infty]$ heißt dabei

- *lokalendlich* oder *Borel-Maß*, falls jedes $x \in \mathcal{X}$ eine offene Umgebung $U \ni x$ mit $\mu(U) < \infty$ besitzt; Elstrodt (2005, Def. VIII.1.1)
- *regulär von innen*, falls
$$\mu(A) = \sup\{\mu(K) : K \subset A \text{ kompakt}\} \qquad (A \in \mathcal{B}(\mathcal{X}));$$
- *regulär von außen*, falls
$$\mu(A) = \inf\{\mu(U) : U \supset A \text{ offen}\} \qquad (A \in \mathcal{B}(\mathcal{X}));$$
- *regulär*, falls μ von innen und außen regulär ist;
- *Radon-Maß*, falls μ ein von innen reguläres Borel-Maß ist.

Um mühevolle Einzelprüfungen dieser Eigenschaften zu vermeiden, betrachtet man in der W-Theorie gerne topologische Räume, auf denen Borel-Maße *grundsätzlich* regulär sind.

Definition 1.1. Separable, vollständig metrisierbare Räume $\mathcal{X}$ heißen *polnisch*. Elstrodt (2005, §VIII.1.5)

Beispiele sind die euklidischen Räume $\mathbb{R}^n$ (die zusätzlich auch noch lokalkompakt sind). Im folgenden sprechen wir kurz von Maßen auf dem topologischen Raum $\mathcal{X}$, wenn wir solche auf dem borelschen Meßraum $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ meinen.

Date: 26. April 2010.

[1] Auf diesen topologischen Räumen sind kompakte Mengen abgeschlossen und daher borelsch.

Satz 1.2 (Ulam 1939)**.** *Borel-Maße auf polnischen Räumen sind σ-endlich und regulär.*

Elstrodt (2005, Satz VIII.1.16)

Für polnische Räume $\mathcal{X}$ führen wir folgende Mengen von Maßen ein:

$$\mathcal{M}(\mathcal{X}) = \{\mu : \mu \text{ Borel'sches Maß auf } \mathcal{X}\};$$
$$\mathcal{M}_1(\mathcal{X}) = \{\mu : \mu \text{ W-Maß auf } \mathcal{X}\};$$
$$\mathcal{M}_{\leqslant 1}(\mathcal{X}) = \{\mu : \mu \text{ Maß auf } \mathcal{X} \text{ mit } \mu(\mathcal{X}) \leqslant 1\}.$$

Die Maße in $\mathcal{M}_{\leqslant 1}(\mathcal{X})$ heißen auch *Sub-Wahrscheinlichkeitsmaße*. Es gilt

$$\mathcal{M}_1(\mathcal{X}) \subset \mathcal{M}_{\leqslant 1}(\mathcal{X}) \subset \mathcal{M}(\mathcal{X}).$$

Ferner führen wir noch die Menge der signierten Maße auf $\mathcal{X}$ ein:

$$\mathcal{M}_s(\mathcal{X}) = \{\mu : \mu \text{ signiertes Maß auf } \mathcal{X}\};$$

mit der Totalvariationsnorm ist dies ein Banachraum. Wir benötigen noch einige Räume stetiger Funktionen, auf denen Maße operieren und mit deren Hilfe wir „schwache“ Konzepte einer Konvergenz von Maßen testen werden:

$$C_b(\mathcal{X}) = \{f \in C(\mathcal{X}) : \ f \text{ ist beschränkt}\};$$
$$C_c(\mathcal{X}) = \{f \in C(\mathcal{X}) : \ f \text{ hat kompakten Träger}\};$$
$$C_0(\mathcal{X}) = \text{ Vervollständigung von } C_c(\mathcal{X}) \text{ bezüglich Supremumsnorm.}$$

Mit der Supremumsnorm versehen sind $C_b(\mathcal{X})$ und $C_0(\mathcal{X})$ Banachräume; es gilt

$$C_0(\mathcal{X}) \subset C_b(\mathcal{X}).$$

Der Anwendung funktionalanalytischer Methoden dient der folgende Satz.

