Kapitel 8

Schwache Konvergenz

Im Hinblick auf die funktionalen Grenzwertsätze wird in diesem Kapitel die Theorie der *schwachen Konvergenz* endlicher Maße auf separablen metrischen, vorallem polnischen Räumen entwickelt. In dieser Situation ist die schwache Topologie auf dem Raum der endlichen Maße *metrisierbar*; Begriffe wie Konvergenz oder Kompaktheit lassen sich also auf Eigenschaften von Folgen zurückführen. Das resultierende Konzept (Konvergenz entlang einer *verteilungsbestimmenden Klasse* + *Straffheit* $\Rightarrow$ schwache Konvergenz) ist in vielen Situation verhältnismäßig gut zu handhaben – die zentrale Kette an Argumenten geht auf Prohorov 1956 zurück.

Wir beginnen mit einer Wiederholung der

Definition 8.1. Sei $(\mathcal{E}, d)$ ein metrischer Raum und $\mathcal{C}_b(\mathcal{E})$ die Menge der stetigen beschränkten reellen Funktionen in $\mathcal{E}$. Die *schwache Topologie* im Raum der endlichen Maße in $\mathcal{B}(\mathcal{E})$ ist die gröbste Topologie, unter der die Abbildungen $\mu \mapsto \mu(f) := \int f(x) d\mu(x)$ für jedes $f \in \mathcal{C}_b(\mathcal{E})$ stetig sind. Entsprechend sagt man, eine *Folge* $(\mu_n)_{n\in\mathbb{N}}$ *konvergiere schwach* gegen ein Maß μ ($\mu_n \Rightarrow \mu$), falls

$$\int f d\mu_n \longrightarrow \int f d\mu \quad \forall\, f \in \mathcal{C}_b(\mathcal{E}).$$

Eine Folge $(X_n)_{n\in\mathbb{N}}$ von ZVAs mit Werten in $\mathcal{E}$ heißt schwach konvergent mit Grenzwert X ($X_n \to_{\mathcal{D}} X$), falls $\mathbb{P}^{X_n} \Rightarrow \mathbb{P}^X$.

Bemerkung 8.2. (i) Für $\mathcal{E} = \mathbb{R}$ ist die schwache Konvergenz äquivalent zur Konvergenz der Verteilungsfunktionen $F_n(x) \to F(x)$ in allen Stetigkeitspunkten x von F (Satz von Helly-Bray; W-Theorie I).

(ii) Wegen Satz 6.17 ist der Grenzwert eindeutig.

(iii) Zur schwachen Konvergenz $\mu_n \Rightarrow \mu$ ist äquivalent: $\int f d\mu_n \to \int f d\mu$ für alle f einer bezüglich der sup-Norm dichten Teilmenge von $\mathcal{C}_b(\mathcal{E})$, da das Maßintegral stabil gegenüber gleichmäßiger Konvergenz ist:

$$\Big| \int f \; d(\mu_n - \mu) \; - \; \int g \; d(\mu_n - \mu) \Big| \leq 2\|f - g\|_{\sup}.$$

Weitere äquivalente Charakterisierungen enthält

Satz 8.3. (Portemanteau-Theorem) *Seien μ_n, $n \in \mathbb{N}$, und μ W-Maße auf $\mathcal{B}(\mathcal{E})$. Folgende Bedingungen sind äquivalent:*

(i) $\mu_n \Rightarrow \mu$;

(ii) $\liminf_{n\to\infty} \mu_n(O) \geq \mu(O)$ *für alle offenen* $O \subset \mathcal{E}$;

(iii) $\limsup_{n\to\infty} \mu_n(A) \leq \mu(A)$ *für alle abgeschlossenen* $A \subset \mathcal{E}$;

(iv) $\mu_n(B) \to \mu(B)$ *für alle* $B \in \mathcal{B}(\mathcal{E})$ *mit* $\mu(\partial B) = 0$.

BEWEIS.

