

2. Konvergenz von Wahrscheinlichkeitsmassen

In diesem Kapitel bezeichnet (S, d) einen metrischen Raum und $\mathcal{P}(S)$ bezeichnet die Menge aller Wahrscheinlichkeitsmasse auf S bezüglich der Borel σ -Algebra $\mathfrak{B}(S)$.

2.1. Die Prohorov-Metrik

Definition 2.1. Die **Prohorov-Metrik** ρ auf $\mathcal{P}(S)$ ist definiert als

$$\rho(\mathbb{P}, Q) = \inf\{\varepsilon > 0 : \mathbb{P}[F] \leq Q[F^\varepsilon] + \varepsilon \text{ für alle } F \in \mathcal{C}\},$$

wobei $\mathcal{C}$ die Klasse aller abgeschlossenen Mengen von S ist, und

$$F^\varepsilon = \{x \in S : \inf_{y \in F} d(x, y) < \varepsilon\}$$

ist die ε -Umgebung von F . Wir sagen eine Folge $\{\mathbb{P}_n\}$ konvergiert nach $\mathbb{P}$, falls $\rho(\mathbb{P}_n, \mathbb{P})$ nach 0 konvergiert.

Wir wollen zuerst zeigen, dass ρ wirklich eine Metrik ist. Dazu brauchen wir den folgenden

Hilfssatz 2.2. Seien $\mathbb{P}, Q \in \mathcal{P}(S)$ und $\alpha, \beta > 0$. Falls

$$\mathbb{P}[F] \leq Q[F^\alpha] + \beta$$

für alle $F \in \mathcal{C}$ gilt, dann gilt auch

$$Q[F] \leq \mathbb{P}[F^\alpha] + \beta$$

für alle $F \in \mathcal{C}$.

Beweis. Sei $F_1 \in \mathcal{C}$ und $F_2 = S \setminus F_1^\alpha$. Dann ist $F_2 \in \mathcal{C}$ und $F_1 \subset S \setminus F_2^\alpha$. Also gilt

$$\mathbb{P}[F_1^\alpha] = 1 - \mathbb{P}[F_2] \geq 1 - Q[F_2^\alpha] - \beta \geq Q[F_1] - \beta.$$

Dies zeigt die Aussage. □

Hilfssatz 2.3. *Die Prohorov-Metrik ist eine Metrik auf $\mathcal{P}(S)$.*

Beweis. Aus Hilfssatz 2.2 folgt sofort, dass $\rho(\mathbb{P}, Q) = \rho(Q, \mathbb{P})$. Ist $\rho(\mathbb{P}, Q) = 0$, so folgt aus der monotonen Konvergenz, dass $\mathbb{P}[F] \leq Q(F)$ für alle $F \in \mathcal{C}$. Da die abgeschlossenen Mengen alle Mengen in $\mathcal{B}(S)$ erzeugen, muss also $\mathbb{P}[F] \leq Q(F)$ für alle $F \in \mathcal{B}(S)$ gelten. Wegen der Symmetrie folgt dann $\mathbb{P}[F] = Q(F)$ für alle $F \in \mathcal{B}(S)$, also $\mathbb{P} = Q$. Seien nun $\mathbb{P}, Q, R \in \mathcal{P}(S)$. Nehmen wir an, dass $\rho(\mathbb{P}, Q) < \varepsilon$ und $\rho(Q, R) < \delta$. Dann ist für $F \in \mathcal{C}$

$$\mathbb{P}[F] \leq Q[F^\varepsilon] + \varepsilon \leq Q[\overline{F^\varepsilon}] + \varepsilon \leq R[\overline{F^{\varepsilon+\delta}}] + \delta + \varepsilon \leq R[F^{\varepsilon+\delta}] + \delta + \varepsilon .$$

Somit ist $\rho(\mathbb{P}, R) \leq \delta + \varepsilon$. Dies beweist die Dreiecksungleichung. $\square$

Hilfssatz 2.4. *Sei μ ein endliches Borelmaß auf S , $n \in \mathbb{N} \setminus \{0\}$ und sei $p_i \geq 0$ und $A_i \in \mathcal{B}(S)$ für $i = 1, \dots, n$. Nehmen wir an, dass*

$$\sum_{i \in I} p_i \leq \mu \left[\bigcup_{i \in I} A_i \right] \quad (2.1)$$

für alle $I \subset \{1, \dots, n\}$. Dann gibt es Borelmasse $\lambda_1, \dots, \lambda_n$ auf S , so dass $\lambda_i(A_i) = \lambda_i(S) = p_i$ für alle i und $\sum_{i=1}^n \lambda_i(A) \leq \mu(A)$ für alle $A \in \mathcal{B}(S)$.

Beweis. Wir können annehmen, dass $p_i > 0$ für alle i . Wir beweisen die Aussage mittels vollständiger Induktion.

Für $n = 1$ definieren wir $\lambda_1(A) = p_1 \mu(A \cap A_1) / \mu(A_1)$. Dann ist $\lambda_1(A_1) = \lambda_1(S) = p_1$, und, da $p_1 \leq \mu(A_1)$, $\lambda_1(A) = \lambda_1(A \cap A_1) \leq \mu(A \cap A_1) \leq \mu(A)$.

Nehmen wir nun an, die Aussage gilt für $n = 1, \dots, m-1$. Definieren wir η durch $\eta(A) = \mu(A \cap A_m) / \mu(A_m)$. Definieren wir

$$\varepsilon_0 = \sup \left\{ \varepsilon : \sum_{i \in I} p_i \leq (\mu - \varepsilon \eta) \left(\bigcup_{i \in I} A_i \right) \text{ für alle } I \subset \{1, \dots, m-1\} \right\} .$$

Nehmen wir zuerst an, dass $\varepsilon_0 \geq p_m$. Setzen wir $\lambda_m = p_m \eta$ und $\mu' = \mu - \lambda_m$. Da $p_m \leq \mu(A_m)$ folgt, dass μ' ein Maß auf $\mathcal{B}(S)$ ist. Nach der Induktionsvoraussetzung gibt es Maße $\lambda_1, \dots, \lambda_{m-1}$, so dass $\lambda_i(A_i) = \lambda_i(S) = p_i$ und $\sum_{i=1}^{m-1} \lambda_i(A) \leq \mu'(A)$ für alle $A \in \mathcal{B}(S)$. Weiter gilt $\lambda_m(A_m) = \lambda_m(S) = p_m$. Durch die Konstruktion hat $\lambda_1, \dots, \lambda_m$ die benötigten Eigenschaften.

Nehmen wir nun $\varepsilon_0 < p_m$ an. Wir setzen dann $\mu' = \mu - \varepsilon_0 \eta$. Es gibt dann ein $\emptyset \neq I_0 \subset \{1, \dots, m-1\}$, so dass $\sum_{i \in I} p_i \leq \mu'(\cup_{i \in I} A_i)$ für alle $I \subset I_0$, wobei für $I = I_0$

die Gleichheit gilt. Insbesondere haben wir $I \neq \{1, \dots, m-1\}$. Es gibt dann Masse λ_i auf S für $i \in I_0$, so dass $\lambda_i(A_i) = \lambda_i(S) = p_i$ für $i \in I_0$ und $\sum_{i \in I_0} \lambda_i(A) \leq \mu'(A)$ für alle $A \in \mathfrak{B}(S)$. Sei $p'_i = p_i$ für $i \leq m-1$ und $p'_m = p_m - \varepsilon_0$. Wir setzen $B_0 = \cup_{i \in I_0} A_i$ und definieren $\mu''(A) = \mu'(A) - \mu'(A \cap B_0)$. Sei $I_1 = \{1, \dots, m\} \setminus I_0$. Sei $I \subset I_1$. Wir bemerken zuerst, dass falls $m \in I$,

$$\sum_{i \in I \cup I_0} p'_i \leq \mu\left(\bigcup_{i \in I \cup I_0} A_i\right) - \varepsilon_0 = \mu'\left(\bigcup_{i \in I \cup I_0} A_i\right). \quad (2.2)$$

Es gilt dann für beliebiges $I \subset I_1$

$$\begin{aligned} \sum_{i \in I} p'_i + \mu'(B_0) &= \sum_{i \in I \cup I_0} p'_i \leq \mu'\left(\bigcup_{i \in I \cup I_0} A_i\right) \\ &= \mu'\left(\bigcup_{i \in I} A_i\right) + \mu'(B_0) - \mu'\left(\left(\bigcup_{i \in I} A_i\right) \cap B_0\right) = \mu''\left(\bigcup_{i \in I} A_i\right) + \mu'(B_0). \end{aligned}$$

Hier haben wir die Definition von I_0 verwendet, und die Ungleichung gilt wegen (2.2), falls $m \in I$ und der Definition von ε_0 , falls $m \notin I$. Also haben wir gezeigt, dass für jedes $I \subset I_1$

$$\sum_{i \in I} p'_i \leq \mu''\left(\bigcup_{i \in I} A_i\right).$$

Wegen der Induktionsvoraussetzung gibt es Masse λ'_i auf S , so dass $\lambda'_i(A_i) = \lambda'_i(S) = p'_i$ für $i \in I_1$ und $\sum_{i \in I_1} \lambda'_i(A) \leq \mu''(A)$. Wir setzen nun $\lambda_i = \lambda'_i$ für $i \in I_1 \setminus \{m\}$ und $\lambda_m = \lambda'_m + \varepsilon_0 \eta$. Dann ist $\lambda_i(A_i) = \lambda_i(S) = p_i$ für $i \in I_1$. Weiter gilt für $A \in \mathfrak{B}(S)$

$$\begin{aligned} \sum_{i=1}^m \lambda_i(A) &= \sum_{i \in I_0} \lambda_i(A) + \sum_{i \in I_1} \lambda_i(A) = \sum_{i \in I_0} \lambda_i(A \cap B_0) + \sum_{i \in I_1} \lambda'_i(A) + \varepsilon_0 \eta(A) \\ &\leq \mu'(A \cap B_0) + \mu''(A) + \varepsilon_0 \eta(A) = \mu'(A) + \varepsilon_0 \eta(A) = \mu(A). \end{aligned}$$

Somit hat $\lambda_1, \dots, \lambda_m$ die geforderten Eigenschaften. □

Korollar 2.5. Sei μ ein endliches Borelmaß auf S , $n \in \mathbb{N} \setminus \{0\}$ und $p_i \geq 0$ und $A_i \in \mathfrak{B}(S)$ für $i = 1, \dots, n$. Sei $\varepsilon > 0$, und nehmen wir an, dass

$$\sum_{i \in I} p_i \leq \mu\left[\bigcup_{i \in I} A_i\right] + \varepsilon$$

für alle $I \subset \{1, \dots, n\}$. Dann gibt es Borelmasse $\lambda_1, \dots, \lambda_n$ auf S , so dass $\lambda_i(A_i) = \lambda_i(S) \leq p_i$ für $i = 1, \dots, n$, $\sum_{i=1}^n \lambda_i(S) \geq \sum_{i=1}^n p_i - \varepsilon$ und $\sum_{i=1}^n \lambda_i(A) \leq \mu(A)$ für alle $A \in \mathfrak{B}(S)$.

Beweis. Sei $S' = S \cup \{\Delta\}$, wobei Δ ein isolierter Punkt ist, der nicht zu S gehört. Wir erweitern das Mass μ auf S' , so dass $\mu(\{\Delta\}) = \varepsilon$. Wir setzen $A'_i = A_i \cup \{\Delta\}$. Dann gilt

$$\sum_{i \in I} p_i \leq \mu \left[\bigcup_{i \in I} A'_i \right]$$

für alle $I \subset \{1, \dots, n\}$. Nach Hilfssatz 2.4 gibt es Masse $\lambda'_1, \dots, \lambda'_n$ auf S' , so dass $\lambda'_i(A'_i) = \lambda'_i(S') = p_i$ und $\sum_{i=1}^n \lambda'_i(A) \leq \mu(A)$ für alle $A \in \mathfrak{B}(S')$. Wir setzen nun $\lambda_i = \lambda'_i|_S$. Dann ist $\lambda_i(A_i) = \lambda'_i(A_i) \leq \lambda'_i(A'_i) = p_i$ und $\lambda_i(S \setminus A_i) = \lambda'_i(S \setminus A_i) = 0$. Also gilt

$$\sum_{i=1}^n \lambda_i(S) = \sum_{i=1}^n \lambda'_i(S' \setminus \{\Delta\}) = \sum_{i=1}^n [p_i - \lambda'_i(\{\Delta\})] \geq \sum_{i=1}^n p_i - \mu(\{\Delta\}) = \sum_{i=1}^n p_i - \varepsilon ,$$

wobei wir $\sum_{i=1}^n \lambda'_i(\{\Delta\}) \leq \mu(\{\Delta\})$ verwendet haben. Schliesslich gilt $\sum_{i=1}^n \lambda_i(A) = \sum_{i=1}^n \lambda'_i(A) \leq \mu(A)$ für alle $A \in \mathfrak{B}(S)$. $\square$

Hilfssatz 2.6. Sei S separabel, $\varepsilon > 0$, $\mathbb{P}, Q \in \mathcal{P}(S)$, so dass $\rho(\mathbb{P}, Q) < \varepsilon$, und $\delta > 0$. Nehmen wir an, die Mengen $E_1, \dots, E_N \in \mathfrak{B}(S)$ sind disjunkt mit Durchmesser kleiner als δ und $\mathbb{P}[E_0] < \delta$, wobei $E_0 = S \setminus (\cup_{k=1}^N E_k)$. Dann gibt es Konstanten $c_1, \dots, c_N \in [0, 1]$ und unabhängige Zufallsvariablen $X, Y_0, \dots, Y_N$ (mit Werten in S) und ζ (mit Werten in $[0, 1]$), die auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \nu)$ definiert sind, so dass X die Verteilung $\mathbb{P}$ hat, ζ auf $[0, 1]$ gleichverteilt ist, und

$$Y = \begin{cases} Y_i, & \text{auf } \{X \in E_i, \zeta \geq c_i\}, i = 1, \dots, N, \\ Y_0, & \text{sonst,} \end{cases} \quad (2.3)$$

die Verteilung Q hat, und

$$\{d(X, Y) \geq \delta + \varepsilon\} \subset \{X \in E_0\} \cup \left\{ \zeta < \max \left\{ \frac{\varepsilon}{\mathbb{P}[E_i]} : \mathbb{P}[E_i] > 0 \right\} \right\} ,$$

und

$$\nu[d(X, Y) \geq \delta + \varepsilon] \leq \delta + \varepsilon .$$

Beweis. Sei $p_i = \mathbb{P}[E_i]$ und $A_i = E_i^\varepsilon$ für $i = 1, \dots, N$. Dann gilt

$$\sum_{i \in I} p_i \leq \mathbb{P} \left[\overline{\bigcup_{i \in I} E_i} \right] \leq Q \left[\bigcup_{i \in I} A_i \right] + \varepsilon$$

für alle $I \subset \{1, \dots, N\}$. Nach Korollar 2.5 gibt es dann Borelmasse $\lambda_1, \dots, \lambda_N$ auf S , so dass $\lambda_i[A_i] = \lambda_i[S] \leq p_i$ für $i = 1, \dots, N$, $\sum_{i=1}^N \lambda_i[S] \geq \sum_{i=1}^N p_i - \varepsilon$ und