Satz 1.3 (Darstellungssatz von Riesz)**.** *Es sei $\mathcal{X}$ ein lokalkompakter polnischer Raum. Dann ist*

Elstrodt (2005, Satz VIII.2.26)

$$C_0(\mathcal{X})' \simeq \mathcal{M}_s(\mathcal{X})$$

vermittels des ordnungstreuen isometrischen Isomorphismus $\Phi : \mathcal{M}_s(\mathcal{X}) \to C_0(\mathcal{X})'$ definiert durch

$$\langle \Phi(\mu), f \rangle = \int_{\mathcal{X}} f \, d\mu \qquad (\mu \in \mathcal{M}_s(\mathcal{X}), f \in C_0(\mathcal{X})).$$

Hierbei (d.h. für lokalkompakte polnische Räume $\mathcal{X}$) ist der Prädualraum $C_0(\mathcal{X})$ separabel.

Die Separabilität findet sich in Bauer (1978, Sätze 44.1 u. 44.3)

In Zukunft werden wir ein (signiertes) Maß μ daher mit der induzierten Linearform $\Phi(\mu)$ kurzerhand identifizieren und schreiben

$$\langle \mu, f \rangle = \int_{\mathcal{X}} f \, d\mu.$$

Wegen der Separabilität des Prädualraums ist nach dem Satz von Banach-Alaoglu die Einheitskugel von $\mathcal{M}_s(\mathcal{X})$ – und damit erst recht die Menge der Sub-W-Maße $\mathcal{M}_{\leqslant 1}(\mathcal{X})$ – schwach-* folgenkompakt. Beachte, dass es wegen der Beschränktheit von Sub-W-Maßen ausreicht, die schwach-* Konvergenz solcher Maße mit Funktionen aus der dichten Teilmenge $C_c(\mathcal{X})$ von $C_0(\mathcal{X})$ zu testen. In der Maßtheorie bezeichnet man diese Form der schwach-* Konvergenz allgemein als *vag*.

Definition 1.4. Ein Folge $\mu_n \in \mathcal{M}_{\leqslant 1}(\mathcal{X})$ von Sub-W-Maßen konvergiert *vag* gegen das Sub-W-Maß $\mu \in \mathcal{M}_{\leqslant 1}(\mathcal{X})$, falls für $n \to \infty$

Elstrodt (2005, Def. VIII.4.7)

$$\langle \mu_n, f \rangle \to \langle \mu, f \rangle \qquad (f \in C_c(\mathcal{X}));$$

in Zeichen

$$\mu_n \xrightarrow{v} \mu \qquad (n \to \infty).$$

Aus der Äquivalenz mit der schwach-* Konvergenz folgt auch, dass der vage Grenzwert *eindeutig* bestimmt ist.

Satz 1.5 (Auswahlsatz von Helly). *Jede Folge μ_n von Sub-W-Maßen auf einem lokalkompakten polnischen Raum $\mathcal{X}$ besitzt eine vag konvergente Teilfolge.*

Elstrodt (2005, Satz VIII.4.16)

Für W-Maße ist die vage Konvergenz für sich allein in der Regel zu schwach, um zu garantieren, dass der Limes wieder ein W-Maß ist. Zum Beispiel gilt auf $\mathcal{X} = \mathbb{R}$, dass $\delta_n \xrightarrow{v} 0$. Im allgemeinen folgt aus $\mu_n \xrightarrow{v} \mu$ nämlich nur

$$\mu(\mathcal{X}) \leqslant \liminf_{n\to\infty} \mu_n(\mathcal{X}).$$

2. Schwache Konvergenz von Massen

Definition 2.1. Es sei $\mathcal{X}$ ein polnischer Raum. Die Folge $\mu_n \in \mathcal{M}_{\leqslant 1}(\mathcal{X})$ konvergiert *schwach* gegen das Maß $\mu \in \mathcal{M}_{\leqslant 1}(\mathcal{X})$, falls für $n \to \infty$