(i) $\Rightarrow$ (ii): Zu einer offenen Teilmenge O bilden wir die beschränkte stetige Funktion $f_{n,O}(x) := \min\big(n \cdot d(x, O^c),\ 1\big)$ (vgl. Kapitel 6, Beweis Satz 6.17); es folgt $f_{n,O} \nearrow 1_O$. Daher ist $\mu_n(O) \geq \int f_{m,O}\, d\mu_n$ ($m \in \mathbb{N}$) und folglich

$$\begin{aligned}\liminf_{n\to\infty} \mu_n(O) &\geq \lim_{n\to\infty} \int f_{m,O}\, d\mu_n \\ &= \int f_{m,O}\, d\mu.\end{aligned}$$

Dies gilt für alle $m \in \mathbb{N}$, also auch im Grenzwert $m \to \infty$, womit $\liminf_n \mu_n(O) \geq \mu(O)$.

(ii) $\Leftrightarrow$ (iii): Die Äquivalenz ist leicht nachzuweisen (man bilde Komplemente).

(ii), (iii) $\Rightarrow$ (iv): Mit dem Abschluss $\bar{B}$ und dem Inneren B^o von $B \in \mathcal{B}(\mathcal{E})$ folgt aus den Aussagen (ii) un (iii), dass

$$\left.\begin{array}{lllll}\limsup\limits_{n\to\infty} \mu_n(B) & \leq & \limsup\limits_{n\to\infty} \mu_n(\bar{B}) & \leq & \mu(\bar{B}) \\ \liminf\limits_{n\to\infty} \mu_n(B) & \geq & \liminf\limits_{n\to\infty} \mu_n(B^o) & \geq & \mu(B^o)\end{array}\right\} = \mu(B),$$

sofern $\mu(\partial B) = 0$.

(iv) $\Rightarrow$ (i): Sei $\mathcal{A}_0$ die Algebra aller $A \in \mathcal{B}(\mathcal{E})$, für welche $\mu(\partial A) = 0$. Sei E_0 der Raum aller $\mathcal{A}_0$-messbaren Treppenfunktionen. Für $g \in E_0$ gilt nach (iv) $\int g d\mu_n \to \int g d\mu$ ($n \to \infty$). Sei $f \in \mathcal{C}_b(\mathcal{E})$. Aufgrund von $\partial\{f > \lambda\} \subset \{f = \lambda\}$ und $\{f = \lambda\} \cap \{f = \lambda'\} = \emptyset$ für $\lambda \neq \lambda'$ ist $\mu(\partial\{f > \lambda\}) = 0$ bis auf höchstens abzählbar viele Werte von λ, bzw. $1_{\{f>\lambda\}} \in E_0$ bis auf höchstens abzählbar viele Werte von λ. Damit existieren zu jedem $\varepsilon > 0$ Treppenfunktionen $g^+ \leq f_+ \leq h^+$ und $g^- \leq f_- \leq h^-$ aus E_0, so dass $\int h^+ d\mu - \int g^+ d\mu < \varepsilon$ und $\int h^- d\mu - \int g^- d\mu < \varepsilon$. Es folgt

$$\begin{array}{ccccc}\int g^+ d\mu_n & \leq & \int f_+ d\mu_n & \leq & \int h^+ d\mu_n \\ \downarrow & & & & \downarrow \\ \int g^+ d\mu & \leq & \int f_+ d\mu & \leq & \int h^+ d\mu,\end{array}$$

und $\int h^+ d\mu_n - \int g^+ d\mu_n < \varepsilon$ für hinreichend großes n (analog für den Negativteil f_-), woraus die Behauptung folgt. $\square$

Bemerkung 8.4. Bedingung (ii) zeigt, dass der Begriff der schwachen Konvergenz *nicht* von der Einbettung abhängt, denn die offenen Mengen in $(B, d_{|B})$ sind von der Form $O \cap B$ mit O offen in $(\mathcal{E}, d)$. Mit anderen Worten, sind μ_n, $n \in \mathbb{N}$, μ Borel-Maße mit $\mu_n(B) = \mu(B) = 1$ für eine Borel-Teilmenge $B \subset \mathcal{E}$, so sind äquivalent:

(i) $\int f d\mu_n \to \int f d\mu$ für alle $f \in \mathcal{C}_b(\mathcal{E})$ und

(ii) $\int f d\mu_n \to \int f d\mu$ für alle $f \in \mathcal{C}_b(B)$.

Man beachte, dass im allg. eine beschränkte stetige Funktion auf B *nicht* die Restriktion einer stetigen beschränkten Funktion auf $\mathcal{E}$ ist.