$\sum_{i=1}^N \lambda_i[A] \leq Q[A]$ für alle $A \in \mathfrak{B}(S)$. Insbesondere gilt $p_i - \lambda_i(S) \leq \varepsilon$. Wir setzen nun $c_i = (p_i - \lambda_i[S])/p_i \in [0, 1]$ mit der Interpretation $0/0 = 0$. Wir haben, $c_i \leq \varepsilon/p_i$ falls $p_i > 0$. Es gilt $(1 - c_i)\mathbb{P}[E_i] = \lambda_i(S)$ und

$$\mathbb{P}[E_0] + \sum_{i=1}^N c_i \mathbb{P}[E_i] = 1 - \sum_{i=1}^N \mathbb{P}[E_i] + \sum_{i=1}^N c_i \mathbb{P}[E_i] = 1 - \sum_{i=1}^N \lambda_i[S] .$$

Wir definieren nun die Masse

$$Q_i[B] = \frac{\lambda_i[B]}{(1 - c_i)\mathbb{P}[E_i]}$$

für $i = 1, \dots, N$ und

$$Q_0[B] = \frac{Q[B] - \sum_{i=1}^N \lambda_i[B]}{\mathbb{P}[E_0] + \sum_{i=1}^N c_i \mathbb{P}[E_i]} .$$

Wir definieren nun unabhängige Zufallsvariablen $X, Y_0, \dots, Y_N, \zeta$ auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \nu)$ mit den Verteilungen $\mathbb{P}, Q_0, \dots, Q_N$ und uniform auf $[0, 1]$. Die Variablen $Y_1, \dots, Y_N$ nehmen dann Werte in $A_1, \dots, A_N$ an. Definieren wir nun Y durch (2.3). Dann haben wir für alle $B \in \mathfrak{B}(S)$

$$\begin{aligned} \nu[Y \in B] &= \sum_{i=1}^N Q_i(B)(1 - c_i)\mathbb{P}[E_i] + Q_0[B] \left(\mathbb{P}[E_0] + \sum_{i=1}^N c_i \mathbb{P}[E_i] \right) \\ &= \sum_{i=1}^N \lambda_i[B] + \left(Q[B] - \sum_{i=1}^N \lambda_i[B] \right) = Q[B] . \end{aligned}$$

Also hat Y die Verteilung Q . Es gilt nun $\{X \in E_i, \zeta \geq c_i\} \subset \{X \in E_i, Y \in A_i\} \subset \{d(X, Y) < \delta + \varepsilon\}$. Also erhalten wir

$$\begin{aligned} \{d(X, Y) \geq \delta + \varepsilon\} &\subset \{X \in E_0\} \cup \bigcup_{i=1}^N \{X \in E_i, \zeta < c_i\} \subset \{X \in E_0\} \cup \left\{ \zeta < \bigvee_{i=1}^N c_i \right\} \\ &\subset \{X \in E_0\} \cup \left\{ \zeta < \max \left\{ \frac{\varepsilon}{\mathbb{P}[E_i]} : \mathbb{P}[E_i] > 0 \right\} \right\} . \end{aligned}$$

Die obige Formel ergibt auch

$$\nu[d(X, Y) \geq \delta + \varepsilon] \leq \mathbb{P}[E_0] + \sum_{i=1}^N c_i \mathbb{P}[E_i] = \mathbb{P}[E_0] + \sum_{i=1}^N (p_i - \lambda_i[S]) \leq \delta + \varepsilon ,$$

was die Aussage beweist. □

Satz 2.7. Sei (S, d) separabel und seien $\mathbb{P}, Q \in \mathcal{P}(S)$. Sei $\mathcal{M}(\mathbb{P}, Q)$ die Menge aller $\mu \in \mathcal{P}(S \times S)$ mit Randverteilungen $\mathbb{P}$ und Q . Dann gilt

$$\rho(\mathbb{P}, Q) = \inf_{\mu \in \mathcal{M}(\mathbb{P}, Q)} \inf\{\varepsilon > 0 : \mu[\{(x, y) : d(x, y) \geq \varepsilon\}] \leq \varepsilon\}.$$

Beweis. Aus $\mu[\{(x, y) : d(x, y) \geq \varepsilon\}] \leq \varepsilon$ für ein $\mu \in \mathcal{M}(\mathbb{P}, Q)$ folgt für $F \in \mathfrak{B}(S)$

$$\mathbb{P}[F] = \mu[F \times S] \leq \mu[(F \times S) \cap \{(x, y) : d(x, y) < \varepsilon\}] + \varepsilon \leq \mu[S \times F^\varepsilon] + \varepsilon = Q[F^\varepsilon] + \varepsilon.$$

Somit gilt $\rho(\mathbb{P}, Q) \leq \varepsilon$. Dies beweist die eine Ungleichung. Die andere Ungleichung folgt aus Hilfssatz 2.6. $\square$

Korollar 2.8. Sei (S, d) separabel. Seien $\{X_n\}$ und X Zufallsvariablen mit Werten in S die auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \nu)$ definiert sind mit Verteilungen $\mathbb{P}_n$ und $\mathbb{P}$. Gilt $d(X_n, X) \xrightarrow{\mathbb{P}} 0$ für $n \rightarrow \infty$, dann gilt $\mathbb{P}_n \rightarrow \mathbb{P}$.

Beweis. Sei μ_n die gemeinsame Verteilung von X_n und X . Dann konvergiert $\mu_n[\{(x, y) : d(x, y) \geq \varepsilon\}] \rightarrow 0$ für jedes $\varepsilon > 0$. Somit folgt das Resultat aus Satz 2.7. $\square$

Satz 2.9. Ist S separabel, dann ist auch $\mathcal{P}(S)$ separabel. Ist zusätzlich (S, d) vollständig, dann ist auch $(\mathcal{P}(S), \rho)$ vollständig.

Beweis. Sei $\{x_n\}$ eine dichte Teilmenge von S , und sei $\delta_x \in \mathcal{P}(S)$ das Dirac-Mass in x . Sei

$$\mathcal{M} = \left\{ \sum_{k=1}^N a_k \delta_{x_k} : N \in \mathbb{N} \setminus \{0\}, a_k \in \mathbb{Q}_+, \sum_{k=1}^N a_k = 1 \right\}.$$

Es ist nicht schwer zu zeigen, dass $\mathcal{M}$ dicht ist.

Sei nun $\{\mathbb{P}_n\} \subset \mathcal{P}(S)$ eine Folge mit $\rho(\mathbb{P}_{n-1}, \mathbb{P}_n) < 2^{-n}$ für $n \geq 2$. Wir wählen nun disjunkte Mengen $E_1^{(n)}, \dots, E_{N_n}^{(n)} \in \mathfrak{B}(S)$ mit Durchmesser kleiner als 2^{-n} , so dass $\mathbb{P}_{n-1}[E_0^{(n)}] < 2^{-n}$, wobei $E_0^{(n)} = S \setminus (\cup_{k=1}^{N_n} E_k^{(n)})$. Nach Hilfssatz 2.6 gibt es einen Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \nu)$, auf dem die unabhängigen Variablen $Y_0^{(n)}, \dots, Y_{N_n}^{(n)}$ (S -wertig) und $\zeta^{(n)}$ (gleichverteilt auf $[0, 1]$) sowie eine Variable X_1 mit Verteilung $\mathbb{P}_1$. Weiter gibt es Konstanten $c_1^{(n)}, \dots, c_{N_n}^{(n)} \in [0, 1]$, so dass für $n \geq 2$ die Zufallsvariablen

$$X_n = \begin{cases} Y_i^{(n)}, & \text{auf } \{X_{n-1} \in E_i^{(n)}, \zeta^{(n)} \geq c_i^{(n)}\}, i = 1, \dots, N_n, \\ Y_0^{(n)}, & \text{sonst,} \end{cases}$$

die Verteilung $\mathbb{P}_n$ haben, und $\nu[d(X_{n-1}, X_n) \geq 2^{-n+1}] \leq 2^{-n+1}$. Aus dem Borel-Cantelli-Lemma folgern wir, dass

$$\nu \left[\sum_{n=2}^{\infty} d(X_{n-1}, X_n) < \infty \right] = 1 .$$

Da (S, d) vollständig ist, gibt es eine Variable X , so dass $X_n \rightarrow X$. Sei $\mathbb{P}$ die Verteilung von X . Es folgt nun aus Korollar 2.8, dass $\rho(\mathbb{P}_n, \mathbb{P})$ nach 0 konvergiert. Somit ist $(\mathcal{P}(S), \rho)$ vollständig. $\square$

Satz 2.10. *Sei (S, r) separabel und $\mathbb{P}_n$ und $\mathbb{P}$ seien aus $\mathcal{P}(S)$ mit $\rho(\mathbb{P}_n, \mathbb{P}) \rightarrow 0$. Dann gibt es einen Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \nu)$ auf dem Variablen X_n und X mit den Verteilungen $\mathbb{P}_n$ und $\mathbb{P}$ definiert sind, und $X_n \rightarrow X$.*

Beweis. Für $k \in \mathbb{N} \setminus \{0\}$ gibt es disjunkte messbare Mengen $E_1^{(k)}, \dots, E_{N_k}^{(k)}$ mit Durchmesser kleiner als 2^{-k} und $\mathbb{P}[E_0^{(k)}] \leq 2^{-k}$, wobei $E_0^{(k)} = S \setminus (\bigcup_{i=1}^{N_k} E_i^{(k)})$. Wir dürfen annehmen, dass $\varepsilon_k = \min_{1 \leq i \leq N_k} \mathbb{P}[E_i^{(k)}] > 0$. Setzen wir $k_n = \max\{1\} \cup \{k \geq 1 : \rho(\mathbb{P}_n, \mathbb{P}) < \varepsilon_k/k\}$. Wir können annehmen, dass $k_n < \infty$, da wir für die anderen n einfach $X_n = X$ setzen können. Wir wenden nun Hilfssatz 2.6 an, wobei wir $Q = \mathbb{P}_n$, $\varepsilon = \varepsilon_{k_n}$ für $k_n > 1$ und $\varepsilon = \rho(\mathbb{P}_n, \mathbb{P}) + 1/n$ für $k_n = 1$, $\delta = 2^{-k_n}$, $E_i = E_i^{(k_n)}$ und $N = N_{k_n}$ wählen. Es gibt dann also einen Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \nu)$ auf dem $Y_0^{(n)}, \dots, Y_{N_{k_n}}^{(n)}, \zeta$ gleichverteilt auf $[0, 1]$ und X mit Verteilung $\mathbb{P}$ existieren, die alle unabhängig sind, so dass es Konstanten $c_1^{(n)}, \dots, c_{N_{k_n}}^{(n)}$ in $[0, 1]$ gibt, so dass

$$X_n = \begin{cases} Y_i^{(n)}, & \text{auf } \{X \in E_i^{(k_n)}, \zeta \geq c_i^{(n)}\}, i = 1, \dots, N_{k_n}, \\ Y_0^{(n)}, & \text{sonst,} \end{cases}$$

die Verteilung $\mathbb{P}_n$ hat, und, falls $k_n > 1$,

$$\{r(X_n, X) \geq 2^{-k_n} + \varepsilon_{k_n}/k_n\} \subset \{X \in E_0^{(k_n)}\} \cup \{\zeta < 1/k_n\} .$$

Setzen wir $K_n = \min_{m \geq n} k_m$. Ist $K_n > 1$, so erhalten wir

$$\begin{aligned} \nu \left[\bigcup_{m=n}^{\infty} \{r(X_m, X) \geq 2^{-k_m} + \varepsilon_{k_m}/k_m\} \right] &\leq \sum_{k=K_n}^{\infty} \nu[X \in E_0^{(k)}] + \nu[\zeta < 1/k_n] \\ &\leq 2^{-K_n} + \frac{1}{K_n} . \end{aligned}$$

Da $\lim_{n \rightarrow \infty} K_n = \infty$, folgt dass $X_n \rightarrow X$. $\square$

Korollar 2.11. (Continuous Mapping Theorem) Seien (S, r) und (S', r') separable metrische Räume und sei $h : S \rightarrow S'$ Borel-messbar. Seien $\mathbb{P}_n$ und $\mathbb{P}$ in $\mathcal{P}(S)$ so dass $\rho(\mathbb{P}_n, \mathbb{P}) \rightarrow 0$. Dann sind $Q_n = \mathbb{P}_n h^{-1}$ und $Q = \mathbb{P} h^{-1}$ in $\mathcal{P}(S')$. Sei $C_h \subset S$ die Menge der Punkte, in denen h stetig ist. Ist $\mathbb{P}[C_h] = 1$, dann gilt $\rho'(Q_n, Q) \rightarrow 0$, wobei ρ' die Prohorov Metrik auf $\mathcal{P}(S')$ ist.

Beweis. Nach Satz 2.10 gibt es einen Wahrscheinlichkeitsraum, auf dem Variablen X_n und X definiert sind, so dass $X_n \rightarrow X$. Dann gilt auch $h(X_n) \rightarrow h(X)$. Nach Korollar 2.8 folgt daraus, dass $\rho'(Q_n, Q) \rightarrow 0$. $\square$

2.2. Der Satz von Prohorov

Definition 2.12. Ein Mass $\mathbb{P} \in \mathcal{P}(S)$ heisst **straff**, falls es für jedes $\varepsilon > 0$ eine kompakte Menge $K \subset S$ gibt, so dass $\mathbb{P}[K] \geq 1 - \varepsilon$. Eine Menge $\mathcal{M} \subset \mathcal{P}(S)$ heisst **straff**, falls es für jedes $\varepsilon > 0$ eine kompakte Menge $K \subset S$ gibt, so dass $\mathbb{P}[K] \geq 1 - \varepsilon$ für alle $\mathbb{P} \in \mathcal{M}$.

Hilfssatz 2.13. Falls (S, r) vollständig und separabel ist, dann ist jedes $\mathbb{P} \in \mathcal{P}(S)$ straff.

Beweis. Sei $\{x_k\}$ dicht in S und wählen wir $\varepsilon > 0$. Dann gibt es natürliche Zahlen N_n , so dass

$$\mathbb{P}\left[\bigcup_{k=1}^{N_n} B(x_k, 1/n)\right] \geq 1 - \frac{\varepsilon}{2^n}.$$

Sei K der Abschluss von $\bigcap_{n \geq 1} \bigcup_{k=1}^{N_n} B(x_k, 1/n)$. Dann ist K total beschränkt und abgeschlossen, und daher kompakt. Weiter gilt

$$\mathbb{P}[K] \geq 1 - \sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} = 1 - \varepsilon.$$

$\square$

Satz 2.14. (Satz von Prohorov) Sei (S, r) vollständig und separabel, und $\mathcal{M} \subset \mathcal{P}(S)$. Dann sind folgende Aussage äquivalent:

i) $\mathcal{M}$ ist straff.

ii) Für jedes $\varepsilon > 0$ gibt es eine kompakte Menge $K \subset S$, so dass

$$\inf_{P \in \mathcal{M}} \mathbb{P}[K^\varepsilon] \geq 1 - \varepsilon .$$

iii) $\mathcal{M}$ ist relativ kompakt.

Beweis. “i) $\Rightarrow$ ii)” Trivial.