Elstrodt (2005, Def. VIII.4.5)

$$\langle \mu_n, f \rangle \to \langle \mu, f \rangle \qquad (f \in C_b(\mathcal{X}));$$

in Zeichen

$$\mu_n \xrightarrow{w} \mu \qquad (n \to \infty).$$

Der Grenzwert einer schwach konvergenten Folge ist *eindeutig* bestimmt. Auf kompakten metrischen Räumen $\mathcal{X}$ fallen vage und schwache Konvergenz von Maßen zusammen, da dann $C_b(\mathcal{X}) = C_c(\mathcal{X})$ ist.

Wegen $\mathcal{M}_{\leqslant 1}(\mathcal{X}) \subset C_b(\mathcal{X})'$ (isometrisch eingebettet) ist die schwache Konvergenz äquivalent zur schwach-* Konvergenz in $C_b(\mathcal{X})'$. Da der Prädualraum $C_b(\mathcal{X})$ hier im allgemeinen nicht separabel ist, folgt daraus aber nicht viel; insbesondere lassen sich anders als bei der vagen Konvergenz die untenstehenden Kompaktheitsresultate nicht aus allgemeinen funktionalanalytischen Resultaten gewinnen.

Der schwache Limes von W-Maßen ist wiederum ein W-Maß: Aus $1 \in C_b(\mathcal{X})$ folgt nämlich

$$1 = \mu_n(\mathcal{X}) = \langle \mu_n, 1 \rangle \to \langle \mu, 1 \rangle = \mu(\mathcal{X}),$$

also $\mu(\mathcal{X}) = 1$. Einige Charakterisierungen der schwachen Konvergenz werden gerne im folgenden Theorem gebündelt („portmanteau“ (engl.) = Handkoffer; wobei dieser Handkoffer bei verschiedenen Autoren gerne unterschiedlich gefüllt ist). Dabei steht U_f für die Menge der Unstetigkeitspunkte einer Funktion $f : \mathcal{X} \to \mathbb{R}$; diese Menge ist (unabhängig von der Messbarkeit von f) stets Borel-messbar.

Billingsley (1995, Fußnote auf S. 334)

Satz 2.2 (Portmanteau-Theorem). *Es sei $\mathcal{X}$ ein polnischer Raum und $\mu, \mu_n \in \mathcal{M}_{\leqslant 1}(\mathcal{X})$. Dann sind äquivalent:*

Klenke (2008, Thm. 13.16)

(i) $\mu_n \xrightarrow{w} \mu$ *für* $n \to \infty$;

(ii) $\langle \mu_n, f \rangle \to \langle \mu, f \rangle$ *für alle beschränkten stetigen* f;

(iii) $\langle \mu_n, f\rangle \to \langle \mu, f\rangle$ *für alle beschränkten, Lipschitz-stetigen* f;

(iv) $\langle \mu_n, f\rangle \to \langle \mu, f\rangle$ *für alle beschränkten, messbaren* f *mit* $\mu(U_f) = 0$;

(v) $\mu_n(A) \to \mu(A)$ *für alle messbaren* A *mit* $\mu(\partial A) = 0$.

Ist $\mathcal{X}$ *zusätzlich lokalkompakt, so ist zudem mit (i)–(v) äquivalent:*

(vi) $\mu_n \xrightarrow{v} \mu$ *und* $\mu(\mathcal{X}) \geqslant \limsup_{n\to\infty} \mu_n(\mathcal{X})$.

Für die Betrachtung zufälliger Maße (etwa empririscher Verteilungen) ist es hilfreich zu wissen, dass sich die schwache Konvergenz anhand einer festen, *abzählbaren* Menge stetiger Funktionen charakterisieren lässt.