Bevor wir die schwache Konvergenz genauer untersuchen, noch drei Analogien zu reellwertigen Zufallsvariablen (der Beweis sei dem Leser als Übungsaufgabe überlassen).

Bemerkung 8.5. Seien X_n, Y_n, $n \in \mathbb{N}$, X Zufallsvariablen mit Werten in einem metrischen Raum $(\mathcal{E}, d)$.

(i) Gilt $X_n \to_{\mathcal{D}} X$ und ist $h : (\mathcal{E}, d) \to \mathbb{R}$ stetig, so folgt $h(X_n) \to_{\mathcal{D}} h(X)$ (Continuous-mapping-Theorem);

(ii) Eine Folge (X_n) von Zufallsvariablen mit Werten in einem separablen metrischen Raum konvergiert stochastisch gegen X ($X_n \to_P X$), falls $d(X_n, X) \to_P 0$. Vorsicht ist geboten, dass $d : \mathcal{E} \times \mathcal{E} \to \mathbb{R}$ überhaupt $\mathcal{B}(\mathcal{E}) \otimes \mathcal{B}(\mathcal{E}) - \mathcal{B}(\mathbb{R})$-messbar ist! Ist $(\mathcal{E}, d)$ separabel, so ist d messbar, denn dann hat man $\mathcal{B}(\mathcal{E} \times \mathcal{E}) = \mathcal{B}(\mathcal{E}) \otimes \mathcal{B}(\mathcal{E})$ (vgl. Kapitel 6, Bem. 6.5), und d ist als stetige Funktion natürlich $\mathcal{B}(\mathcal{E} \times \mathcal{E}) - \mathcal{B}(\mathbb{R})$-messbar;

(iii) Gilt $X_n \to_{\mathcal{D}} X$ und $d(X_n, Y_n) \to_P 0$, so folgt $Y_n \to_{\mathcal{D}} X$ (Satz von Slutzky).

Metrisierbarkeit der schwachen Konvergenz

Ziel ist es zu zeigen, dass die schwache Topologie metrisiert werden kann. Beweistechnisch ist der Vorzug, dass Begriffe wie Konvergenz oder (relative) Kompaktheit auf Eigenschaften von Folgen zurückgeführt werden können. Für Anwendungen in der mathematischen Statistik wie beispielsweise in der Dimensionsasymptotik als adäquater Approximation im Zusammenhang mit Microarray-Daten oder dem sog. Bootstrap (Resampling-Verfahren) ermöglicht die Metrisierbarkeit die einfache Charakterisierung einer *schwachen Approximation* zweier Folgen von Zufallsvariablen, *ohne* dass ein Grenzwert existieren muss.

Bezeichnungen. Für eine Funktion $f : \mathcal{E} \to \mathbb{R}$ seien

$$\|f\|_L := \sup_{x \neq y} \frac{f(x) - f(y)}{d(x,y)} \quad \text{(Lipschitz-Seminorm)}$$

und

$$\|f\|_{BL} := \|f\|_L + \|f\|_{\sup} \quad \text{(bounded Lipschitz-Norm)}$$

$\|f\|_L$ ist die Lipschitz-Konstante von f; ist $\|f\|_L$ endlich, so heißt f Lipschitzfunktion zur Lipschitz-Konstante $\|f\|_L$. $\|\cdot\|_{BL}$ ist eine Norm auf $BL = \{f : \|f\|_{BL} < \infty\}$. BL ist vollständig, eine Algebra mit $\|fg\|_{BL} \leq \|f\|_{BL} \cdot \|g\|_{BL}$ und ein Vektorverband. Die Normungleichung folgt leicht aus

$$\begin{aligned}|f(x)g(x) - f(y)g(y)| &\leq \big|\, f(x)\big(g(x) - g(y)\big) + g(y)\big(f(x) - f(y)\big)\,\big| \\ &\leq \|f\|_{\sup}\|g\|_L d(x,y) + \|g\|_{\sup}\|f\|_L d(x,y).\end{aligned}$$

Proposition 8.6. (duale BL-Metrik) *Auf dem Raum der endlichen Maße über $\mathcal{B}(\mathcal{E})$ definiert $d_{BL}(.,.)$, punktweise gegeben durch*

$$d_{BL}(\mu_1, \mu_2) := \|\mu_1 - \mu_2\|^*_{BL} = \sup\left\{ \int f d(\mu_1 - \mu_2) : \ \|f\|_{BL} \leq 1 \right\},$$

eine Metrik.

