Z. Wahrscheinlichkeitstheorie verw. Geb. 28, 37–42 (1973)

Konvergenz stetiger stochastischer Prozesse mit polnischem Zustands- und Parameterraum*

Ludwig Fahrmeir

1. Einleitung

In [8] werden Bedingungen dafür angegeben, daß fast alle Pfade eines reellen Prozesses mit Parameter aus einem pseudometrischen Raum stetig sind. [5, 9, 12] beschäftigen sich mit schwacher Konvergenz von reellen Prozessen mit mehrdimensionalem Zeitparameter.

Wir setzen nun stochastische Prozesse, deren Pfade fast alle stetig sind, voraus und lassen als Zustandsraum E und Parameterraum T polnische (d.h. metrische, separable und vollständige) Räume zu; T ist zusätzlich lokalkompakt. Solche Prozesse lassen sich hinreichend charakterisieren durch zugeordnete Wahrscheinlichkeitsmaße auf dem Raum $C(T, E)$ der stetigen Abbildungen von T in E (3.). Mit Hilfe dieser Konstruktion können wir die Begriffe „schwache Konvergenz" bzw. „Verteilungskonvergenz" der zugrundeliegenden Prozesse einführen (4.). In Verallgemeinerung von Ergebnissen in [3, 6, 13] für $T=[0, 1]$ oder $T=[0, \infty)$ gelangt man so zu Konvergenzkriterien für stetige stochastische Prozesse mit polnischem Zustands- und Parameterraum.

2. Eigenschaften des Raumes $C(T, E)$

(T, d_T) sei im folgenden, wenn nicht anders gesagt, immer ein metrischer, separabler, lokalkompakter Raum mit Metrik d_T, (E, d_E) metrisch, vollständig und separabel mit Metrik d_E, E^T der Raum aller Abbildungen von T in E, $C \equiv C(T, E)$ der Raum aller stetigen derartigen Abbildungen.

Bemerkung. Aus der Lokalkompaktheit und Separabilität von (T, d_T) folgt, daß eine T überdeckende, isotone Folge $(A_i)_{i \in N}$ kompakter Mengen existiert. Jede kompakte Menge ist dann in einem A_i enthalten ([11]). Außerdem ist T vollständig ([2]). In der Terminologie von [2, 4] sind also E und T polnische Räume.

Wir definieren mit Hilfe von

$$(2.1) \qquad r_i(x, y) := \sup_{t \in A_i} d_E(x(t), y(t)), \qquad (x, y) \in C \times C$$

die Abbildung $r\colon C \times C \to R$ durch

$$(2.2) \qquad r(x, y) := \sum_{i=1}^{\infty} 2^{-i} r_i(x, y)/(1 + r_i(x, y)).$$

Satz 2.1. a) *r ist eine Metrik auf C.*

---

* Diese Arbeit ist Teil der Dissertation des Autors bei Prof. J. Heinhold am Institut für Angew. Math. der TU München.

b) *für jede Folge* $(x_n)_{n\in N}$, $x_n \in C$, *gilt:*

$$\lim_{n\to\infty} r(x_n, x) = 0 \qquad \text{genau dann, wenn } \lim_{n\to\infty} r_i(x_n, x) \text{ für alle } i.$$

c) (C, r) *ist vollständig und separabel.*

Beweis. a) und b) sind leicht nachzuprüfen. b) besagt zusammen mit (2.1), daß die von r auf C induzierte uniforme Struktur die der gleichmäßigen Konvergenz auf kompakten Teilmengen ist ([11]). Die Vollständigkeit und Separabilität folgt nun aus [4].

Wir haben also noch erhalten das

Korollar. *Die Topologie von* (C, r) *ist die Topologie der kompakten Konvergenz.*

Es seien nun $\mathfrak{B}$ bzw. $\mathfrak{B}_C$ die σ-Algebra der Borelschen Mengen von (E, d_E) bzw. (C, r), $\mathfrak{B}^T$ die Produkt-σ-Algebra in E^T, $\sigma(\pi_t, t\in T)$ die von den Projektionen $\pi_t\colon C \to E$, $x \mapsto x(t)$, $t\in T$, erzeugte σ-Algebra in C.

Über die Beziehungen zwischen den verschiedenen angegebenen σ-Algebren gilt

Satz 2.2. *Es ist*

$$\mathfrak{B}_C = C \cap \mathfrak{B}^T = \sigma(\pi_t, t\in T). \tag{2.3}$$

Beweis. 1. $C \cap \mathfrak{B}^T = \sigma$ ergibt sich sofort aus $\pi_t^{-1}(B) = pr_t^{-1}(B) \cap C$, $B\in\mathfrak{B}$ $(pr_t\colon E^T \to E$ ist die Projektionsabbildung $x \mapsto x(t))$.