Satz 2.3. *Es sei* $\mathcal{X}$ *ein polnischer Raum. Dann gibt es eine abzählbare Familie* f_k *in* $C_b(\mathcal{X})$, *so dass für Maße* μ *und* μ_n *in* $\mathcal{M}_{\leqslant 1}(\mathcal{X})$ *gilt:*

Parthasarathy (1967, Thm. II.6.6)

$$\mu_n \xrightarrow{w} \mu \qquad \Leftrightarrow \qquad \langle \mu_n, f_k\rangle \to \langle \mu, f_k\rangle \quad (k \in \mathbb{N}).$$

Für das zentrale Kompaktheitsresultat benötigen wir den Begriff der *Straffheit*.

Definition 2.4. Eine Familie $\mathcal{F} \subset \mathcal{M}_{\leqslant 1}(\mathcal{X})$ heißt *straff*, falls es zu jedem $\epsilon > 0$ eine kompakte Menge $K \subset \mathcal{X}$ gibt mit

Elstrodt (2005, Def. VIII.4.18)

$$\sup\{\mu(K^c) : \mu \in \mathcal{F}\} < \epsilon.$$

Dabei bezeichnet A^c das Komplement der Menge A. Auf einem polnischen Raum $\mathcal{X}$ ist jedes einzelne Maß für sich allein straff; ist $\mathcal{X}$ kompakt, so ist natürlich jede Familie endlicher Maße straff.

Satz 2.5 (Prohorov 1956)**.** *Für polnischen Räume* $\mathcal{X}$ *gilt: Eine Familie* $\mathcal{F} \subset \mathcal{M}_{\leqslant 1}(\mathcal{X})$ *von Sub-W-Maßen ist genau dann schwach relativ folgenkompakt, wenn sie straff ist.*

Elstrodt (2005, Satz VIII.4.23)

Auf die Straffheit kann dann und nur dann verzichtet werden, wenn der polnische Raum $\mathcal{X}$ selbst (folgen-) kompakt ist.

Parthasarathy (1967, Thm. II.6.4)

Auf einem polnischen Raum kann die schwache Konvergenz von W-Maßen durch eine Metrik induziert werden; eine besonders handliche solche Metrik wird durch

Folgt aus Stroock (1993, Lemma 3.1.10) und Satz 2.2(iii).

$$d(\mu, \nu) = \sup_{f \in \mathrm{Lip}_1(\mathcal{X};[-1,1])} |\langle \mu, f\rangle - \langle \nu, f\rangle|$$

definiert, wobei $\mathrm{Lip}_1(\mathcal{X}; [-1, 1])$ die Menge jener Lipschitz-stetigen Funktionen $f : \mathcal{X} \to [-1, 1]$ bezeichnet, deren Lipschitz-Konstante jeweils durch 1 beschränkt ist. Tatsächlich ist $\mathcal{M}_1(\mathcal{X})$ dann selbst ein polnischer Raum; dies erlaubt es, die Konstruktion topologischer Maße zu „stapeln“ und etwa die schwache Konvergenz von W-Maßen auf W-Maßen etc. zu betrachten.

Stroock (1993, Thm. 3.1.11)

3. Schwache Konvergenz von Zufallsvariablen

Für eine (reell-wertige) Zufallsvariable X auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$ definieren wir die Verteilung als das W-Maß $\mu_X \in \mathcal{M}_1(\mathbb{R})$ durch

$$\mu_X(A) = \mathbb{P}(X \in A) \qquad (A \in \mathcal{B}(\mathbb{R})).$$

Beachte, dass $\mathbb{R}$ ein lokalkompakter polnischer Raum ist und unsere bisherigen Konzepte daher Anwendung finden können. Die zugehörige Verteilungsfunktion $F_X : \mathbb{R} \to [0, 1]$ ist

$$F_X(x) = \mathbb{P}(X \leqslant x) = \mu_X((-\infty, x]) \qquad (x \in \mathbb{R}).$$

Für eine Folge X_n von Zufallsvariablen definieren wir die *schwache Konvergenz* $X_n \xrightarrow{w} X$ (auch *Konvergenz der Verteilung nach* genannt) durch $\mu_{X_n} \xrightarrow{w} \mu_X$.