BEWEIS. Einzig zu zeigen ist: $\|\mu_1 - \mu_2\|^*_{BL} = 0 \Rightarrow \mu_1 = \mu_2$. Sei $A \subset \mathcal{E}$ abgeschlossen. Dann gilt $f_{n,A} \searrow 1_A$ mit $f_{n,A}(x) := \big(1 - n \cdot d(x, A)\big)^+$. Wegen $\|f_{n,A}\|_{BL} < \infty$ gilt $\int f_{n,A} d\mu_1 = \int f_{n,A} d\mu_2$ und folglich $\mu_1(A) = \mu_2(A)$, also $\mu_1 = \mu_2$. □

Satz 8.7. *Sei $(\mathcal{E}, d)$ ein separabler metrischer Raum. Für endliche Maße μ_n, $n \in \mathbb{N}$, und μ sind äquivalent:*

(i) $\mu_n \Rightarrow \mu$;

(ii) $\int f d\mu_n \to \int f d\mu$ *für alle* $f \in BL$;

(iii) $d_{BL}(\mu_n, \mu) \to 0$.

Bemerkung 8.8. Es ist Aussage des Satzes, dass die duale bounded-Lipschitz-Metrik die schwache Konvergenz metrisiert. Entsprechend kann man sagen, zwei Folgen (X_n) und (Y_n) von Zufallsvariablen approximieren sich schwach, wenn gilt $d_{BL}(\mathbb{P}^{X_n}, \mathbb{P}^{Y_n}) \to 0$ $(n \to \infty)$.

BEWEIS. (i) $\Rightarrow$ (ii): unmittelbar.

(ii) $\Rightarrow$ (iii): Setzt man voraus, dass $(\mathcal{E}, d)$ polnisch ist, so existiert zu $\varepsilon > 0$ eine kompakte Menge mit $\mu(K) > 1 - \varepsilon$ – eine Konsequenz aus Ulams Theorem (Satz 6.18). Ist nur die schwächere Bedingung des Satzes gegeben, so vervollständigt man $(\mathcal{E}, d)$ zu einem polnischen Raum, wobei die Lipschitzfunktionen als gleichmäßig stetige Funktionen stetig auf die Vervollständigung fortgesetzt werden. Man betrachtet μ_n, μ als Funktionale auf den so fortgesetzten Funktionen $f \in BL$.

Sei $B := \{f : \|f\|_{BL} \leq 1\}$. Diese Menge von Funktionen, eingeschränkt auf K, ist abgeschlossen bzgl. $\|.\|_{K,\sup}$ und folglich kompakt nach dem Satz von Arzelà-Ascoli (siehe Kapitel 9). Also existieren $f_1, ..., f_k \in B$ so dass für alle $f \in B$ gilt: $\|f - f_j\|_{K,\sup} = \sup_{x\in K} |f(x) - f_j(x)| < \varepsilon$ für ein $j = 1, ..., k$. Automatisch ist dann $\sup_{x\in K_\varepsilon} |f(x) - f_j(x)| < 3\varepsilon$ für ein $j = 1, ..., k$, mit $K_\varepsilon := \{x \in \mathcal{E} : \ d(x, K) < \varepsilon\}$.

Weiter gilt $\mu_n(K_\varepsilon) > 1 - \varepsilon$ für hinreichend großes $n \in \mathbb{N}$, wie sich mit $\mu(K) > 1 - \varepsilon$ aus der Betrachtung der Funktionen $1_K \leq g \leq 1_{K_\varepsilon}$, $g \in BL$, gegeben durch $g(x) := \big(1 - d(x, K)/\varepsilon\big)^+$, ergibt. Damit ist mit obigen j

$$\Big| \int (f - f_j) d(\mu_n - \mu) \Big| \ \leq \ \int_{K_\varepsilon} |f - f_j| d|\mu_n - \mu| + \int_{K_\varepsilon^c} |f - f_j| d(\mu_n + \mu) \ \leq \ 3\varepsilon \cdot 2 + 2 \cdot 2\varepsilon = 10\varepsilon.$$

Es folgt $\big| \int f d(\mu_n - \mu) \big| \ \leq \ \big| f_j d(\mu_n - \mu) \big| + 10\varepsilon$, wobei der erste Term der rechten Seite nach Voraussetzung (ii) gegen Null strebt.