“ii) $\Rightarrow$ iii)” Da $\overline{\mathcal{M}}$ vollständig ist (Satz 2.9), genügt es zu zeigen, dass $\mathcal{M}$ total beschränkt ist. Sei $\varepsilon > 0$ und $0 < \delta < \varepsilon/2$. Wir wählen K , so dass $\inf_{P \in \mathcal{M}} \mathbb{P}[K^\delta] \geq 1 - \delta$. Es gibt dann eine endliche Menge $\{x_1, \dots, x_n\}$, so dass für $B_i = B(x_i, 2\delta)$ gilt, dass $K^\delta \subset \bigcup_{i=1}^n B_i$. Dies ist der Fall, da K mit endlich vielen δ -Umgebungen überdeckt werden kann. Wir setzen $E_i = B_i \setminus \bigcup_{j=1}^{i-1} B_j$. Sei $x_0 \in S$ und $m \geq n/\delta$. Sei

$$\mathcal{N} = \left\{ \sum_{i=0}^n \frac{k_i}{m} \delta_{x_i} : 0 \leq k_i \leq m, \sum_{i=0}^n k_i = m \right\} .$$

Sei $Q \in \mathcal{M}$. Wir setzen $k_i = \lfloor mQ[E_i] \rfloor$ und $k_0 = m - \sum_{i=1}^n k_i$. Da die Mengen $\{E_i\}$ disjunkt sind, gilt $k_0 \geq 0$. Mit $\mathbb{P} = \sum_{i=0}^n (k_i/m) \delta_{x_i} \in \mathcal{N}$ erhalten wir für eine abgeschlossene Menge F

$$Q[F] \leq Q\left[\bigcup_{F \cap E_i \neq \emptyset} E_i\right] + \delta \leq \sum_{F \cap E_i \neq \emptyset} \frac{\lfloor mQ[E_i] \rfloor}{m} + \frac{n}{m} + \delta \leq \mathbb{P}[F^{2\delta}] + 2\delta .$$

Somit ist $\rho(Q, \mathbb{P}) \leq 2\delta < \varepsilon$. Da $\mathcal{N}$ eine endliche Menge ist, ist $\mathcal{M}$ total beschränkt.

“iii) $\Rightarrow$ i)” Sei $\varepsilon > 0$. Da $\mathcal{M}$ total beschränkt ist, gibt es für jedes $n \in \mathbb{N} \setminus \{0\}$ eine endliche Menge $\mathcal{M}_n \subset \mathcal{M}$, so dass $\mathcal{M} \subset \{Q : \inf_{\mathbb{P} \in \mathcal{M}_n} \rho(\mathbb{P}, Q) < \varepsilon 2^{-n-1}\}$. Aus Hilfssatz 2.13 schliessen wir, dass wir für jedes n eine kompakte Menge $K_n \subset S$ wählen können, so dass $\inf_{\mathbb{P} \in \mathcal{M}_n} \mathbb{P}[K_n] \geq 1 - \varepsilon 2^{-n-1}$. Für $Q \in \mathcal{M}$ gilt dann, dass es ein $\mathbb{P} \in \mathcal{M}_n$ gibt, so dass

$$Q[K_n^{\varepsilon 2^{-n-1}}] \geq \mathbb{P}[K_n] - \varepsilon 2^{-n-1} \geq 1 - \varepsilon 2^{-n} .$$

Wir setzen $K = \overline{\bigcap_{n=1}^{\infty} K_n^{\varepsilon 2^{-n-1}}}$. Dann ist K kompakt, da K abgeschlossen und total beschränkt ist. Wir haben für $Q \in \mathcal{M}$,

$$Q[K] \geq 1 - \sum_{i=1}^{\infty} \varepsilon 2^{-n} = 1 - \varepsilon .$$

Dies beweist die Behauptung. □

Proposition 2.15. Seien (S_k, r_k) metrische Räume und $S = \prod_{k=1}^{\infty} S_k$ sei der Raum mit der Metrik $r(x, y) = \sum_{k=1}^{\infty} 2^{-k}(r_k(x_k, y_k) \wedge 1)$. Sei $\{\mathbb{P}_\alpha\} \subset \mathcal{P}(S)$, und $\mathbb{P}_\alpha^k = \mathbb{P}_\alpha \pi_k^{-1} \in \mathcal{P}(S_k)$ sei die Randverteilung, wobei π_k die Koordinatenabbildung bezeichnet. Dann ist $\{\mathbb{P}_\alpha\}$ genau dann straff, wenn $\{\mathbb{P}_\alpha^k\}$ für jedes k straff ist.

Beweis. Sei $\{\mathbb{P}_\alpha^k\}$ für alle k straff und sei $\varepsilon > 0$. Dann gibt es eine kompakte Menge $K_k \subset S_k$, so dass $\inf_\alpha \mathbb{P}_\alpha^k[K_k] \geq 1 - \varepsilon 2^{-k}$. Dann ist $K = \cap_{k=1}^{\infty} \pi_K^{-1}(K_k)$ kompakt in S , und

$$\mathbb{P}_\alpha[K] \geq 1 - \sum_{k=1}^{\infty} (1 - \mathbb{P}_\alpha^k[K_k]) \geq 1 - \varepsilon .$$

Also ist $\{\mathbb{P}_\alpha\}$ straff.

Sei nun $\{\mathbb{P}_\alpha\}$ straff. Ist $K \subset S$ kompakt, so ist $\pi_k(K)$ kompakt in S_k und $\mathbb{P}_\alpha^k[\pi_k(K)] \geq \mathbb{P}_\alpha[K]$. Somit ist $\{\mathbb{P}_\alpha^k\}$ für jedes k straff. $\square$

2.3. Schwache Konvergenz

Wir bezeichnen mit $\bar{C}(S)$ dem Raum der reellen beschränkten stetigen Funktionen auf einem metrischen Raum (S, r) . Auf $\bar{C}(S)$ verwenden wir die Supremumsnorm. Ist $F \in \mathfrak{B}(S)$, so ist $r(x, F) = \inf\{r(x, y) : y \in F\}$ der Abstand von x zu F .

Definition 2.16. Eine Folge $\{\mathbb{P}_n\} \subset \mathcal{P}(S)$ **konvergiert schwach** gegen ein Mass $\mathbb{P} \in \mathcal{P}(S)$, falls für jedes $f \in \bar{C}(S)$ der Grenzwert $\lim_{n \rightarrow \infty} \int f \, d\mathbb{P}_n = \int f \, d\mathbb{P}$ gilt. Wir schreiben $\mathbb{P}_n \Rightarrow \mathbb{P}$. Eine Folge von S -wertigen Zufallsvariablen $\{X_n\}$ **konvergiert schwach** gegen X , falls die entsprechenden Masse konvergieren. Wir schreiben $X_n \Rightarrow X$.

Eine Menge A heisst $\mathbb{P}$ -stetig, falls $\mathbb{P}[\partial A] = 0$.

Satz 2.17. Sei (S, r) ein metrischer Raum und $\{\mathbb{P}_n\} \subset \mathcal{P}(S)$ und $\mathbb{P} \in \mathcal{P}(S)$. Dann impliziert

$$\text{i) } \lim_{n \rightarrow \infty} \rho(\mathbb{P}_n, \mathbb{P}) = 0$$

die folgenden äquivalenten Bedingungen:

$$\text{ii) } \mathbb{P}_n \Rightarrow \mathbb{P} ,$$

- iii) $\lim_{n \rightarrow \infty} \int f \, d\mathbb{P}_n = \int f \, d\mathbb{P}$ für alle gleichmässig stetigen $f \in \bar{C}(S)$,
- iv) $\overline{\lim}_{n \rightarrow \infty} \mathbb{P}_n[F] \leq \mathbb{P}[F]$ für alle abgeschlossenen Mengen $F \subset S$,
- v) $\underline{\lim}_{n \rightarrow \infty} \mathbb{P}_n[G] \geq \mathbb{P}[G]$ für alle offenen Mengen $G \subset S$,
- vi) $\lim_{n \rightarrow \infty} \mathbb{P}_n[A] = \mathbb{P}[A]$ für alle $\mathbb{P}$ -stetigen Mengen $A \subset S$.

Ist S separabel, dann sind alle Bedingungen i)–vi) äquivalent.

Beweis. “i) $\Rightarrow$ ii)” Wir setzen $\varepsilon_n = \rho(\mathbb{P}_n, \mathbb{P}) + 1/n$. Sei $f \in \bar{C}(S)$ mit $f \geq 0$. Dann gilt

$$\int f \, d\mathbb{P}_n = \int_0^{\|f\|} \mathbb{P}_n[f \geq t] \, dt \leq \int_0^{\|f\|} \mathbb{P}[\{f \geq t\}^{\varepsilon_n}] \, dt + \varepsilon_n \|f\|.$$

Somit gilt

$$\overline{\lim}_{n \rightarrow \infty} \int f \, d\mathbb{P}_n \leq \lim_{n \rightarrow \infty} \int_0^{\|f\|} \mathbb{P}[\{f \geq t\}^{\varepsilon_n}] \, dt = \int_0^{\|f\|} \mathbb{P}[f \geq t] \, dt = \int f \, d\mathbb{P}.$$

Insbesondere haben wir gezeigt, dass für jedes $f \in \bar{C}(S)$

$$\|f\| + \overline{\lim}_{n \rightarrow \infty} \int f \, d\mathbb{P}_n = \overline{\lim}_{n \rightarrow \infty} \int (\|f\| + f) \, d\mathbb{P}_n \leq \int (\|f\| + f) \, d\mathbb{P} = \|f\| + \int f \, d\mathbb{P}$$

und

$$\|f\| - \underline{\lim}_{n \rightarrow \infty} \int f \, d\mathbb{P}_n = \overline{\lim}_{n \rightarrow \infty} \int (\|f\| - f) \, d\mathbb{P}_n \leq \int (\|f\| - f) \, d\mathbb{P} = \|f\| - \int f \, d\mathbb{P}.$$

Dies beweist die Behauptung.

“ii) $\Rightarrow$ iii)” Trivial.

“iii) $\Rightarrow$ iv)” Sei F abgeschlossen und $\varepsilon > 0$. Wir betrachten die Funktion $f_\varepsilon(x) = (1 - \varepsilon^{-1}r(x, F))^+$. Dann gilt $\overline{\lim}_{n \rightarrow \infty} \mathbb{P}_n[F] \leq \overline{\lim}_{n \rightarrow \infty} \int f_\varepsilon \, d\mathbb{P}_n \leq \mathbb{P}[F^\varepsilon]$. Lassen wir $\varepsilon \rightarrow 0$ folgt die Behauptung.

“iv) $\Leftrightarrow$ v)” Trivial.

“iv) und v) $\Rightarrow$ vi)” Wir erhalten

$$\overline{\lim}_{n \rightarrow \infty} \mathbb{P}_n[A] \leq \overline{\lim}_{n \rightarrow \infty} \mathbb{P}_n[\bar{A}] \leq \mathbb{P}[\bar{A}] = \mathbb{P}[A]$$

und

$$\underline{\lim}_{n \rightarrow \infty} \mathbb{P}_n[A] \geq \underline{\lim}_{n \rightarrow \infty} \mathbb{P}_n[\mathring{A}] \geq \mathbb{P}[\mathring{A}] = \mathbb{P}[A].$$

“vi) $\Rightarrow$ ii)” Sei $f \in \bar{C}(S)$ und $f \geq 0$. Dann ist $\partial\{f \geq t\} \subset \{f = t\}$. Somit gibt es höchstens abzählbar viele t für die $\{f \geq t\}$ keine $\mathbb{P}$ -stetige Menge ist. Aus der beschränkten Konvergenz erhalten wir dann

$$\lim_{n \rightarrow \infty} \int f \mathbb{P}_n = \lim_{n \rightarrow \infty} \int_0^{\|f\|} \mathbb{P}_n[f \geq t] dt = \int_0^{\|f\|} \mathbb{P}[f \geq t] dt = \int f d\mathbb{P}.$$

Die Aussage folgt nun aus $f = f^+ - f^-$.

“v) $\Rightarrow$ i)” Sei S separabel. Sei $0 < \varepsilon < 2$ und $E_1, E_2, \dots \in \mathfrak{B}(S)$ eine Partition von S , so dass der Durchmesser von E_i echt kleiner als $\varepsilon/2$ ist. Sei $N = \inf\{n \in \mathbb{N} : \mathbb{P}[\bigcup_{i=1}^n E_i] > 1 - \varepsilon/2\}$. Seien $\mathcal{G}$ die Mengen der Form $(\bigcup_{i \in I} E_i)^{\varepsilon/2}$ mit $I \subset \{1, \dots, N\}$. Es gibt dann also ein $n_0 \in \mathbb{N}$, so dass $\mathbb{P}[G] < \mathbb{P}_n[G] + \varepsilon/2$ für alle $n \geq n_0$ und $G \in \mathcal{G}$. Für $F \in \mathcal{C}$ sei $F_0 = \bigcup\{E_i : i \leq N, E_i \cap F \neq \emptyset\}$. Dann ist $F_0^{\varepsilon/2} \in \mathcal{G}$ und

$$\mathbb{P}[F] \leq \mathbb{P}[F_0^{\varepsilon/2}] + \varepsilon/2 \leq \mathbb{P}_n[F_0^{\varepsilon/2}] + \varepsilon \leq \mathbb{P}_n[F^\varepsilon] + \varepsilon$$

für alle $n \geq n_0$. Also ist $\rho(\mathbb{P}_n, \mathbb{P}) \leq \varepsilon$ für alle $n \geq n_0$. $\square$

Korollar 2.18. Sei $\{\mathbb{P}_n\} \subset \mathcal{P}(S)$ und $\mathbb{P} \in \mathcal{P}(S)$ und $S' \in \mathfrak{B}(S)$. Nehmen wir an, dass $\mathbb{P}_n[S'] = \mathbb{P}[S'] = 1$. Wir bezeichnen mit $\mathbb{P}'_n$ beziehungsweise $\mathbb{P}'$ die Einschränkungen auf $\mathfrak{B}(S')$. Dann gilt $\mathbb{P}_n \Rightarrow \mathbb{P}$ genau dann, wenn $\mathbb{P}'_n \Rightarrow \mathbb{P}'$.