Satz 3.1 (Portmanteau-Theorem für Zufallsvariable). *Es seien X und X_n Zufallsvariable. Dann sind äquivalent:*

Direkte Umformulierung von Satz 2.2, siehe auch Klenke (2008, Thm. 13.23 und Cor. 13.24).

(i) $X_n \xrightarrow{w} X$;

(ii) $\mathbb{E}(f(X_n)) \to \mathbb{E}(f(X))$ *für alle beschränkten, stetigen f;*

(iii) $\mathbb{E}(f(X_n)) \to \mathbb{E}(f(X))$ *für alle beschränkten, Lipschitz-stetigen f;*

(iv) $\mathbb{E}(f(X_n)) \to \mathbb{E}(f(X))$ *für alle beschränkten, messbaren f mit $\mathbb{P}(X \in U_f) = 0$;*

(v) $\mathbb{P}(X_n \in A) \to \mathbb{P}(X \in A)$ *für alle $A \in \mathcal{B}(\mathbb{R})$, für welche $X \notin \partial A$ fast sicher;*

(vi) $F_{X_n}(x) \to F_X(x)$ *für alle Stetigkeitspunkte x von F_X.*

Ist F_X zusätzlich stetig, so ist zudem mit (i)–(vi) äquivalent:

van der Vaart (1998, Lemma 2.11)

(vii) $\|F_{X_n} - F_X\|_\infty \to 0$, *d.h. die Konvergenz in (vi) ist gleichmäßig.*

4. Stochastische und fast sichere Konvergenz von Zufallsvariablen

Neben dem Konzept der schwachen Konvergenz (das ja nur die Wahrscheinlichkeitsverteilungen der Zufallsvariablen benutzt) benutzen wir noch weitere Konvergenzkonzepte, welche die Zufallsvariablen als reelle Funktionen betrachten.

Definition 4.1. Es seien X und X_n Zufallsvariablen auf dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$.

Grimmett and Stirzaker (2001, Def. 7.2(1))

- X_n konvergiert *stochastisch* gegen X, in Zeichen $X_n \xrightarrow{\mathbb{P}} X$, falls für $\epsilon > 0$
$$\mathbb{P}(|X_n - X| \geqslant \epsilon) \to 0 \qquad (n \to \infty).$$
- X_n konvergiert *fast sicher* gegen X, in Zeichen $X_n \xrightarrow{\text{a.s.}} X$, falls
$$\mathbb{P}\left(\{\omega \in \Omega : X_n(\omega) \to X(\omega) \text{ für } n \to \infty\}\right) = 1.$$

Es gilt folgender Satz vom „abelschen“ Typ (d.h. ohne Zusatzbedingungen) über die wechselseitige Beziehung der Konvergenzkonzepte.

Satz 4.2. *Es gelten die Implikationen*

Grimmett and Stirzaker (2001, Thm. 7.2(3))

$$X_n \xrightarrow{a.s.} X \quad \Rightarrow \quad X_n \xrightarrow{\mathbb{P}} X \quad \Rightarrow \quad X_n \xrightarrow{w} X.$$

Die Umkehrungen sind nur unter Zusatzbedingungen richtig. Wir geben zwei einfache Beispiele für solche Sätze vom „tauberschen“ Typ. Das erste folgt fast unittelbar aus den Definitionen:

Satz 4.3. *Ist $c \in \mathbb{R}$ eine Konstante, so gilt die Implikation*

Grimmett and Stirzaker (2001, Thm. 7.2(4a))

$$X_n \xrightarrow{w} c \quad \Rightarrow \quad X_n \xrightarrow{\mathbb{P}} c.$$

Das zweite Beispiel ist ein einfacher Spezialfall des Lemmas von Borel–Cantelli.