(iii) $\Rightarrow$ (i): Aus (iii) folgt sofort (ii). (i) ergibt sich daraus folgendermaßen: Sei $A \subset \mathcal{E}$ abgeschlossen und $g_{\varepsilon,A}$ gegeben durch $g_{\varepsilon,A}(x) := \big(1 - d(x, A)/\varepsilon\big)^+$. $g_{\varepsilon,A}$ ist eine beschränkte Lipschitzfunktion mit $1_A \leq g_{\varepsilon,A}$ und $\int g_{\varepsilon,A} d\mu \searrow \mu(A)$ für $\varepsilon \searrow 0$. Ferner gilt $\int g_{\varepsilon,A} d\mu_n \to \int g_{\varepsilon,A} d\mu$ $(n \to \infty)$ nach (ii) und desweiteren $\limsup_n \mu_n(A) \leq \limsup_n \int g_{\varepsilon,A} d\mu_n = \int g_{\varepsilon,A} d\mu$, womit (i) bewiesen ist. □

(Relativ) schwache Kompaktheit

Beispiel 8.9. (Maße in $\mathcal{B}(\mathcal{C}[0,1])$) Da die endlichdimensionalen Projektionen wegen $\mathcal{B}(\mathcal{C}[0,1]) = \mathcal{B}^{[0,1]} \cap \mathcal{C}[0,1]$ ein $\cap$-stabiles Erzeugendensystem definieren (Proposition 6.19 und nachfolgende Bemerkung), ist ein W-Maß auf $\mathcal{B}(\mathcal{C}[0,1])$ eindeutig durch seine endlichdimensionalen Verteilungen (“fidis” – für finite dimensional distributions) festgelegt. *Dennoch folgt aus der schwachen Konvergenz der fidis nicht die schwache Konvergenz in* $\mathcal{B}(\mathcal{C}[0,1])$, wie folgendes Beispiel zeigt: Sei $\mathbb{P}_n$ das Dirac-Maß an der Funktion

$$x_n(t) := \begin{cases} nt & \text{für } t \in [0, 1/n] \\ 2 - nt & \text{für } t \in [1/n, 2/n] \\ 0 & \text{für } t \in [2/n, 1] \end{cases}$$

und $\mathbb{P}$ das Dirac-Maß in $x(.) = 0$. Wegen $\int f d\mathbb{P}_n = f(x_n(.))$ muss für die schwache Konvergenz die Konvergenz von $f \circ x_n$ gegen $f \circ x = f(0)$ für alle $f \in \mathcal{C}_b(\mathcal{C}[0,1])$ nachgewiesen werden. Dieses gilt aber beispielsweise nicht für $f(x(.)) := \min\big(\|x\|_{\sup}, 1\big)$. Andererseits konvergiert $x_n(t)$ f.s. für alle t, d.h. die fidis konvergieren.

Definition 8.10. Sei $(\mathcal{E}, d)$ ein metrischer Raum. Sei $\mathcal{P}$ die Menge aller W-Maße auf $\mathcal{B}(\mathcal{E})$. Eine Teilmenge $\mathcal{Q} \subset \mathcal{P}$ heißt

(i) *relativ schwach folgenkompakt*, wenn jede Folge (μ_n) in $\mathcal{Q}$ eine schwach konvergente Teilfolge mit Grenzwert in $\mathcal{P}$ besitzt;

(ii) *schwach folgenkompakt*, wenn jede Folge (μ_n) in $\mathcal{Q}$ eine schwach konvergente Teilfolge mit Grenzwert in $\mathcal{Q}$ besitzt.

Bemerkung 8.11. Wegen der Metrisierbarkeit der schwachen Topologie im Raum der endlichen Maße über der Borel-σ-Algebra eines *separablen* metrischen Raumes sind in diesem Falle die Begriffe *(relativ) schwach kompakt* und *(relativ) schwach folgenkompakt* äquivalent. (Eine Menge ist relativ kompakt, wenn ihr Abschluss kompakt ist.)