Beweis. Nehmen wir $\mathbb{P}_n \Rightarrow \mathbb{P}$ an. Sei $G' \subset S'$ offen in S' . Dann existiert $G \subset S$ offen in S , so dass $G' = G \cap S'$. Also gilt

$$\varliminf_{n \rightarrow \infty} \mathbb{P}'_n[G'] = \varliminf_{n \rightarrow \infty} \mathbb{P}_n[G \cap S'] = \varliminf_{n \rightarrow \infty} \mathbb{P}_n[G] \geq \mathbb{P}[G] = \mathbb{P}[G'] = \mathbb{P}'[G'].$$

Also gilt $\mathbb{P}'_n \Rightarrow \mathbb{P}'$. Nehmen wir $\mathbb{P}'_n \Rightarrow \mathbb{P}'$ an. Sei G offen in S . Dann ist $G' = G \cap S'$ offen in S' . Also haben wir

$$\varliminf_{n \rightarrow \infty} \mathbb{P}_n[G] = \varliminf_{n \rightarrow \infty} \mathbb{P}_n[G'] = \varliminf_{n \rightarrow \infty} \mathbb{P}'_n[G'] \geq \mathbb{P}'[G'] = \mathbb{P}[G'] = \mathbb{P}[G].$$

Also gilt auch $\mathbb{P}_n \Rightarrow \mathbb{P}$. $\square$

2.4. Trennende Mengen

Proposition 2.19. *Ist S separabel, so gibt es eine Folge $\{f_n\} \subset \bar{C}(S)$ mit $f_n \geq 0$, so dass jede Funktion aus $B(S)$ erhalten werden kann als punktweiser Grenzwert einer beschränkten Folge von Funktionen aus dem Raum der Funktionen, die von $\{f_n\}$ aufgespannt wird.*

Beweis. Sei $\{x_i\} \subset S$ dicht in S . Für jede offene Menge $G \subset S$, die als endlicher Durchschnitt von Kugeln $B(x_i, 1/k)$ geschrieben werden kann, wählen wir eine beschränkte Folge von Funktionen $\{f_n^G\} \subset \bar{C}(S)$, so dass f_n^G punktweise nach $\mathbb{1}_G$ konvergiert. Die Klasse der Mengen B , für die $\mathbb{1}_B$ als punktweiser Grenzwert erhalten werden kann, bildet eine Dynkin Klasse. Daraus folgt die Aussage. $\square$

Definition 2.20. *Eine Menge $M \subset \bar{C}(S)$ heisst **trennend**, falls es für zwei Masse $\mathbb{P}, Q \in \mathcal{P}(S)$ mit $\mathbb{P} \neq Q$ eine Funktion $f \in M$ gibt, so dass $\int f \, d\mathbb{P} \neq \int f \, dQ$.*

*M heisst **punkttrennend**, falls es für jedes $x, y \in S$ mit $x \neq y$ eine Funktion $f \in M$ gibt, so dass $f(x) \neq f(y)$.*

Proposition 2.21. *Seien (S_k, r_k) separable metrische Räume, $S = \prod_{k=1}^{\infty} S_k$ der Produktraum mit der Produktmetrik $r(x, y) = \sum_{k=1}^{\infty} 2^{-k}(d_k(x_k, y_k) \wedge 1)$. Für $M_k \subset \bar{C}(S_k)$ definieren wir*

$$M = \left\{ \prod_{k=1}^n f_k(x_k) : n \geq 1, f_k \in M_k \cup \{1\} \right\}.$$

Ist M_k trennend für alle k , so ist M trennend.

Beweis. Seien $\mathbb{P}$ und Q zwei Masse auf S . Nehmen wir an, dass

$$\int f_1(x_1) \cdots f_n(x_n) \, d\mathbb{P}(x) = \int f_1(x_1) \cdots f_n(x_n) \, dQ(x)$$

für alle $n \geq 1$ und $f_i \in M_i \cup \{1\}$. Wir definieren die signierten Masse

$$d\mu(x) = f_2(x_2) \cdots f_n(x_n) \, d\mathbb{P}(x), \quad d\nu(x) = f_2(x_2) \cdots f_n(x_n) \, dQ(x).$$

Seien μ^1 und ν^1 die Randverteilungen von μ und ν auf S_1 . Dann ist $\int f_1(x) \, d\mu^1(x) = \int f_1(x) \, d\nu^1(x)$. Da M_1 trennend ist, gilt dies auch für die endlichen signierten Masse. Also ist $\mu^1 = \nu^1$. Insbesondere gilt

$$\int \mathbb{1}_{A_1}(x_1) f_2(x_2) \cdots f_n(x_n) \, d\mathbb{P}(x) = \int \mathbb{1}_{A_1}(x_1) f_2(x_2) \cdots f_n(x_n) \, dQ(x)$$

für alle $A_1 \in \mathcal{B}(S_1)$. Gehen wir analog weiter, erhalten wir

$$\int \mathbb{1}_{A_1}(x_1) \cdots \mathbb{1}_{A_n}(x_n) \, d\mathbb{P}(x) = \int \mathbb{1}_{A_1}(x_1) \cdots \mathbb{1}_{A_n}(x_n) \, dQ(x).$$

Da $\mathcal{B}(S)$ von den Mengen der Form $A_1 \times \cdots \times A_n$ erzeugt wird, gilt $\mathbb{P} = Q$. $\square$

2.5. Der Raum $D_S[0, \infty)$

Definition 2.22. Der Raum der cadlag Funktionen $[0, \infty) \rightarrow S$ wird mit $D_S[0, \infty)$ bezeichnet. Sei Λ' die Menge aller stetigen und strikt wachsenden Funktionen $\lambda : [0, \infty) \rightarrow [0, \infty)$ mit $\lambda(0) = 0$ und $\lim_{t \rightarrow \infty} \lambda(t) = \infty$. Mit Λ bezeichnen wir die Menge aller $\lambda \in \Lambda'$, so dass

$$\gamma(\lambda) = \sup_{0 \leq t < s} \left| \log \frac{\lambda(s) - \lambda(t)}{s - t} \right| < \infty .$$

Wir definieren den Abstand

$$d(x, y) = \inf_{\lambda \in \Lambda} \left\{ \gamma(\lambda) \vee \int_0^\infty e^{-u} d(x, y, \lambda, u) \, du \right\}$$

für $x, y \in D_S[0, \infty)$, wobei $d(x, y, \lambda, u) = \sup_{t \geq 0} r(x(t \wedge u), y(\lambda(t) \wedge u)) \wedge 1$.

Hilfssatz 2.23. Ist $x \in D_S[0, \infty)$, so sind die Unstetigkeitsstellen von x abzählbar.

Beweis. Sei $A_n = \{t > 0 : r(x(t), x(t-)) > n^{-1}\}$. Die Menge A_n kann keine Häufungspunkte enthalten, da x cadlag ist. Also ist A_n abzählbar. Damit ist auch die Menge der Unstetigkeiten $\cup_n A_n$ abzählbar. $\square$

Hilfssatz 2.24. d ist eine Metrik für $D_S[0, \infty)$.

Beweis. Bezeichnen wir mit $\lambda^{-1}(s) = \inf\{t : \lambda(t) \geq s\}$ die Umkehrfunktion von λ . Wir bemerken, dass $\gamma(\lambda) = \gamma(\lambda^{-1})$. Es gilt dann $\sup_{t \geq 0} r(x(t \wedge u), y(\lambda(t) \wedge u)) = \sup_{t \geq 0} r(x(\lambda^{-1}(t) \wedge u), y(t \wedge u))$ und somit $d(x, y, \lambda, u) = d(y, x, \lambda^{-1}, u)$. Also haben wir $d(x, y) = d(y, x)$.

Sei $d(x, y) = 0$. Dann gibt es eine Folge $\{\lambda_n\} \subset \Lambda$, so dass $\gamma(\lambda_n) \rightarrow 0$ und $\int_0^\infty e^{-u} d(x, y, \lambda_n, u) \, du \rightarrow 0$. Sei $\lambda \in \Lambda$. Falls $\lambda(t) < t$, so ist

$$\frac{t - \lambda(t)}{t} = 1 - e^{-|\log(\lambda(t)/t)|} \leq 1 - e^{-\gamma(\lambda)} \leq \gamma(\lambda) . \quad (2.4a)$$

Ist $\lambda(t) > t$, so folgt

$$\frac{\lambda(t) - t}{t} = \frac{\lambda(t) - \lambda(\lambda^{-1}(t))}{t - \lambda^{-1}(t)} \frac{t - \lambda^{-1}(t)}{t} \leq e^{\gamma(\lambda)} \gamma(\lambda) . \quad (2.4b)$$

Insbesondere haben wir

$$\lim_{n \rightarrow \infty} \sup_{0 \leq t \leq T} |\lambda_n(t) - t| = 0 \quad (2.5)$$

für alle $T > 0$. Ist also t ein Punkt, an dem $y(t)$ stetig ist, so muss $x(t) = y(t)$ gelten, da $d(x, y, \lambda_n, u)$ für fast alle u gegen 0 konvergiert. Da y cadlag ist, ist y fast überall stetig (Hilfssatz 2.23). Da auch x rechtsstetig ist, muss $x = y$ gelten.

Seien $x, y, z \in D_S[0, \infty)$, $\lambda_1, \lambda_2 \in \Lambda$ und $u \geq 0$. Wir haben dann wegen der Dreiecksungleichung

$$r(x(t \wedge u), z(\lambda_2(\lambda_1(t)) \wedge u)) \leq r(x(t \wedge u), y(\lambda_1(t) \wedge u)) + r(y(\lambda_1(t) \wedge u), z(\lambda_2(\lambda_1(t)) \wedge u)) .$$

Diese Ungleichung gilt auch für r ersetzt durch $r \wedge 1$. Da $\lambda_1(t)$ alle positiven Zahlen durchläuft, erhalten wir

$$d(x, z, \lambda_2 \circ \lambda_1, u) \leq d(x, y, \lambda_1, u) + d(y, z, \lambda_2, u) .$$

Im weiteren erhalten wir für $s > t \geq 0$

$$\begin{aligned} \left| \log \frac{\lambda_2(\lambda_1(s)) - \lambda_2(\lambda_1(t))}{s - t} \right| &= \left| \log \frac{\lambda_2(\lambda_1(s)) - \lambda_2(\lambda_1(t))}{\lambda_1(s) - \lambda_1(t)} + \log \frac{\lambda_1(s) - \lambda_1(t)}{s - t} \right| \\ &\leq \gamma(\lambda_2) + \gamma(\lambda_1) . \end{aligned} \quad (2.6)$$

Also gilt $\gamma(\lambda_2 \circ \lambda_1) \leq \gamma(\lambda_1) + \gamma(\lambda_2)$. Setzen wir diese Ungleichungen zusammen, ergibt sich die Dreiecksungleichung $d(x, z) \leq d(x, y) + d(y, z)$. $\square$

Proposition 2.25. Sei $\{x_n\} \subset D_S[0, \infty)$ und $x \in D_S[0, \infty)$. Folgende Aussagen sind äquivalent:

- i) $\lim_{n \rightarrow \infty} d(x_n, x) = 0$.
- ii) Es gibt $\{\lambda_n\} \subset \Lambda$, so dass $\lim_{n \rightarrow \infty} \gamma(\lambda_n) = 0$ und

$$\lim_{n \rightarrow \infty} \sup_{0 \leq t \leq T} r(x_n(t), x(\lambda_n(t))) = 0 \quad (2.7)$$

für alle $T > 0$.

- iii) Für jedes $T > 0$ gibt es eine Folge $\{\lambda_n\} \subset \Lambda'$ (kann von T abhängen), so dass (2.5) und (2.7) gilt.

Beweis. “i) $\Rightarrow$ ii)” Es gibt $\{\lambda_n\} \subset \Lambda$ und $\{u_n\} \subset (0, \infty)$, so dass $\gamma(\lambda_n) \rightarrow 0$, $u_n \rightarrow \infty$ und $d(x_n, y_n, \lambda_n, u_n) \rightarrow 0$. Insbesondere gilt

$$\lim_{n \rightarrow \infty} \sup_{t \geq 0} r(x_n(t \wedge u_n), x(\lambda_n(t) \wedge u_n)) = 0 .$$

Da $\lambda_n(T) \leq T + \gamma(\lambda_n)e^{\gamma(\lambda_n)}T$, siehe (2.4), ist für n gross genug, $u_n \geq \max\{T, \lambda_n(T)\}$. Somit gilt die Aussage.

“ii) $\Rightarrow$ i)” Es folgt, dass

$$\lim_{n \rightarrow \infty} \sup_{t \geq 0} r(x_n(t \wedge u), x(\lambda_n(t) \wedge u)) = 0$$

für jedes $u > 0$. Da $\int_u^\infty e^{-s} d(x_n, x, \lambda_n, s) ds \leq e^{-u}$, folgt die Aussage.

“ii) $\Rightarrow$ iii)” Trivial, da (2.5) für alle $\{\lambda_n\} \subset \Lambda$ gilt, für die $\gamma(\lambda_n) \rightarrow 0$.

“iii) $\Rightarrow$ ii)” Sei $N \in \mathbb{N} \setminus \{0\}$. Wir wählen $\{\lambda_n^N\} \subset \Lambda'$, so dass die Bedingungen für $T = N$ erfüllt sind, und so dass für jedes n , $\lambda_n^N(t) = \lambda_n^N(N) + t - N$ für $t \geq N$. Definieren wir $\tau_0^N = 0$ und

$$\tau_k^N = \inf\{t > \tau_{k-1}^N : r(x(t), x(\tau_{k-1}^N)) > N^{-1}\}, \quad (2.8)$$

mit der Interpretation dass $\tau_k^N = \infty$, falls $\tau_{k-1}^N = \infty$. Da x Grenzwerte von links hat, gibt es keine Häufungspunkte von $\{\tau_k^N : k \geq 0\}$. Wir setzen $u_{k,n}^N = (\lambda_n^N)^{-1}(\tau_k^N)$ mit $(\lambda_n^N)^{-1}(\infty) = \infty$ und definieren

$$\mu_n^N(t) = \tau_k^N + \frac{t - u_{k,n}^N}{u_{k+1,n}^N - u_{k,n}^N} (\tau_{k+1}^N - \tau_k^N)$$

für $t \in [u_{k,n}^N, u_{k+1,n}^N) \cap [0, N]$, mit der Interpretation $\infty/\infty = 1$, und $\mu_n^N(t) = \mu_n^N(N) + t - N$ für $t > N$. Dann ist $\mu_n^N \in \Lambda$ mit

$$\gamma(\mu_n^N) = \max\{|\log(\tau_{k+1}^N - \tau_k^N) - \log(u_{k+1,n}^N - u_{k,n}^N)| : u_{k,n}^N < N\}.$$

Es gilt nun für alle n

$$\begin{aligned} \sup_{0 \leq t \leq N} r(x_n(t), x(\mu_n^N(t))) &\leq \sup_{0 \leq t \leq N} r(x_n(t), x(\lambda_n^N(t))) + \sup_{0 \leq t \leq N} r(x(\lambda_n^N(t)), x(\mu_n^N(t))) \\ &\leq \sup_{0 \leq t \leq N} r(x_n(t), x(\lambda_n^N(t))) + \frac{2}{N}. \end{aligned}$$

Da $\lim_{n \rightarrow \infty} u_{k,n}^N = \tau_k^N$ konvergiert $\gamma(\mu_n^N)$ nach 0. Also können wir eine Folge $1 < n_1 < n_2 < \dots$ wählen, so dass $\gamma(\mu_n^N) \leq N^{-1}$ und $\sup_{0 \leq t \leq N} r(x_n(t), x(\mu_n^N(t))) \leq 3N^{-1}$ für alle $n \geq n_N$. Für $1 \leq n < n_1$ wählen wir $\hat{\lambda}_n \in \Lambda$ beliebig. Für $n_N \leq n < n_{N+1}$ wählen wir $\hat{\lambda}_n = \mu_n^N$. Dann ist $\{\hat{\lambda}_n\} \in \Lambda$ und erfüllt die Bedingungen. $\square$

Satz 2.26. *Ist S separabel, dann ist $D_S[0, \infty)$ separabel. Ist (S, r) vollständig, dann ist auch $(D_S[0, \infty), d)$ vollständig.*

Beweis. Da $r \wedge 1$ die gleich Topologie erzeugt, und die Vollständigkeit nicht verändert wird, können wir annehmen, dass $r \leq 1$. Wir nehmen an, dass S separabel ist. Sei $\{\alpha_i : i \in \mathbb{N}\} \subset S$ eine abzählbare dichte Menge. Sei Γ die Klasse von Funktionen der Form $X(t) = \sum_{k=1}^n \alpha_{i_k} \mathbb{1}_{t_{k-1} \leq t < t_k} + \alpha_{i_{n+1}} \mathbb{1}_{t \geq t_n}$, wobei $n \geq 1$, $0 = t_0 < t_1 < \dots < t_n$ rationale Zahlen sind und $i_n \in \mathbb{N}$. Dann ist Γ abzählbar. Es ist nicht schwer zu sehen, dass Γ dicht ist in $D_S[0, \infty)$.