Satz 4.4. *Gilt für jedes* $\epsilon > 0$

Grimmett and Stirzaker (2001, Thm. 7.2(4c))

$$\sum_{n=1}^{\infty} \mathbb{P}(|X_n - X| \geqslant \epsilon) < \infty,$$

(und damit natürlich unmittelbar $X_n \xrightarrow{\mathbb{P}} X$*), so ist* $X_n \xrightarrow{a.s.} X$.

Mit Hilfe der Tschebyscheff'schen Ungleichung

Billingsley (1995, (21.13))

$$\mathbb{P}(|X_n - \mathbb{E}X_n| \geqslant \epsilon) \leqslant \frac{\mathrm{Var}(X_n)}{\epsilon^2}$$

folgt aus diesem Theorem die sehr nützliche Implikation

$$\sum_{n=1}^{\infty} \mathrm{Var}(X_n) < \infty \quad \Rightarrow \quad X_n - \mathbb{E}X_n \xrightarrow{\text{a.s}} 0.$$

Die folgenden beiden Sätze ermöglichen es oft, in Beweisen die stochastische und die schwache Konvergenz durch die fast sichere Konvergenz zu ersetzen; sie führen zu einfachen und transparenten Argumentationen.

Satz 4.5. *Es seien* X *und* X_n *Zufallsvariable. Dann ist* $X_n \xrightarrow{\mathbb{P}} X$ *äquivalent dazu, dass jede Teilfolge* $X_{n'}$ *eine weitere Teilfolge* $X_{n''}$ *mit* $X_{n''} \xrightarrow{a.s.} X$ *besitzt.*

Billingsley (1995, Thm. 20.5)

Satz 4.6 (Darstellungssatz von Skorokhod). *Es seien* X *und* X_n *Zufallsvariable. Dann ist* $X_n \xrightarrow{w} X$ *äquivalent dazu, dass es Zufallsvariable* Y_n *und* Y *auf einem Wahrscheinlichkeitsraum* $(\Omega', \mathcal{A}', \mathbb{P}')$ *gibt mit*

Billingsley (1995, Thm. 25.6)

(i) $\mu_{Y_n} = \mu_{X_n}$ *und* $\mu_Y = \mu_X$,

(ii) $Y_n \xrightarrow{a.s.} Y$.

Wir stellen noch drei „Stetigkeitsresultate" für die Konvergenz von Zufallsvariablen zusammen.

Satz 4.7 („Continuous Mapping Theorem"). *Es seien* X *und* X_n *Zufallsvariable und* $\phi : \mathbb{R} \to \mathbb{R}$ *messbar. Gilt* $X \notin U_\phi$ *fast sicher, so ist*

van der Vaart (1998, Thm. 2.3)

$$X_n \to X \quad \Rightarrow \quad \phi(X_n) \to \phi(X)$$

sowohl für die schwache, als auch für die stochastische und die fast sichere Konvergenz von Zufallsvariablen.

Satz 4.8. *Es seien* X, X_n, Y *und* Y_n *Zufallsvariable. Dann gilt*

Grimmett and Stirzaker (2001, Thm. 7.3(9))

- $X_n \xrightarrow{a.s.} X$ *und* $Y_n \xrightarrow{a.s.} Y$ *impliziert* $X_n + Y_n \xrightarrow{a.s.} X + Y$;
- $X_n \xrightarrow{\mathbb{P}} X$ *und* $Y_n \xrightarrow{\mathbb{P}} Y$ *impliziert* $X_n + Y_n \xrightarrow{\mathbb{P}} X + Y$.

Für die schwache Konvergenz ist die Addition von Zufallsvariablen hingegen in der Regel nicht stetig, es gilt nur folgender schwächerer Satz.