Satz 8.12. *Ist $\mathcal{E}$ kompakt, dann ist $\mathcal{P}$ schwach kompakt.*

BEWEIS. Sei $\mu_n \in \mathcal{P}$. $\mathcal{C}(\mathcal{E})$ ist separabel. Sei $g_1, g_2, ...$ eine bzgl. der sup-Norm dichte Teilmenge in $\mathcal{C}(\mathcal{E})$. Dann wählen wir nach dem Diagonalverfahren eine Teilfolge $(n(m))$, so dass $\int g_k d\mu_{n(m)}$ konvergiert ($k = 1, 2, ...; m \to \infty$). Aufgrund von

$$\left| \int g_k d\mu_{n(m)} - \int g d\mu_{n(m)} \right| \leq \|g - g_k\|_{\sup}$$

konvergiert dann auch $\int g d\mu_{n(m)}$ für alle $g \in \mathcal{C}(\mathcal{E})$. Bezeichne $l(g)$ den Limes: $\int g d\mu_{n(m)} \to l(g)$, so ist l eine positive Linearform auf $\mathcal{C}(\mathcal{E})$ mit $l(1) = 1$ – also gibt es ein (eindeutiges) Maß $\mu \in \mathcal{P}$ mit $l(g) = \int g d\mu$ (Rieszscher Darstellungssatz). Das heißt $\mu_{n(m)} \Rightarrow \mu$. □

Definition 8.13. Sei $(\mathcal{E}, d)$ ein metrischer Raum und $\mathcal{P}$ die Menge aller Borel-W-Maße auf $\mathcal{E}$. Eine Menge $\mathcal{Q} \subset \mathcal{P}$ heißt *straff* (engl. *tight*), wenn es für alle $\varepsilon > 0$ eine kompakte Menge K gibt mit $\mu(K) \geq 1 - \varepsilon$ für alle $\mu \in \mathcal{Q}$.

Satz 8.14. (Prohorov, 1956) *(i) Sei $(\mathcal{E}, d)$ ein metrischer Raum und $\mathcal{Q}$ eine Menge von Borel-W-Maßen. Dann gilt:*

$$\mathcal{Q} \text{ straff} \Rightarrow \mathcal{Q} \text{ relativ folgenkompakt.}$$

(ii) Ist $(\mathcal{E}, d)$ polnisch, gilt auch die Umkehrung:

$$\mathcal{Q} \text{ relativ kompakt} \Rightarrow \mathcal{Q} \text{ straff.}$$

Wir zeigen die erste Aussage nur im Falle eines polnischen Raumes $(\mathcal{E}, d)$. Ein für nicht-separable metrische Räume deutlich aufwendigerer Beweis von Aussage (i) findet sich in Billingsley 1999 (Convergence of Probability Measures).

BEWEIS. Insbesondere bei Teil (ii) nutzen wir entscheidend aus, dass $\mathcal{E}$ als Borelmenge, genauer als ein G_δ, eines kompakten metrischen Raumes $(\mathcal{E}^*, d^*)$ betrachtet werden kann (siehe Kapitel 6).

(i) Sei $K \subset \mathcal{E}$ kompakt mit $\mu(K) > 1-\varepsilon$ $(\mu \in \mathcal{Q})$. Sei (μ_n) eine Folge in $\mathcal{Q}$. Nach Satz 8.12 existieren eine Teilfolge $(\mu_{n(k)})$ und ein Maß μ auf $\mathcal{E}^*$ mit $\mu_{n(k)} \Rightarrow \mu$. Es bleibt zu zeigen, dass $\mu(\mathcal{E}) = 1$. Dies folgt aber aus $1-\varepsilon < \limsup_k \mu_{n(k)}(K) \leq \mu(K) \leq \mu(\mathcal{E})$, da K als kompakte Menge in $\mathcal{E}^*$ abgeschlossen ist. Folglich ist $\mu(\mathcal{E}) = 1$; μ kann also als W-Maß auf $(\mathcal{E}, d)$ betrachtet werden.