Sei nun (S, r) vollständig und $\{x_n\} \subset D_S[0, \infty)$ eine Cauchyfolge. Es gibt Zahlen $1 \leq N_1 < N_2 < \dots$, so dass für $n, m \geq N_k$ die Ungleichung $d(x_n, x_m) \leq 2^{-k-1}e^{-k}$ gilt. Wir setzen dann $y_k = x_{N_k}$. Wir wählen $u_k > k$ und $\lambda_k \in \Lambda$, so dass $\gamma(\lambda_k) \leq 2^{-k}$ und $d(y_k, y_{k+1}, \lambda_k, u_k) \leq 2^{-k}$. Aus (2.6) folgt, dass $\overline{\lim}_{n \rightarrow \infty} \gamma(\lambda_{k+n} \circ \dots \circ \lambda_{k+1} \circ \lambda_k) < \infty$. Insbesondere ist die Folge $\lambda_{k+n} \circ \dots \circ \lambda_{k+1} \circ \lambda_k$ beschränkt, und hat daher auf jedem beschränkten Intervall eine konvergente Teilfolge. Aus (2.4) folgt nun aber, dass $\mu_k = \lim_{n \rightarrow \infty} \lambda_{k+n} \circ \dots \circ \lambda_{k+1} \circ \lambda_k$ gleichmässig auf beschränkten Intervallen existiert, Lipschitz-stetig ist, und $\gamma(\mu_k) \leq \sum_{i=k}^{\infty} \gamma(\lambda_i) \leq 2^{-k+1}$. Also ist $\mu_k \in \Lambda$. Es gilt nun

$$\begin{aligned} & \sup_{t \geq 0} r(y_k(\mu_k^{-1}(t) \wedge u_k), y_{k+1}(\mu_{k+1}^{-1}(t) \wedge u_k)) \\ &= \sup_{t \geq 0} r(y_k(\mu_k^{-1}(t) \wedge u_k), y_{k+1}(\lambda_k(\mu_k^{-1}(t)) \wedge u_k)) \\ &= \sup_{t \geq 0} r(y_k(t \wedge u_k), y_{k+1}(\lambda_k(t) \wedge u_k)) \leq 2^{-k}. \end{aligned}$$

Als der Vollständigkeit von (S, r) folgt also, dass $y_k \circ \mu_k^{-1}$ gegen eine Funktion $y : [0, \infty) \rightarrow S$ konvergiert, und zwar gleichmässig auf beschränkten Intervallen. Dann muss aber $y \in D_S[0, \infty)$ gelten. Wir haben nun $\lim_{n \rightarrow \infty} \gamma(\mu_k^{-1}) = 0$ und $\lim_{k \rightarrow \infty} \sup_{0 \leq t \leq T} r(y_k(\mu_k^{-1}(t)), y(t)) = 0$ für alle $T > 0$. Also folgt aus Proposition 2.25, dass $\lim_{n \rightarrow \infty} y_n = y$. Da $\{x_k\}$ eine Cauchyfolge ist, folgt auch $\lim_{n \rightarrow \infty} x_n = y$. $\square$

2.6. Die kompakten Mengen von $D_S[0, \infty)$

Betrachten wir eine Elementarfunktion $x \in D_S[0, \infty)$ setzen wir $s_0(x) = 0$ und $s_k(x) = \inf\{t > s_{k-1}(x) : x(t) \neq x(t-)\}$.

Hilfssatz 2.27. Sei $\Gamma \subset S$ kompakt und $\delta > 0$. Wir definieren die Menge der Elementarfunktionen $A(\Gamma, \delta)$ mit der Eigenschaft, dass $x(t) \in \Gamma$ für alle $t \geq 0$ und $s_k(x) - s_{k-1}(x) > \delta$ falls $s_{k-1} \neq \infty$. Dann ist der Abschluss von $A(\Gamma, \delta)$ kompakt.

Beweis. Sei $\{x_n\} \subset A(\Gamma, \delta)$. Durch ein Diagonalisierungsargument gibt es eine Teilfolge $\{y_m\}$, so dass entweder $s_k(y_m) < \infty$ für jedes m und $\lim_{m \rightarrow \infty} s_k(y_m) = t_k \in [0, \infty]$ existiert und $\lim_{m \rightarrow \infty} y_m(s_k(y_m)) = \alpha_k$ existiert, oder $s_k(y_m) = \infty$ für alle m . Da $s_k(y_m) - s_{k-1}(y_m) > \delta$ folgt nun, dass y_m gegen eine Elementarfunktion $y \in D_S[0, \infty)$ konvergiert, mit $y(t) = \alpha_k$ für $t_k \leq t < t_{k+1}$. Also hat $\{x_n\}$ eine konvergente Teilfolge und somit ist die Aussage bewiesen. $\square$

Wir definieren für $x \in D_S[0, \infty)$, $\delta > 0$ und $T > 0$ die Grösse

$$w'(x, \delta, T) = \inf_{\{t_i\}} \max_i \sup_{r, s \in [t_{i-1}, t_i]} r(x(s), x(t)) ,$$

wobei $\{t_i\}$ über alle $0 = t_0 < t_1 < \dots < t_{n-1} < T \leq t_n$ mit $\min_{1 \leq i \leq n} (t_i - t_{i-1}) > \delta$ und $n \geq 1$. Man beachte, dass $w'(x, \delta, T)$ wachsend ist in δ und T . Man hat ferner

$$w'(x, \delta, T) \leq w'(y, \delta, T) + 2 \sup_{0 \leq s < T + \delta} r(x(s), y(s)) . \quad (2.9)$$

Hilfssatz 2.28.

- i) Für jedes $x \in D_S[0, \infty)$ und $T > 0$ ist die Funktion $w'(x, \delta, T)$ rechtsstetig in δ und $\lim_{\delta \downarrow 0} w'(x, \delta, T) = 0$.
- ii) Gilt für eine Folge $\{x_n\} \subset D_S[0, \infty)$ und eine Funktion $x \in D_S[0, \infty)$ der Grenzwert $\lim_{n \rightarrow \infty} d(x_n, x) = 0$, so haben wir

$$\overline{\lim}_{n \rightarrow \infty} w'(x_n, \delta, T) \leq w'(x, \delta, T + \varepsilon)$$

für alle $\delta > 0$, $T > 0$ und $\varepsilon > 0$.

Beweis. i) Zu jedem $\delta > 0$ und jeder Partition gibt es ein $\delta' > \delta$, so dass die Partition auch für δ' zulässig ist. Somit muss $w'(x, \delta, T)$ rechtsstetig in δ sein.

Sei $N \geq 1$ und $\{\tau_k^N\}$ wie in (2.8) definiert. Für $0 < \delta < \min\{\tau_{k+1}^N - \tau_k^N : \tau_k^N \leq T\}$ (existiert, da cadlag), gilt $w'(x, \delta, T) \leq 2N^{-1}$.

ii) Aus Proposition 2.25 folgt, dass es $\{\lambda_n\} \subset \Lambda'$ gibt, so dass (2.5) und (2.7) gelten, wenn man T durch $T + \delta$ ersetzt. Sei $y_n(t) = x(\lambda_n(t))$ und $\delta_n = \sup_{0 \leq t \leq T} (\lambda_n(t + \delta) - \lambda_n(t))$. Dann folgt aus (2.9) und i)

$$\begin{aligned} \overline{\lim}_{n \rightarrow \infty} w'(x_n, \delta, T) &= \overline{\lim}_{n \rightarrow \infty} w'(y_n, \delta, T) \leq \overline{\lim}_{n \rightarrow \infty} w'(x, \delta_n, \lambda_n(T)) \\ &\leq \lim_{n \rightarrow \infty} w'(x, \delta_n \vee \delta, T + \varepsilon) = w'(x, \delta, T + \varepsilon) . \end{aligned}$$

$\square$

Satz 2.29. Sei (S, d) vollständig. Dann ist der Abschluss von $A \subset D_S[0, \infty)$ genau dann kompakt, wenn die folgenden zwei Bedingungen erfüllt sind:

- i) Für jedes $t \in \mathbb{Q}_+$ gibt es eine kompakte Menge $\Gamma_t \in S$, so dass $x(t) \in \Gamma_t$ für alle $x \in A$.
- ii) Für jedes $T > 0$ gilt

$$\limsup_{\delta \rightarrow 0} \sup_{x \in A} w'(x, \delta, T) = 0 .$$

Beweis. Nehmen wir an, A erfüllt die beiden Bedingungen. Sei $\ell \geq 1$. Wir wählen $\delta_\ell \in (0, \frac{1}{2})$, so dass $\sup_{x \in A} w'(x, \delta_\ell, \ell) \leq \ell^{-1}$. Weiter wählen wir $m_\ell > \delta_\ell^{-1} \geq 2$. Wir definieren $\Gamma^{(\ell)} = \bigcup_{i=0}^{(\ell+1)m_\ell} \Gamma_{i/\ell_\ell}$ und $A_\ell = A(\Gamma^{(\ell)}, \delta_\ell)$, wobei wir die Notation aus Hilfssatz 2.27 verwenden. Für $x \in A$ gibt es eine Partition $0 = t_0 < t_1 < \dots < t_{n-1} < \ell \leq t_n < \ell + 1 < t_{n+1} = \infty$ mit $\min_{1 \leq i \leq n} (t_i - t_{i-1}) > \delta_\ell$, so dass $\max_{1 \leq i \leq n} \sup_{s, t \in [t_{i-1}, t_i]} r(x(s), x(t)) \leq 2\ell^{-1}$. Wir definieren $x' \in A_\ell$ durch $x'(t) = x(\lfloor m_\ell t + 1 \rfloor / m_\ell)$ für $t_i \leq t < t_{i+1}$. Dann ist $\sup_{0 \leq t < \ell} r(x'(t), x(t)) \leq 2\ell^{-1}$. Also

$$d(x', x) \leq \int_0^\infty e^{-u} \sup_{t \geq 0} (r(x'(t \wedge u), x(t \wedge u)) \wedge 1) du \leq 2\ell^{-1} + e^{-\ell} < 3\ell^{-1} .$$

Somit ist $A \subset A_\ell^{3/\ell}$. Wir wissen aus Hilfssatz 2.27, dass $\bar{A}_\ell$ kompakt ist. Da damit $A \subset \bigcap_{\ell \geq 1} A_\ell^{3/\ell}$ total beschränkt ist, hat A einen kompakten Abschluss.

Nehmen wir nun an, dass der Abschluss von A kompakt ist. Sei $\{x_n(t)\} \subset S$ eine Folge mit $x_n \in A$. Da A einen kompakten Abschluss hat, können wir annehmen, dass $\{x_n\}$ einen Grenzwert x hat. Es gibt nach Proposition 2.25 eine Folge $\{\lambda_n\} \subset \Lambda$, so dass $r(x_n(t), x(\lambda_n(t))) \rightarrow 0$. Wir wissen (siehe (2.5)), dass $\lambda_n(t) \rightarrow t$. Da aber der Grenzwert von $s \rightarrow t$ von $x(s)$ von beiden Seiten existiert, gibt es eine Teilfolge, so dass $x(\lambda_n(t))$ konvergiert. Somit gibt es eine konvergente Teilfolge von $\{x_n(t)\}$, das heisst, der Abschluss Γ_t von $\{y(t) : y \in A\}$ ist kompakt.

Nehmen wir an, dass es ein $\eta > 0$, $T > 0$ und eine Folge $\{x_n\} \subset A$ gibt, so dass $w'(x_n, n^{-1}, T) \geq \eta$ für alle n . Da A einen kompakten Abschluss hat, können wir annehmen, dass es ein $x \in D_S[0, \infty)$ gibt, so dass $d(x_n, x) \rightarrow 0$. Aus Hilfssatz 2.28 ii) schliessen wir, dass $\eta \leq \lim_{n \rightarrow \infty} w'(x, \delta, T) \leq w'(x, \delta, T+1)$ für jedes $\delta > 0$. Lassen wir nun aber $\delta \rightarrow 0$, ergibt sich aus Aus Hilfssatz 2.28 i) ein Widerspruch. $\square$

2.7. Konvergenz in Verteilung auf $D_S[0, \infty)$

Sei (E, r) ein metrischer Raum. Wir bezeichnen mit $\mathcal{S}_E$ die Borel- σ -Algebra auf $D_E[0, \infty)$ und mit $\mathcal{S}'_E$ die Borel σ -Algebra, die durch die Koordinatenabbildungen erzeugt wird.

Proposition 2.30. *Sei D eine dichte Teilmenge von $[0, \infty)$. Es gilt $\sigma(\pi_t : t \in D) = \mathcal{S}'_E \subset \mathcal{S}_E$. Ist E separabel, so ist $\mathcal{S}'_E = \mathcal{S}_E$.*

Beweis. Für $\varepsilon > 0$, $t \geq 0$ und $f \in \bar{C}(E)$ definiert $f_t^\varepsilon(x) = \varepsilon^{-1} \int_t^{t+\varepsilon} f(\pi_s(x)) \, ds$ eine stetige Funktion $f_t^\varepsilon(x)$ auf $D_E[0, \infty)$. Da $\lim_{\varepsilon \downarrow 0} f_t^\varepsilon(x) = f(\pi_t(x))$, ist $f \circ \pi_t$ Borel-messbar für jedes $f \in \bar{C}(E)$, und damit auch für jedes $f \in B(E)$. Also ist

$$\pi_t^{-1}(\Gamma) = \{x \in D_E[0, \infty) : \mathbb{1}_\Gamma(\pi_t(x)) = 1\} \in \mathcal{S}_E.$$

Da die Mengen $\pi_t^{-1}(\Gamma) \in \mathcal{S}'_E$ erzeugen, gilt $\mathcal{S}'_E \subset \mathcal{S}_E$. Für $t \geq 0$ sei $\{t_n\} \subset D \cap [t, \infty)$, so dass $\lim_{n \rightarrow \infty} t_n = t$. Dann ist $\pi_t = \lim_{n \rightarrow \infty} \pi_{t_n}$ wegen der Rechtsstetigkeit und somit $\sigma(\pi_s : s \in D)$ -messbar. Damit ist der erste Teil der Aussage gezeigt.