Satz 4.9 (Lemma von Slutsky). *Es gelte* $X_n \xrightarrow{w} X$ *und* $Y_n \xrightarrow{w} c$, *wobei* $c \in \mathbb{R}$ *eine Konstante ist. Dann gilt*

van der Vaart (1998, Lemma 2.8)

- $X_n + Y_n \xrightarrow{w} X + c$;
- $Y_n \cdot X_n \xrightarrow{w} c \cdot X$.

5. Die Momentenmethode

Da für das kompakte Intervall $[a,b]$ die Polynome in $C[a,b]$ bzgl. der Supremumsnorm *dicht* liegen, kann die schwache Konvergenz in $\mathcal{M}_1([a,b])$ einfach durch Polynome getestet werden. Da nun aber

$$\langle \mu, x^k \rangle = m_k(\mu)$$

das k-te Moment des W-Maßes $\mu \in \mathcal{M}_1([a,b])$ ist, so gilt in $\mathcal{M}_1([a,b])$

$$\mu_n \xrightarrow{w} \mu \quad \Leftrightarrow \quad m_k(\mu_n) \to m_k(\mu) \quad (k \in \mathbb{N}).$$

Der folgende Satz zeigt, dass dieses Resultat unter einer zusätzlichen „tauberschen" Bedingung (gemeint ist die Eindeutigkeit in (ii)) auch für $\mathbb{R}$ statt $[a,b]$ gültig ist.

Satz 5.1. *Sei X_n eine Folge von Zufallsvariablen mit gleichmäßig beschränkten Momenten:*

$$\sup_n |\mathbb{E}(X_n^k)| \leqslant M_k < \infty \qquad (k \in \mathbb{N}).$$

Dann gilt:

(i) $X_n \xrightarrow{w} X$ *impliziert* $\mathbb{E}(X_n^k) \to \mathbb{E}(X^k)$ *für alle* $k \in \mathbb{N}$; Billingsley (1995, Cor. 25.12)

(ii) *wenn die Verteilung von X durch ihre Momente eindeutig festgelegt ist und* $\mathbb{E}(X_n^k) \to \mathbb{E}(X^k)$ *für alle* $k \in \mathbb{N}$, *so folgt* $X_n \xrightarrow{w} X$. Billingsley (1995, Thm. 30.2)

Wir benötigen also handliche Kriterien dafür, dass ein W-Maß $\mu \in \mathcal{M}_1(\mathbb{R})$ eindeutig durch seine Momente bestimmt ist.

Satz 5.2. *Das W-Maß $\mu \in \mathcal{M}_1(\mathbb{R})$ besitze die endlichen Momente $m_k = \int_{-\infty}^{\infty} x^k \, \mu(dx)$ zu jeder Ordnung $k \in \mathbb{N}$. Dann ist μ das einzige W-Maß mit diesen Momenten, falls eines der beiden folgenden Kriterien erfüllt ist:*

- *(Carleman'sches Kriterium)* Akhiezer (1965, p. 85)

$$\sum_{k=1}^{\infty} m_{2k}^{-1/2k} = \infty$$

oder

- *die Potenzreihe* Billingsley (1995, Thm. 30.1)

$$\sum_{k=1}^{\infty} m_k \frac{z^k}{k!}$$

besitzt positiven Konvergenzradius.

Das zweite Kriterium wird für die Verteilung einer Zufallsvariablen X besonders handlich, wenn die Momenten-erzeugende Funktion

$$M(z) = \mathbb{E}(e^{zX})$$

für alle reellen $|z| < r$ ($r > 0$) existiert. Dann besitzt nämlich X endliche Momente $\mathbb{E}(X^k)$ jeder Ordnung $k \in \mathbb{N}$ und es gilt Billingsley (1995, (21.22))

$$M(z) = \sum_{k=0}^{\infty} \mathbb{E}(X^k) \frac{z^k}{k!} \qquad (|z| < r).$$

Der Konvergenzradius beträgt also mindestens r, so dass die Verteilung von X nach dem zweiten Kriterium eindeutig aus den Momenten bestimmt ist. Die

Konvergenz von Momenten gegen diejenigen von X würde also die schwache Konvergenz gegen X sichern.