(ii) $\mathcal{E}$ ist ein G_δ in $\mathcal{E}^*$, d.h. von der Form $\mathcal{E} = \bigcap_{i\in\mathbb{N}} G_i$ mit offenen $G_i \subset \mathcal{E}^*$. Es ist $\mu(G_i) = 1$ für alle i $(\mu \in \mathcal{Q})$. Es genügt daher, kompakte Mengen $K_i \subset G_i$ mit $\mu(K_i) > 1-\varepsilon/2^i$ $(\mu \in \mathcal{Q})$ zu finden, weil dann für die kompakte Menge $K := \bigcap_{i\in\mathbb{N}} K_i \subset \bigcap_{i\in\mathbb{N}} G_i = \mathcal{E}$ folgt: $\mu(K) > 1-\varepsilon$ $(\mu \in \mathcal{Q})$. Somit reicht es, die Behauptung in dem Fall zu beweisen, dass $\mathcal{E} \subset \mathcal{E}^*$ *offen* ist. In diesem Fall kann man $\mathcal{E}$ schreiben als

$$\mathcal{E} = \bigcup_{i\in\mathbb{N}} B_i, \quad B_i \text{ offen}, \quad B_i \subset B_{i+1}, \quad \bar{B}_i \subset \mathcal{E} \;\text{ kompakt}$$

(etwa mit $B_i = \{y : d^*(y, \mathcal{E}^c) > 1/i\}$). Wir zeigen: Zu jedem $\varepsilon > 0$ existiert ein Index i, so dass $\mu(B_i) > 1-\varepsilon$ $(\mu \in \mathcal{Q})$. Das genügt, da dann $\mu(\bar{B}_i) > 1-\varepsilon$ für das kompakte $\bar{B}_i$. Wäre dies nicht so, gäbe es zu allen $n \in \mathbb{N}$ ein $\mu_n \in \mathcal{Q}$ mit $\mu_n(B_n) \leq 1-\varepsilon$. Gemäß Voraussetzung würde dann eine Teilfolge $\mu_{n(k)}$ existieren mit schwachem Grenzwert μ_∞ auf $\mathcal{E}$ $(\mu_\infty(\mathcal{E}) = 1)$, und es folgte $\mu_\infty(B_i) \leq \liminf_k \mu_{n(k)}(B_i) \leq \liminf_k \mu_{n(k)}(B_{n(k)}) \leq 1-\varepsilon$ für alle i. Damit wäre aber nach dem Satz von der monotonen Konvergenz $\mu_\infty(\mathcal{E}) \leq 1-\varepsilon$ im Widerspruch zur Voraussetzung. □

Die Räume $(\mathcal{C}[0,1], \sup)$ und $(\mathcal{D}[0,1], d_S)$

Der Satz von Prohorov führt nun zu einem Konzept für den Nachweis schwacher Konvergenz von W-Maßen auf polnischen Räumen. Am wichtigsten ist der Bildraum der Brownschen Bewegung $\mathcal{C}[0,1]$, versehen mit der Topologie der gleichmäßigen Konvergenz.

Satz 8.15. (Schwache Konvergenz in $\mathcal{C}[0,1]$) *Seien X_n, $n \in \mathbb{N}$, und X Zufallsvariablen mit Werten in $(\mathcal{C}[0,1], \sup)$. Dann gilt $X_n \to_{\mathcal{D}} X$ genau dann, wenn gilt $\pi \circ X_n \to_{\mathcal{D}} \pi \circ X$ für alle endlichdimensionalen Projektionen π und die Folge $(\mathbb{P}^{X_n})$ straff ist.*

BEWEIS. Da die Projektionen stetig sind (Kapitel 6, Satz 6.19), folgt aus $X_n \to_{\mathcal{D}} X$ die Konvergenz der fidis sowie aus Satz 8.14 die Straffheit. Gelten umgekehrt die schwache Konvergenz der fidis und die Straffheit, so impliziert Satz 8.14 (i) die relative Kompaktheit der Folge $(\mathbb{P}_n)$, $\mathbb{P}_n := \mathbb{P}^{X_n}$. Sei $\mathbb{P}_{n(m)}$ eine Teilfolge. Dann existiert eine Teilteilfolge $(\mathbb{P}_{n(m_j)})_{j\in\mathbb{N}}$ und ein W-Maß $\mathbb{P}$ mit $\mathbb{P}_{n(m_j)} \Rightarrow \mathbb{P}$. Wegen der Stetigkeit der endlichdimensionalenProjektionen folgt die schwache Konvergenz der endlichdimensionalen Verteilungen von