Sei nun E separabel. Sei $n \geq 1$ und $0 = t_0 < t_1 < \dots < t_n < t_{n+1} = \infty$. Für $\alpha_0, \alpha_1, \dots, \alpha_n \in E$ setzen wir $\eta(\alpha_0, \dots, \alpha_n)(t) = \alpha_i$ für $t_i \leq t < t_{i+1}$. Dann ist $\eta(\alpha_0, \dots, \alpha_n) \in D_E[0, \infty)$. Aus $d(\eta(\alpha_0, \dots, \alpha_n), \eta(\alpha'_0, \dots, \alpha'_n)) \leq \max_i r(\alpha_i, \alpha'_i)$, folgt, dass η eine stetige Abbildung von E^{n+1} nach $D_E[0, \infty)$ ist. Sei $z \in D_E[0, \infty)$. Die Abbildung $d(z, \cdot)$ ist stetig. Somit ist $d(z, \eta \circ (\pi_{t_0}, \dots, \pi_{t_n}))$ auch $\mathcal{S}'_E$ -messbar. Für $m \geq 1$ sei η_m die Funktion η mit $n = m^2$ und $t_i = i/m$. Dann folgt, dass $\lim_{m \rightarrow \infty} d(z, \eta_m(\pi_{t_0}(x), \dots, \pi_{t_{m^2}}(x))) = d(z, x)$. Also ist $d(z, x)$ auch $\mathcal{S}'_E$ -messbar. Das bedeutet, dass jeder offene Ball in $D_E[0, \infty)$ auch in $\mathcal{S}'_E$ ist. Nach Satz 2.26 ist $D_E[0, \infty)$ separabel. Also enthält $\mathcal{S}'_E$ alle offenen Mengen in $D_E[0, \infty)$, und somit auch $\mathcal{S}_E$. $\square$

Satz 2.31. *Sei (S, r) vollständig und separabel und sei $\{X_\alpha\}$ eine Familie von Prozessen mit Pfaden in $D_S[0, \infty)$. Dann ist $\{X_\alpha\}$ genau dann relativ kompakt, wenn die beiden folgenden Bedingungen erfüllt sind:*

- i) Für jedes $\eta > 0$ und $t \in \mathbb{Q}_+$ gibt es eine kompakte Menge $\Gamma_{\eta, t} \subset S$, so dass $\inf_\alpha \mathbb{P}[X_\alpha(t) \in \Gamma_{\eta, t}] \geq 1 - \eta$.
- ii) Für jedes $\eta > 0$ und $T > 0$ gibt es ein $\delta > 0$, so dass $\sup_\alpha \mathbb{P}[w'(X_\alpha, \delta, T) \geq \eta] \leq \eta$.

Beweis. Ist $\{X_\alpha\}$ relativ kompakt, so folgen die beiden Bedingungen aus den Sätzen 2.26, 2.14 und 2.29.

Nehmen wir nun die beiden Bedingungen an. Sei $0 < \varepsilon < 1$, $T \in \mathbb{N}$, so dass $e^{-T} < \varepsilon/2$, $\eta = \varepsilon/4$ und $\delta > 0$, so dass $\sup_\alpha \mathbb{P}[w'(X_\alpha, \delta, T) \geq \eta] \leq \eta$. Wir wählen $m > 1/\delta$ und setzen $\Gamma = \bigcup_{i=0}^{mT} \Gamma_{\varepsilon 2^{-i-2}, i/m}$. Aus $\mathbb{P}[X_\alpha(i/m) \notin \Gamma^{\varepsilon/4}] \leq \varepsilon 2^{-i-2}$ schliessen wir, dass

$$\inf_\alpha \mathbb{P}[X_\alpha(i/m) \in \Gamma^{\varepsilon/4}, i = 0, 1, \dots, mT] \geq 1 - \frac{\varepsilon}{2}.$$

Wir setzen $A = A(\Gamma, \delta)$ in der Notation von Hilfssatz 2.27, und erinnern uns, dass A einen kompakten Abschluss hat. Sei nun $x \in D_S[0, \infty)$ mit $w'(x, \delta, T) < \varepsilon/4$ und $x(i/m) \in \Gamma^{\varepsilon/4}$ für $i = 0, 1, \dots, mT$. Wir wählen $0 = t_0 < t_1 < t_{n-1} < T \leq t_n$, so dass $\min_{1 \leq i \leq n} (t_i - t_{i-1}) > \delta$ und $\max_{1 \leq i \leq n} \sup_{t, s \in [t_{i-1}, t_i]} r(x(s), x(t)) < \varepsilon/4$. Weiter wählen wir $\{y_i\} \subset \Gamma$, so dass $r(x(i/m), y_i) < \varepsilon/4$. Definieren wir nun $x'(t) = y_{\lfloor mt_{i-1} + 1 \rfloor}$ für $t_{i-1} \leq t < t_i$ und $i \leq n-1$ und $x'(t) = y_{\lfloor mt_{n-1} + 1 \rfloor}$ für $t \geq t_{n-1}$. Dann gilt $\sup_{0 \leq t < T} r(x(t), x'(t)) < \varepsilon/2$ und damit $d(x, x') < \varepsilon/2 + e^{-T} < \varepsilon$. Also ist $x \in A^\varepsilon$. Also haben wir gezeigt, dass $\inf_\alpha \mathbb{P}[X_\alpha \in A^\varepsilon] \geq 1 - \varepsilon$. Aus den Sätzen 2.26 und 2.14 folgt dann, dass $\{X_\alpha\}$ relativ kompakt ist. $\square$

Für einen Prozess X bezeichnen wir mit $D(X)$ die Menge der Punkte t , an denen $\mathbb{P}[X(t) = X(t-)] = 1$.

Hilfssatz 2.32. Sei X ein Prozess mit Pfaden in $D_S[0, \infty)$. Dann gibt es höchstens abzählbar viele Punkte in $[0, \infty)$, die nicht in $D(X)$ sind.

Beweis. Seien $\varepsilon, \delta, T > 0$. Nehmen wir an, dass es unendliche viele Punkte $\{t_n\}$ in $\{0 \leq t \leq T : \mathbb{P}[r(X(t), X(t-)) \geq \varepsilon] \geq \delta\}$ gibt, dann

$$\mathbb{P}\left[\bigcap_{n \geq 1} \bigcup_{m \geq n} \{r(X(t_m), X(t_m-)) \geq \varepsilon\}\right] = \lim_{n \rightarrow \infty} \mathbb{P}\left[\bigcup_{m \geq n} \{r(X(t_m), X(t_m-)) \geq \varepsilon\}\right] \geq \delta.$$

Dann wäre aber $\mathbb{P}[X \in D_S[0, \infty)] < 1$. Also ist $\{t \in (0, \infty) : \mathbb{P}[r(X(t), X(t-)) \geq \varepsilon] \geq \delta\}$ und damit auch $\{t \in (0, \infty) : \mathbb{P}[r(X(t), X(t-)) \geq \varepsilon] > 0\}$ abzählbar. Lassen wir $\varepsilon \rightarrow 0$, folgt die Behauptung. $\square$

Satz 2.33. Sei S separabel und X_n und X seien Prozesse mit Pfaden in $D_S[0, \infty)$.

i) Gilt $X_n \Rightarrow X$, dann gilt

$$(X_n(t_1), \dots, X_n(t_k)) \Rightarrow (X(t_1), \dots, X(t_k)) \quad (2.10)$$

für jede endliche Menge $\{t_1, \dots, t_k\} \subset D(X)$. Weiter, für jede endliche Menge $\{t_1, \dots, t_k\} \subset [0, \infty)$ gibt es Folgen $\{t_i^n\} \subset [t_i, \infty)$, so dass $t_i^n \rightarrow t_i$, so dass $(X_n(t_1^n), \dots, X_n(t_k^n)) \Rightarrow (X(t_1), \dots, X(t_k))$.

ii) Ist $\{X_n\}$ relativ kompakt, und gibt es eine dichte Menge D , so dass (2.10) für jede endliche Menge $\{t_1, \dots, t_k\} \subset D$ gilt, dann gilt $X_n \Rightarrow X$.

Beweis. i) Nehmen wir $X_n \Rightarrow X$ an. Dann ist $D_S[0, \infty)$ separabel (Satz 2.26). Nach Satz 2.10 gibt es einen Wahrscheinlichkeitsraum, auf dem Y_n und Y mit Pfaden in $D_S[0, \infty)$ definiert sind mit den gleichen Verteilungen wie X_n und X , so dass $Y_n \rightarrow Y$. Ist $t \in D(X) = D(Y)$, so ist π_t stetig bezüglich der Verteilung von Y . Korollar 2.11 impliziert dann, dass $Y_n(t) \rightarrow Y(t)$. Somit folgt die erste Behauptung. Die zweite Behauptung folgt aus Hilfssatz 2.32.

ii) Wenn wir eine konvergente Teilfolge betrachten, können wir $X_n \Rightarrow Y$ für einen Prozess Y mit Pfaden in $D_S[0, \infty)$ annehmen. Seien $t_1, \dots, t_k \in D(Y)$ und $f_1, \dots, f_k \in \bar{C}(S)$. Weiter wählen wir $\{t_i^m\} \subset D \cap [t_i, \infty)$, so dass $t_i^m \rightarrow t_i$. Es gibt $n_1 < n_2 < \dots$, so dass $|\mathbb{E}[\prod_{i=1}^k f_i(X(t_i^m))] - \mathbb{E}[\prod_{i=1}^k f_i(X_{n_m}(t_i^m))]| < m^{-1}$. Dann gilt

$$\begin{aligned} \left| \mathbb{E} \left[\prod_{i=1}^k f_i(X(t_i)) \right] - \mathbb{E} \left[\prod_{i=1}^k f_i(Y(t_i)) \right] \right| &\leq \left| \mathbb{E} \left[\prod_{i=1}^k f_i(X(t_i)) \right] - \mathbb{E} \left[\prod_{i=1}^k f_i(X(t_i^m)) \right] \right| \\ &\quad + \left| \mathbb{E} \left[\prod_{i=1}^k f_i(X(t_i^m)) \right] - \mathbb{E} \left[\prod_{i=1}^k f_i(X_{n_m}(t_i^m)) \right] \right| \\ &\quad + \left| \mathbb{E} \left[\prod_{i=1}^k f_i(X_{n_m}(t_i^m)) \right] - \mathbb{E} \left[\prod_{i=1}^k f_i(Y(t_i)) \right] \right| \end{aligned}$$

für jedes $m \geq 1$. Alle drei Terme auf der rechten Seite konvergieren nach 0. Also gilt $\mathbb{E}[\prod_{i=1}^k f_i(X(t_i))] = \mathbb{E}[\prod_{i=1}^k f_i(Y(t_i))]$. Wegen der Rechtsstetigkeit gilt dies für alle $\{t_1, \dots, t_k\} \in [0, \infty)$. Die Mengen A , für die $\mathbb{P}[X \in A] = \mathbb{P}[Y \in A]$ bilden eine Dynkin-Klasse und enthält somit die endlich-dimensionalen Verteilungen. Also haben X und Y die gleiche Verteilung. Da Y beliebig war, muss jede konvergente Teilfolge den Grenzwert X haben. Also haben wir die Aussage bewiesen. $\square$

2.8. Kriterien für relative Kompaktheit in $D_S[0, \infty)$

Sei $x \in D_S[0, \infty)$. Für $\varepsilon > 0$ definieren wir $\tau_0 = \sigma_0 = 0$, und

$$\tau_k = \inf\{t > \tau_{k-1} : r(x(t), x(\tau_{k-1})) > \varepsilon/2\}$$

und

$$\sigma_k = \sup\{t \leq \tau_k : r(x(t), x(\tau_k)) \vee r(x(t-), x(\tau_k)) \geq \varepsilon/2\}$$

falls $\tau_k < \infty$ und $\sigma_k = \infty$, falls $\tau_k = \infty$. Wir bemerken, dass $w'(x, \delta, T) < \varepsilon/2$ impliziert, dass $\min\{\tau_{k+1} - \sigma_k : \tau_k < T\} > \delta$. Wir setzen $s_k = (\sigma_k + \tau_k)/2$. Wir bemerken, dass $\sigma_k \leq s_k \leq \tau_k \leq \sigma_{k+1}$, und damit $2s_{k+1} = \sigma_{k+1} + \tau_{k+1} \geq \tau_k + \tau_{k+1}$. Also gilt auch $2(s_{k+1} - s_k) \geq (\tau_k + \tau_{k+1}) - (\sigma_k + \tau_k) = \tau_{k+1} - \sigma_k$.

Hilfssatz 2.34. Sei (S, r) separabel und $\{X_\alpha\}$ eine Familie von Prozessen mit Pfaden in $D_S[0, \infty)$. Seien die Variablen $\tau_k^{\alpha, \varepsilon}$, $\sigma_k^{\alpha, \varepsilon}$ und $s_k^{\alpha, \varepsilon}$ wie oben für $x = X_\alpha$ definiert. Dann sind folgende Aussagen äquivalent:

- i) $\liminf_{\delta \downarrow 0} \mathbb{P}_\alpha[w'(X_\alpha, \delta, T) < \varepsilon] = 1, \quad \varepsilon > 0, T > 0.$
- ii) $\liminf_{\delta \downarrow 0} \mathbb{P}_\alpha[\min\{s_{k+1}^{\alpha, \varepsilon} - s_k^{\alpha, \varepsilon} : s_k^{\alpha, \varepsilon} < T\} \geq \delta] = 1, \quad \varepsilon > 0, T > 0.$
- iii) $\liminf_{\delta \downarrow 0} \mathbb{P}_\alpha[\min\{\tau_{k+1}^{\alpha, \varepsilon} - \sigma_k^{\alpha, \varepsilon} : \tau_k^{\alpha, \varepsilon} < T\} \geq \delta] = 1, \quad \varepsilon > 0, T > 0.$

Beweis. “i) $\Rightarrow$ iii)” Sei $w'(X_\alpha, \delta, T) < \varepsilon/2$. Dann muss $\tau_{k+1}^{\alpha, \varepsilon} - \tau_k^{\alpha, \varepsilon} > \delta$ sein, falls $\tau_k^{\alpha, \varepsilon} < T$. Also folgt iii), da $\sigma_k^{\alpha, \varepsilon} \leq \tau_k^{\alpha, \varepsilon}$.

“iii) $\Rightarrow$ ii)” Nehmen wir an, dass $\min\{\tau_{k+1}^{\alpha, \varepsilon} - \sigma_k^{\alpha, \varepsilon} : \tau_k^{\alpha, \varepsilon} < T + \delta\} > 2\delta$. Dann haben wir oben gezeigt, dass $s_{k+1}^{\alpha, \varepsilon} - s_k^{\alpha, \varepsilon} > \delta$ für alle k , so dass $\tau_k^{\alpha, \varepsilon} < T + \delta$. Ist $s_k^{\alpha, \varepsilon} < T$ und $\tau_k^{\alpha, \varepsilon} \geq T + \delta$, so ist $s_{k+1}^{\alpha, \varepsilon} - s_k^{\alpha, \varepsilon} > \tau_k^{\alpha, \varepsilon} - T \geq \delta$. Somit gilt $\min\{s_{k+1}^{\alpha, \varepsilon} - s_k^{\alpha, \varepsilon} : s_k^{\alpha, \varepsilon} < T\} > \delta$ und damit ii).