Satz 5.3. *Es seien X_n und X Zufallsvariablen, für welche die zugehörigen Momenten-erzeugenden Funktionen $M_n(z)$ und $M(z)$ für $|z| < r$ mit einem gewissen $r > 0$ existieren. Gilt dann $M_n(z) \to M(z)$ punktweise für $|z| < r$, so folgt $X_n \xrightarrow{w} X$.* Billingsley (1995, p. 390)

Beispiel 5.4. Für $N(m, \sigma^2)$-verteiltes X gilt

$$M(z) = \mathbb{E}(e^{zX}) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{zx} e^{-\frac{(x-m)^2}{2\sigma^2}}\, dx = e^{mz+\sigma^2 z^2/2},$$

eine ganze Funktion, die für alle z existiert. Die Verteilung einer normalverteilten Zufallsvariable ist also durch ihre Momente eindeutig bestimmt; für ein $N(0,1)$-verteiltes X lauten die Momente

$$\mathbb{E}(X^n) = \begin{cases} 0 & n \text{ ungerade;} \\ (2k-1)!! = 1 \cdot 3 \cdot 5 \cdots (2k-3) \cdot (2k-1) & n = 2k \text{ gerade.} \end{cases}$$

Die *Doppelfakultäten* $(2k-1)!!$ besitzen zahlreiche kombinatorische Deutungen, die zu ihrer Identifizierung herangezogen werden können. Sloane (2010, A001147)

Beispiel 5.5. Für eine Zufallsvariable X mit Halbkreisverteilung gilt

$$M(z) = \mathbb{E}(e^{zX}) = \frac{1}{2\pi} \int_{-2}^{2} e^{zx} \sqrt{4 - x^2}\, dx = z^{-1} I_1(2z),$$

ebenfalls eine ganze Funktion, die für alle z existiert. Die Verteilung einer halbkreisverteilten Zufallsvariable X ist also ebenfalls durch ihre Momente eindeutig bestimmt:

$$\mathbb{E}(X^n) = \begin{cases} 0 & n \text{ ungerade;} \\ C_k = \dfrac{\binom{2k}{k}}{k+1} & n = 2k \text{ gerade.} \end{cases}$$

Auch die *Catalan'schen Zahlen* C_k besitzen zahllose kombinatorische Deutungen, die zu ihrer Identifizierung herangezogen werden können. Sloane (2010, A000108) und Koshy (2009)

Literatur

Akhiezer, N. I.: 1965, *The classical moment problem and some related questions in analysis*, Hafner Publishing, New York.

Bauer, H.: 1978, *Wahrscheinlichkeitstheorie und Grundzüge der Masstheorie*, third edn, Walter de Gruyter, Berlin-New York.

Billingsley, P.: 1995, *Probability and measure*, third edn, John Wiley, New York.

Elstrodt, J.: 2005, *Maß- und Integrationstheorie*, fourth edn, Springer-Verlag, Berlin.

Grimmett, G. R. and Stirzaker, D. R.: 2001, *Probability and random processes*, third edn, Oxford University Press, New York.

Klenke, A.: 2008, *Probability theory*, Springer-Verlag, London.

Koshy, T.: 2009, *Catalan numbers with applications*, Oxford University Press, Oxford.

Parthasarathy, K. R.: 1967, *Probability measures on metric spaces*, Academic Press, New York.

Sloane, N. J. A.: 2010, The on-line encyclopedia of integer sequences, `www.research.att.com/~njas/sequences/`.

Stroock, D. W.: 1993, *Probability theory, an analytic view*, Cambridge University Press, Cambridge.

van der Vaart, A. W.: 1998, *Asymptotic statistics*, Cambridge University Press, Cambridge.