$\mathbb{P}_{n(m_j)}$ gegen die von $\mathbb{P}$ sowie von $\mathbb{P}^X$ nach Voraussetzung. Da die endlichdimensionalen Verteilungen nach Satz 6.19 eine Verteilung auf $(\mathcal{C}[0,1], \sup)$ eindeutig festlegen, folgt $\mathbb{P} = \mathbb{P}^X$, d.h. $\mathbb{P}_{n(m_j)} \Rightarrow \mathbb{P}^X$. Alle konvergenten Teilteilfolgen konvergieren also gegen denselben Limes $\mathbb{P}^X$ und damit auch die ganze Folge. □

Mithilfe einer Charakterisierung der relativ kompakten Mengen werden wir im folgenden Kapitel handhabbare Kriterien für die Straffheit von W-Maßen auf $(\mathcal{C}[0,1], \sup)$ herleiten.

Bemerkung 8.16. Die Aussage des vorangehenden Satzes mitsamt Beweis bleiben gültig auch für andere metrische Räume $(\mathcal{E}, d)$, solange die endlichdimensionalen Projektionen stetig sind und die Verteilung eindeutig festlegen (man spricht deswegen von einer *verteilungsbestimmenden Klasse*). Die Situation unterscheidet sich hinsichtlich der Charakterisierung der Straffheit.

Bemerkung 8.17. (Schwache Konvergenz in $\mathcal{D}[0,1]$) Bildraum des Poissonprozesses oder der empirischen Verteilungsfunktion eingeschränkt auf $[0,1]$ ist der Raum der càdlàg-Funktionen auf $[0,1]$, den man gewöhnlich mit $\mathcal{D}[0,1]$ bezeichnet. Dieser Funktionenraum, versehen mit der sup-Norm ist *nicht* separabel, denn zu den Indikatorfunktionen $1_{[t,1]}$ gibt es offenbar keine abzählbar dichte Teilmenge. Sei

$$d_S(x,y) := \inf_{\lambda \in \Lambda} \max \Big\{ \sup_t |t - \lambda(t)|, \ \sup_t |x(t) - y(\lambda(t))| \Big\} \quad \text{mit}$$

$$\Lambda := \big\{ \lambda : [0,1] \to [0,1] |\ \lambda \text{ ist bijektiv und streng monoton wachsend} \big\}$$

(insbesondere ist also $\lambda(0) = 0$ und $\lambda(1) = 1$ für $\lambda \in \Lambda$). d_S definiert auf $\mathcal{D}[0,1]$ eine Metrik (Übungsaufgabe), die die *Skorohod-Topologie* induziert. Man kann zeigen, dass $(\mathcal{D}[0,1], d_S)$ separabel, *nicht* jedoch vollständig ist. Es existiert aber eine zu d_S äquivalente vollständige Metrik d_S^*, d.h. $\mathcal{D}[0,1]$, versehen mit der Skorohod-Topologie, ist polnisch. Schwache Konvergenz ist also äquivalent zu Straffheit + Konvergenz aller Auswertungen $f \in \mathcal{F}$ entlang einer verteilungsbestimmenden Klasse $\mathcal{F} \subset \mathcal{C}_b$. Analog zu $\mathcal{B}(\mathcal{C}[0,1])$ (vgl. Proposition 6.19) gilt

$$\mathcal{B}(\mathcal{D}[0,1]) = \mathcal{B}^{[0,1]} \cap \mathcal{D}[0,1]$$

(insbesondere haben wir $\mathcal{B}(\mathcal{C}[0,1]) = \mathcal{B}(\mathcal{D}[0,1]) \cap \mathcal{C}[0,1]$), womit die endlichdimensionalen Verteilungen ein Maß auf $\mathcal{B}(\mathcal{D}[0,1])$ eindeutig festlegen. Allerdings sind (bis auf π_0 und π_1) die endlichdimensionalen Projektionen im allgemeinen *nicht* stetig, was die Charakterisierung der schwachen Konvergenz etwas verkompliziert. $\mathbb{P}_n \Rightarrow \mathbb{P}$ impliziert also *nicht* immer $\mathbb{P}_n^{\pi_{t_1,\ldots,t_k}} \Rightarrow \mathbb{P}^{\pi_{t_1,\ldots,t_k}}$. Für weitere Details sei der Leser verwiesen auf Billingsley (1999), Abschnitte 12 und 13.