“ii) $\Rightarrow$ i)” Nehmen wir $\min\{s_{k+1}^{\alpha, \varepsilon} - s_k^{\alpha, \varepsilon} : s_k^{\alpha, \varepsilon} < T\} \geq 2\delta > \delta$ an. Sei $s_k^{\alpha, \varepsilon} \leq t < s < s_{k+1}^{\alpha, \varepsilon}$. Ist $t \leq \tau_k^{\alpha, \varepsilon}$, so ist $r(X_\alpha(t), X_\alpha(\tau_k^{\alpha, \varepsilon})) \leq \varepsilon/2$. Ist $t > \tau_k^{\alpha, \varepsilon}$, so ist $t < \tau_{k+1}^{\alpha, \varepsilon}$, also auch $r(X_\alpha(t), X_\alpha(\tau_k^{\alpha, \varepsilon})) \leq \varepsilon/2$. Aus der Dreiecksungleichung erhalten wir $r(X_\alpha(t), X_\alpha(s)) \leq \varepsilon$. Somit gilt i). $\square$

Hilfssatz 2.35. Für jedes α sei $0 = s_0^\alpha < s_1^\alpha < \dots$ eine Folge von Zufallsvariablen mit $\lim_{n \rightarrow \infty} s_n^\alpha = \infty$. Sei $\Delta_k^\alpha = s_{k+1}^\alpha - s_k^\alpha$, $T > 0$ und $K_\alpha(T) = \max\{k \geq 0 : s_k^\alpha < T\}$. Wir setzen $F(t) = \sup_\alpha \sup_k \mathbb{P}[\Delta_k^\alpha < t, s_k^\alpha < T]$. Dann gilt für alle $\delta > 0$ und $L \in \mathbb{N} \setminus \{0\}$

$$F(\delta) \leq \sup_\alpha \mathbb{P}\left[\min_{0 \leq k \leq K_\alpha(T)} \Delta_k^\alpha < \delta\right] \leq LF(\delta) + e^T \int_0^\infty Le^{-Lt} F(t) dt.$$

Somit erhält man

$$\limsup_{\delta \downarrow 0} \mathbb{P}_\alpha\left[\min_{0 \leq k \leq K_\alpha(T)} \Delta_k^\alpha < \delta\right] = 0$$

genau dann, wenn $F(0+) = 0$.

Beweis. Die erste Ungleichung ist trivial. Wir erhalten

$$\begin{aligned}
\mathbb{P}\left[\min_{0 \leq k \leq K_\alpha(T)} \Delta_k^\alpha < \delta\right] &\leq \sum_{k=0}^{L-1} \mathbb{P}[\Delta_k^\alpha < \delta, s_k^\alpha < T] + \mathbb{P}[K_\alpha(T) \geq L] \\
&\leq LF(\delta) + e^T \mathbb{E}\left[\mathbb{1}_{K_\alpha(T) \geq L} \exp\left\{-\sum_{k=0}^{L-1} \Delta_k^\alpha\right\}\right] \\
&\leq LF(\delta) + e^T \prod_{k=0}^{L-1} (\mathbb{E}[\mathbb{1}_{s_k^\alpha < T} e^{-L\Delta_k^\alpha}])^{1/L} \\
&\leq LF(\delta) + e^T \prod_{k=0}^{L-1} \left(\int_0^\infty L e^{-Lt} \mathbb{P}[s_k^\alpha < T, \Delta_k^\alpha < t] dt\right)^{1/L} \\
&\leq LF(\delta) + e^T \int_0^\infty L e^{-Lt} F(t) dt.
\end{aligned}$$

Die zweite Aussage erhält man, wenn man zuerst $\delta \downarrow 0$ gehen lässt, und danach $L \rightarrow \infty$. $\square$

Proposition 2.36. *Unter den Annahmen von Hilfssatz 2.34 sind die Aussagen des Hilfssatzes auch äquivalent zu*

$$\text{iv)} \quad \lim_{\delta \downarrow 0} \sup_{\alpha} \sup_{k \geq 0} \mathbb{P}[\tau_{k+1}^{\alpha, \varepsilon} - \sigma_k^{\alpha, \varepsilon} < \delta, \tau_k^{\alpha, \varepsilon} < T] = 0, \quad \varepsilon > 0, T > 0.$$

Beweis. Wir bemerken zuerst, dass

$$\begin{aligned}
\mathbb{P}[s_{k+1}^{\alpha, \varepsilon} - s_k^{\alpha, \varepsilon} < \delta/2, s_k^{\alpha, \varepsilon} < T] &\leq \mathbb{P}[\tau_{k+1}^{\alpha, \varepsilon} - \sigma_k^{\alpha, \varepsilon} < \delta, \tau_k^{\alpha, \varepsilon} < T + \delta] \\
&\leq \mathbb{P}[s_{k+1}^{\alpha, \varepsilon} - s_k^{\alpha, \varepsilon} < \delta, s_k^{\alpha, \varepsilon} < T + \delta].
\end{aligned}$$

Wir bemerken weiter, dass $s_k^{\alpha, \varepsilon} \rightarrow \infty$. Gelten nun die Aussagen aus Hilfssatz 2.34, so konvergiert nach Hilfssatz 2.35 $\mathbb{P}[s_{k+1}^{\alpha, \varepsilon} - s_k^{\alpha, \varepsilon} < \delta, s_k^{\alpha, \varepsilon} < T + \delta]$ nach 0. Somit folgt aus den obigen Betrachtungen die Behauptung. Gilt iv), so folgt aus den obigen Betrachtungen und aus Hilfssatz 2.35 die Aussage ii). $\square$

Wir bezeichnen mit $S(T)$ die Klasse der $\{\mathcal{F}_{t+}^X\}$ -Stoppszeiten, die durch T begrenzt sind.

Hilfssatz 2.37. *Sei (S, r) separabel und X ein Prozess mit Pfaden in $D_S[0, \infty)$, $T > 0$ und $\beta > 0$. Dann gilt für jedes $\delta > 0$, $0 < \lambda < 1$ und $\tau \in S(T)$*

$$\begin{aligned}
\mathbb{P}\left[\sup_{0 \leq u \leq \delta} r(X(\tau + u), X(\tau)) \geq \lambda, \sup_{0 \leq v \leq \delta \wedge \tau} r(X(\tau), X(\tau - v)) \geq \lambda\right] \\
\leq \lambda^{-2\beta} a_\beta [1 + 2a_\beta^2(1 + 4a_\beta)] C(\delta) \quad (2.11)
\end{aligned}$$

und

$$\mathbb{P} \left[\sup_{0 \leq u \leq \delta} r(X(u), X(0)) \geq \lambda \right] \leq \lambda^{-2\beta} a_\beta \{a_\beta(1 + 4a_\beta)C(\delta) + \mathbb{E}[r^\beta(X(\delta), X(0)) \wedge 1]\} ,$$

wobei

$$C(\delta) = \sup_{\tau \in S(T+2\delta)} \sup_{0 \leq u \leq 2\delta} \mathbb{E} \left[\sup_{0 \leq v \leq 3\delta \wedge \tau} r^\beta(X(\tau+u), X(\tau)) r^\beta(X(\tau), X(\tau-v)) \wedge 1 \right]$$

und $a_\beta = 2^{(\beta-1)^+}$.

Beweis. Wir bemerken zuerst, dass für $c, d \geq 0$ die Ungleichung $(c+d)^\beta \leq a_\beta(c^\beta + d^\beta)$ gilt. Wir dürfen annehmen, dass $r \leq 1$.

Sei $\tau \in S(T+\delta)$ und bezeichnen wir mit $M_\tau(\delta)$ die Klasse der $\{\mathcal{F}_{\tau+}^X\}$ -messbaren Zufallsvariablen mit $0 \leq U \leq \delta$. Sei $U \in M_\tau(\delta)$. Dann gilt

$$\begin{aligned} r^\beta(X(\tau+U), X(\tau)) &\leq \frac{a_\beta}{\delta} \int_\delta^{2\delta} [r^\beta(X(\tau+\theta), X(\tau)) + r^\beta(X(\tau+\theta), X(\tau+U))] d\theta \\ &\leq \frac{a_\beta}{\delta} \left[\int_\delta^{2\delta} r^\beta(X(\tau+\theta), X(\tau)) d\theta + \int_0^{2\delta} r^\beta(X(\tau+U+\theta), X(\tau+U)) d\theta \right]. \end{aligned}$$

Daraus folgern wir mittels der Dreiecksungleichung, dass

$$\begin{aligned} &\sup_{0 \leq v < 2\delta \wedge \tau} r^\beta(X(\tau+U), X(\tau)) r^\beta(X(\tau), X(\tau-v)) \\ &\leq \frac{a_\beta}{\delta} \left[\int_\delta^{2\delta} \sup_{0 \leq v < 2\delta \wedge \tau} r^\beta(X(\tau+\theta), X(\tau)) r^\beta(X(\tau), X(\tau-v)) d\theta \right. \\ &\quad + a_\beta \int_0^{2\delta} \sup_{0 \leq v < 2\delta \wedge \tau} r^\beta(X(\tau+U+\theta), X(\tau+U)) r^\beta(X(\tau+U), X(\tau-v)) d\theta \\ &\quad \left. + a_\beta \int_0^{2\delta} r^\beta(X(\tau+U+\theta), X(\tau+U)) r^\beta(X(\tau+U), X(\tau)) d\theta \right] \\ &\leq \frac{a_\beta}{\delta} \left[\int_\delta^{2\delta} \sup_{0 \leq v < 2\delta \wedge \tau} r^\beta(X(\tau+\theta), X(\tau)) r^\beta(X(\tau), X(\tau-v)) d\theta \right. \\ &\quad + 2a_\beta \int_0^{2\delta} \sup_{0 \leq v < 3\delta \wedge (\tau+U)} r^\beta(X(\tau+U+\theta), X(\tau+U)) \\ &\quad \left. \times r^\beta(X(\tau+U), X(\tau+U-v)) d\theta \right]. \end{aligned}$$

Da $\tau+U \in S(T+2\delta)$, folgern wir

$$\begin{aligned} &\sup_{\tau \in S(T+\delta)} \sup_{U \in M_\tau(\delta)} \mathbb{E} \left[\sup_{0 \leq v < 2\delta \wedge \tau} r^\beta(X(\tau+U), X(\tau)) r^\beta(X(\tau), X(\tau-v)) \right] \\ &\leq a_\beta(1 + 4a_\beta)C(\delta). \end{aligned}$$

Wählen wir $0 < \eta < \lambda$ und $\tau \in S(T)$ und definieren $\Delta = \inf\{t > 0 : r(X(\tau + t), X(\tau)) > \lambda - \eta\}$. Für $0 \leq v \leq \delta \wedge \tau$ haben wir

$$\begin{aligned} r^\beta(X(\tau + \Delta \wedge \delta), X(\tau)) r^\beta(X(\tau), X(\tau - v)) \\ \leq a_\beta r^\beta(X(\tau + \delta), X(\tau)) r^\beta(X(\tau), X(\tau - v)) \\ + a_\beta^2 r^\beta(X(\tau + \delta), X(\tau + \Delta \wedge \delta)) r^\beta(X(\tau + \Delta \wedge \delta), X(\tau)) \\ + a_\beta^2 r^\beta(X(\tau + \delta), X(\tau + \Delta \wedge \delta)) r^\beta(X(\tau + \Delta \wedge \delta), X(\tau - v)) . \end{aligned}$$

Da $\tau + \Delta \wedge \delta \in S(T + \delta)$, $\delta - \Delta \wedge \delta \in M_{\tau + \Delta \wedge \delta}(\delta)$ und $\Delta \wedge \delta + v \leq 2\delta$, erhalten wir

$$\mathbb{E} \left[\sup_{0 \leq v \leq \delta \wedge \tau} r^\beta(X(\tau + \Delta \wedge \delta), X(\tau)) r^\beta(X(\tau), X(\tau - v)) \right] \leq [a_\beta + 2a_\beta^3(1 + 4a_\beta)]C(\delta) .$$

Die linke Seite von (2.11) ist durch $(\lambda - \eta)^{-\beta} \lambda^{-\beta}$ mal die linke Seite der obigen Ungleichung beschränkt. Lassen wir $\eta \downarrow 0$, folgt (2.11).

Wählen wir nun $\tau = 0$. Dann gilt

$$\begin{aligned} r^{2\beta}(X(\Delta \wedge \delta), X(0)) &\leq a_\beta [r^\beta(X(\delta), X(\Delta \wedge \delta)) r^\beta(X(\Delta \wedge \delta), X(0)) \\ &\quad + r^\beta(X(\delta), X(0)) r^\beta(X(\Delta \wedge \delta), X(0))] \\ &\leq a_\beta [r^\beta(X(\delta), X(\Delta \wedge \delta)) r^\beta(X(\Delta \wedge \delta), X(0)) + r^\beta(X(\delta), X(0))] . \end{aligned}$$

Die Aussage folgt nun analog. □

Satz 2.38. Sei (S, r) vollständig und separabel und $\{X_\alpha\}$ eine Familie von Prozessen mit Pfaden in $D_S[0, \infty)$. Es gelte die Bedingung i) aus Satz 2.31. Dann sind die folgenden Aussagen äquivalent:

- i) $\{X_\alpha\}$ ist relativ kompakt.
- ii) Für jedes $T > 0$ gibt es ein $\beta > 0$ und eine Familie $\{\gamma_\alpha(\delta) : 0 < \delta < 1\}$ von positiven Zufallsvariablen, so dass

$$\mathbb{E}[r^\beta(X_\alpha(t+u), X_\alpha(t)) \wedge 1 \mid \mathcal{F}_t^\alpha] (r^\beta(X_\alpha(t), X_\alpha(t-v)) \wedge 1) \leq \mathbb{E}[\gamma_\alpha(\delta) \mid \mathcal{F}_t^\alpha] \quad (2.12)$$

für $0 \leq t \leq T$, $0 \leq u \leq \delta$ und $0 \leq v \leq \delta \wedge t$, wobei $\{\mathcal{F}_t^\alpha\}$ die natürliche Filtration von X_α ist. Und weiter gilt $\lim_{\delta \downarrow 0} \sup_\alpha \mathbb{E}[\gamma_\alpha(\delta)] = 0$ und

$$\lim_{\delta \downarrow 0} \sup_\alpha \mathbb{E}[r^\beta(X_\alpha(\delta), X_\alpha(0)) \wedge 1] = 0 . \quad (2.13)$$

iii) Für jedes $T > 0$ gibt es ein $\beta > 0$, so dass (2.13) gilt und für die Grössen

$$C_\alpha(\delta) = \sup_{\tau \in S_0^\alpha(T)} \sup_{0 \leq u \leq \delta} \mathbb{E} \left[\sup_{0 \leq v \leq \delta \wedge \tau} (r^\beta(X_\alpha(\tau + u), X_\alpha(\tau)) \wedge 1) \right. \\ \left. \times (r^\beta(X_\alpha(\tau), X_\alpha(\tau - v)) \wedge 1) \right]$$

der Grenzwert $\lim_{\delta \downarrow 0} \sup_\alpha C_\alpha(\delta) = 0$ gilt. $S_0^\alpha(T)$ bezeichnet die Klasse der diskreten $\{\mathcal{F}_t^\alpha\}$ -Stoppzeiten, die durch T beschränkt sind.

Beweis. Wir können annehmen, dass $r \leq 1$.

“i) $\Rightarrow$ ii)” Wir haben

$$r(X_\alpha(t + u), X_\alpha(t)) r(X_\alpha(t), X_\alpha(t - v)) \leq r(X_\alpha(t + u), X_\alpha(t)) \wedge r(X_\alpha(t), X_\alpha(t - v)) \\ \leq w'(X_\alpha, 2\delta, T + \delta)$$

und $r(X_\alpha(\delta), X_\alpha(0)) \leq w'(X_\alpha, \delta, T)$. Wählen wir $\gamma_\alpha(\delta) = w'(X_\alpha, 2\delta, T + \delta)$, so folgt die Behauptung aus Satz 2.31.

“ii) $\Rightarrow$ iii)” Sei $\tau \in S_0^\alpha(T)$ fest. Multiplizieren wir in (2.12) beide Seiten mit $\mathbb{1}_{\tau=z}$ und addieren über alle Werte, die τ annimmt, sehen wir, dass wir t durch τ ersetzen können. Weiter können wir die rechte Seite durch das Supremum über $v \in [0, \delta \wedge t] \cap \mathbb{Q}$, und damit über $v \in [0, \delta \wedge t]$ ersetzen. Somit gilt iii).

“iii) $\Rightarrow$ i)” Aus der Konstruktion von $\{\tau_k^{\alpha, \varepsilon}\}$ und $\{\sigma_k^{\alpha, \varepsilon}\}$ und der Definition von $C_\alpha(\delta)$ folgt, dass die Bedingung iv) aus Proposition 2.36 erfüllt ist. Somit gilt die Bedingung i) aus Hilfssatz 2.34. Aus Satz 2.31 folgt dann i). $\square$

2.9. Weitere Kriterien für relative Kompaktheit in $D_S[0, \infty)$

Satz 2.39. Sei (S, d) vollständig und separabel. Sei $\{X_\alpha\}$ eine Familie von Prozessen mit Pfaden in $D_S[0, \infty)$. Nehmen wir an, dass es für jedes $\eta > 0$ und $T > 0$ eine kompakte Menge $\Gamma_{\eta, T} \subset S$ gibt, so dass

$$\inf_\alpha \mathbb{P}[X_\alpha(t) \in \Gamma_{\eta, T} \text{ für } 0 \leq t \leq T] \geq 1 - \eta \quad (2.14)$$

gilt. Sei H eine Teilmenge von $\bar{C}(S)$, die dicht ist in $\bar{C}(S)$ in der Topologie der gleichmässigen Konvergenz auf kompakten Mengen. Dann ist $\{X_\alpha\}$ relativ kompakt genau dann, wenn $\{f \circ X_\alpha\}$ relativ kompakt ist für jedes $f \in H$.

Beweis. Die Abbildung $D_S[0, \infty) \rightarrow D_{\mathbb{R}}[0, \infty)$, $x \mapsto f \circ x$ ist stetig (Übung). Somit folgt aus der relativen Kompaktheit von $\{X_\alpha\}$, dass auch $\{f \circ X_\alpha\}$ relativ kompakt ist.

Sei nun $\{f \circ X_\alpha\}$ relativ kompakt für jedes $f \in H$. Wir können annehmen, dass $r \leq 1$. Aus der Voraussetzung, (2.9) und Satz 2.31 folgt, dass $\{f \circ X_\alpha\}$ relativ kompakt ist für jedes $f \in \bar{C}(S)$. Insbesondere ist für $z \in S$ die Menge der Abbildungen $\{t \mapsto r(X_\alpha(t), z)\}$ relativ kompakt. Sei $\eta > 0$ und $T > 0$. Für jedes $\varepsilon > 0$ gibt es eine endliche Menge $\{z_1, \dots, z_N\} \subset \Gamma_{\eta, T}$, so dass $\min_{1 \leq i \leq N} r(x, z_i) < \varepsilon$ für alle $x \in \Gamma_{\eta, T}$. Ist $y \in \Gamma_{\eta, T}$, so gibt es ein z_i , so dass $r(y, z_i) < \varepsilon$. Also haben wir

$$r(x, y) \leq r(x, z_i) + r(y, z_i) < r(x, z_i) - r(y, z_i) + 2\varepsilon \leq |r(x, z_i) - r(y, z_i)| + 2\varepsilon.$$

Somit gilt für $0 \leq t \leq T$, $\delta \in (0, 1)$, $0 \leq u \leq \delta$ und $0 \leq v \leq \delta \wedge t$

$$\begin{aligned} & r(X_\alpha(t+u), X_\alpha(t))r(X_\alpha(t), X_\alpha(t-v)) \\ & \leq \bigvee_{i=1}^N |r(X_\alpha(t+u), z_i) - r(X_\alpha(t), z_i)| |r(X_\alpha(t), z_i) - r(X_\alpha(t-v), z_i)| \\ & \quad + 4(\varepsilon + \varepsilon^2) + \mathbb{1}_{X_\alpha(s) \notin \Gamma_{\eta, T} \text{ für ein } s \in [0, T]} \\ & \leq \bigvee_{i=1}^N w'(q(\cdot, z_i) \circ X_\alpha, 2\delta, T + \delta) + 4(\varepsilon + \varepsilon^2) + \mathbb{1}_{X_\alpha(s) \notin \Gamma_{\eta, T} \text{ für ein } s \in [0, T]} \\ & \stackrel{\text{def}}{=} \gamma_\alpha(\delta). \end{aligned}$$

Nach Satz 2.31 gilt $\lim_{\delta \downarrow 0} \sup_\alpha \mathbb{E}[w'(q(\cdot, z) \circ X_\alpha, 2\delta, T + \delta)] = 0$ für jedes $z \in S$. Also können wir Funktionen $\eta(\delta)$ und $\varepsilon(\delta)$ wählen, so dass $\lim_{\delta \downarrow 0} \sup_\alpha \mathbb{E}[\gamma_\alpha(\delta)] = 0$. Analog erhalten wir

$$r(X_\alpha(\delta), X_\alpha(0)) \leq \bigvee_{i=1}^N |r(X_\alpha(\delta), z_i) - r(X_\alpha(0), z_i)| + 2\varepsilon + \mathbb{1}_{X_\alpha(0) \notin \Gamma_{\eta, T}}$$

für alle $\delta > 0$. Somit konvergiert auch $\lim_{\delta \downarrow 0} \sup_\alpha \mathbb{E}[r(X_\alpha(\delta), X_\alpha(0))] = 0$. Nach Satz 2.38 ist $\{X_\alpha\}$ relativ kompakt. $\square$

Korollar 2.40. Sei (S, r) vollständig und separabel und $\{X_n\}$ eine Folge von Prozessen mit Pfaden in $D_S[0, \infty)$. Sei $M \subset \bar{C}(S)$ eine punkttrennende Menge. Falls es einen Prozess X mit Pfaden in $D_S[0, \infty)$ gibt, so dass für jede endliche Menge $\{g_1, g_2, \dots, g_n\} \subset M$ gilt, $(g_1, \dots, g_n) \circ X_n \Rightarrow (g_1, \dots, g_n) \circ X$. Dann gilt $X_n \Rightarrow X$.

Beweis. Sei H die kleinste Algebra, die $M \cup \{1\}$ enthält. Nach dem Satz von Stone-Weierstrass ist H dicht in $\bar{C}(E)$ bezüglich der Topologie der gleichmässigen Konvergenz auf kompakten Mengen. Dann gilt $(g_1, \dots, g_n) \circ X_n \Rightarrow (g_1, \dots, g_n) \circ X$ für alle $\{g_1, g_2, \dots, g_n\} \subset H$. Wir müssen somit (2.14) zeigen.

Sei $\Gamma \subset S$ kompakt und $\delta > 0$. Für $x \in S$ wählen wir $\{h_1^x, \dots, h_{k(x)}^x\} \subset H$, so dass $\varepsilon(x) \stackrel{\text{def}}{=} \inf_{y: r(x,y) \geq \delta} \max_{1 \leq i \leq k(x)} |h_i^x(y) - h_i^x(x)| > 0$. Sei $U_x = \{y \in S : \max_{1 \leq i \leq k(x)} |h_i^x(y) - h_i^x(x)| < \varepsilon(x)\}$. Dann ist $\Gamma \subset \bigcup_{x \in \Gamma} U_x \subset \Gamma^\delta$. Da die Funktionen stetig sind ist U_x offen. Also gibt es eine endliche Menge $\{x_1, \dots, x_N\} \subset \Gamma$, so dass $\Gamma \subset \bigcup_{\ell=1}^N U_{x_\ell} \subset \Gamma^\delta$. Definieren wir die Abbildung $\sigma : D_{\mathbb{R}}[0, \infty) \rightarrow D_{\mathbb{R}}[0, \infty)$, $x \mapsto \sigma(x)(t) = \sup_{0 \leq s \leq t} x(s)$. Die Abbildung ist stetig (Proposition 2.25). Sei $g_\ell(x) = \max_{1 \leq i \leq k(x_\ell)} |h_i^{x_\ell}(x) - h_i^{x_\ell}(x_\ell)|$, $Y_n(t) = \min_{1 \leq \ell \leq N} \{g_\ell(X_n(t)) - \varepsilon(x_\ell)\}$ und $Z_n = \sigma(Y_n)$. Aus der Voraussetzung und der Stetigkeit von σ folgt, dass $Z_n \Rightarrow Z$, wobei Z analog zu Z_n aus X konstruiert wird. Für ein $T > 0$, so dass $Z(t)$ stetig in T ist, gilt dann

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{P}[X_n(t) \in \Gamma^\delta \text{ für } 0 \leq t \leq T] &\geq \lim_{n \rightarrow \infty} \mathbb{P}[Z_n(T) < 0] \geq \mathbb{P}[Z(T) < 0] \\ &\geq \mathbb{P}[X(t) \in \Gamma \text{ für } 0 \leq t \leq T], \end{aligned}$$

wobei wir verwendet haben, dass $\sup_{x \in \Gamma} \min_{1 \leq i \leq N} \{g_i(x) - \varepsilon(x_i)\} < 0$. Sei $\eta > 0$, $T > 0$ ein Stetigkeitspunkt von Z und $m \geq 1$. Nach Hilfssatz 2.13 und Satz 2.26 gibt es eine kompakte Menge $\Gamma_{0,m} \subset S$, so dass $\mathbb{P}[X(t) \in \Gamma_{0,m} \text{ für } 0 \leq t \leq T] \geq 1 - \eta 2^{-m-1}$. Dann gibt es ein $n_m \geq 1$, so dass $\inf_{n \geq n_m} \mathbb{P}[X_n(t) \in \Gamma_{0,m}^{1/m} \text{ für } 0 \leq t \leq T] \geq 1 - \eta 2^{-m}$. Für $n = 1, \dots, n_m - 1$ können wir eine kompakte Menge $\Gamma_{n,m} \subset S$ wählen, so dass $\mathbb{P}[X_n(t) \in \Gamma_{n,m}^{1/m} \text{ für } 0 \leq t \leq T] \geq 1 - \eta 2^{-m}$. Die Menge $\Gamma_n = \bigcup_{i=0}^{n_m-1} \Gamma_{i,m}$ ist kompakt, und $\inf_{n \geq 1} \mathbb{P}[X_n(t) \in \Gamma_n^{1/m} \text{ für } 0 \leq t \leq T] \geq 1 - \eta 2^{-m}$. Wir wählen nun $\Gamma_{\eta,T} = \bigcap_{m \geq 1} \Gamma_m^{1/m}$. Die Menge ist abgeschlossen und total beschränkt und damit kompakt, und erfüllt (2.14). Da H dicht in $\bar{C}(S)$ ist, und $\bar{C}(S)$ dicht in $B(E)$ ist, ist es nun leicht zu zeigen, dass die endlich-dimensionalen Verteilungen von X_n konvergieren. Somit folgt die Behauptung aus Satz 2.39. $\square$

Korollar 2.41. Sei S lokal kompakt und separabel, und S^Δ die Einpunkt-Kompaktifizierung. Ist $\{X_n\}$ eine Folge von Prozessen mit Pfaden in $D_S[0, \infty)$ und $\{f \circ X_n\}$ relativ kompakt für jedes $f \in \hat{C}(S)$ (der Raum der stetigen Funktionen, die in Unendlich verschwinden), dann ist $\{X_n\}$ relativ kompakt als Folge von Prozessen mit Pfaden in $D_{S^\Delta}[0, \infty)$. Gilt zusätzlich, dass $(X_n(t_1), \dots, X_n(t_k)) \Rightarrow (X(t_1), \dots, X(t_k))$ für alle endlichen Teilmengen $\{t_1, \dots, t_k\}$ einer dichten Menge $D \subset [0, \infty)$, für einen Prozess X mit Pfaden in $D_S[0, \infty)$, dann gilt $X_n \Rightarrow X$ in $D_S[0, \infty)$.

Beweis. Ist $f \in C(S^\Delta)$, so ist $f - f(\Delta) \mid_S \in \hat{C}(S)$. Also ist $\{f \circ X_n\}$ relativ kompakt für jedes $f \in C(S^\Delta)$. Somit folgt aus Satz 2.39, dass $\{X_n\}$ relativ kompakt in $D_{S^\Delta}[0, \infty)$. Aus Satz 2.33 folgt, dass $X_n \Rightarrow X$ in $D_{S^\Delta}[0, \infty)$. Aus Korollar 2.18 folgt, dass $X_n \Rightarrow X$ in $D_S[0, \infty)$. $\square$

Satz 2.42. Sei $\{X^\alpha\}$ eine Familie von Prozessen mit Pfaden in $D_S[0, \infty)$ adaptiert zu einer Filtration $\{\mathcal{F}_t^\alpha\}$. Sei $\mathcal{L}^\alpha$ der Banach-Raum der $\{\mathcal{F}_t^\alpha\}$ -adaptierten Prozesse Y mit Norm $\|Y\| = \sup\{\mathbb{E}[|Y_t|] : t \geq 0\} < \infty$, und

$$\hat{\mathcal{A}}^\alpha = \left\{ (Y, Z) \in (\mathcal{L}^\alpha)^2 : Y_t - \int_0^t Z_s \, ds \text{ ist ein } \{\mathcal{F}_t^\alpha\}\text{-Martingal} \right\}.$$

Sei C_a eine Unteralgebra von $\bar{C}(S)$ und sei D die Familie der Funktionen $f \in \bar{C}(S)$, so dass es für jedes $\varepsilon > 0$ und $T > 0$ ein $(Y_\alpha, Z_\alpha) \in \hat{\mathcal{A}}^\alpha$ gibt, so dass

$$\sup_\alpha \mathbb{E} \left[\sup_{t \in [0, T] \cap \mathbb{Q}} |Y_\alpha(t) - f(X_\alpha(t))| \right] < \varepsilon$$

und

$$\sup_\alpha \mathbb{E}[\|Z_\alpha\|_{p, T}] < \infty \quad \text{für ein } p \in (1, \infty],$$

wobei $\|h\|_{p, T} = (\int_0^T |h(t)|^p \, dt)^{1/p}$ für $p < \infty$ und $\|h\|_{\infty, T} = \text{ess sup}\{|h(t)| : 0 \leq t \leq T\}$. Falls C_a im Abschluss von D enthalten ist, dann ist $\{(f_1, \dots, f_k) \circ X_\alpha\}$ relativ kompakt in $D_{\mathbb{R}^k}[0, \infty)$ für alle $f_1, \dots, f_k \in C_a$ und $1 \leq k \leq \infty$.

Beweis. Der Satz ist in [1] bewiesen. $\square$