2. $\sigma = \mathfrak{B}_C$ beweist man ähnlich wie in [10], wobei nur der Teil $\mathfrak{B}_C \subset \sigma$ etwas abzuändern ist.

3. Stetige stochastische Prozesse und Wahrscheinlichkeitsmaße auf $C(T, E)$

$(\Omega, \mathfrak{A}, p, (X_t)_{t\in T})$ sei ein stetiger stochastischer Prozeß mit Zustandsraum $(E, \mathfrak{B})$, d.h. p-fast alle Pfade sollen in $C(T, E)$ liegen. Die (meßbare) Abbildung

$$X\colon (\Omega, \mathfrak{A}, p) \to (E^T, \mathfrak{B}^T),$$

die jedem $\omega\in\Omega$ den Pfad $(X_t(\omega))$ zuordnet, erzeugt auf (E^T, B^T) das Bildmaß P von p.

Nun gilt zwar $C \subset E^T$, C ist aber i.allg. nicht aus $\mathfrak{B}^T$, d.h. C kann durch P keine Wahrscheinlichkeit zugeordnet werden. Es gilt jedoch $P^*(C) = 1$, wenn P^* das P zugeordnete äußere Maß ist (zum Beweis der beiden Behauptungen folge man den Ideen von [2], S. 294 ff.).

Durch Einschränkung des auf $(E^T, \mathfrak{B}^T)$ erklärten W-Maßes P kann man aber zu einem W-Maß P_C auf $(C, \mathfrak{B}_C)$ gelangen: Zu jedem $B\in\mathfrak{B}_C$ existiert wegen Satz 2.2 ein $Q\in\mathfrak{B}^T$ mit $B = C\cap Q$. Dann wird P_C erklärt durch

$$P_C(B) = P_C(C\cap Q) := P(Q), \qquad B\in\mathfrak{B}_C. \tag{3.1}$$

Man zeigt leicht, daß $P_C(B)$ unabhängig ist von der Wahl des „Repräsentanten" Q.

Falls umgekehrt auf $(C, \mathfrak{B}_C)$ ein W-Maß P_C gegeben ist, so kann man P_C auf $(E^T, \mathfrak{B}^T)$ ausdehnen:

$$i\colon (C, \mathfrak{B}_C, P_C) \to (E^T, \mathfrak{B}^T)$$

sei die (meßbare) kanonische Injektion. Das Bildmaß von P_C bezüglich i liefert dann ein W-Maß P auf $(E^T, \mathfrak{B}^T)$. Die W-Räume $(C, \mathfrak{B}_C, P_C)$ und $(E^T, \mathfrak{B}^T, P)$ liefern also dieselbe Information über (X_t). Da aber $(C, \mathfrak{B}_C, P_C)$ angenehmer zu handhaben ist, beschränken wir uns darauf, schreiben im folgenden wieder einfach P statt P_C und nennen P die Verteilung von (X_t).

4. Schwache Konvergenz von W-Maßen auf $C(T, E)$ und Verteilungskonvergenz von stetigen stochastischen Prozessen

Nach den Vorbereitungen in 2. und 3. lassen sich nunmehr sämtliche Ergebnisse aus [13] über Straffheit und schwache Konvergenz von W-Maßen auf $C[0, \infty)$ ohne große Mühe verallgemeinern. Wir führen nur die wichtigsten Sätze an, geben aber darüber hinaus noch Kriterien für die Verteilungskonvergenz von stetigen Prozessen an.

Bezeichnungen.

$W(C)$ Menge aller W-Maße auf $(C, \mathfrak{B}_C)$,
$\mathfrak{E} = \mathfrak{E}(T)$ System aller nichtleeren, endlichen Teilmengen von T,
P^I die zu $P \in W(C)$ gehörigen, von den Projektionen $\pi_I: C \to E^I$, $I \in \mathfrak{E}$, erzeugten endlich-dimensionalen Verteilungen,
$C_i = C(A_i, E)$ Menge der stetigen Abbildungen von $A_i \subset T$ in E,
r_i durch (2.1) definierte Metrik auf C_i,
$\mathfrak{B}_{C_i}$ σ-Algebra der Borelschen Mengen von (C_i, r_i),
res_i: $C \to C_i$, $x \mapsto x|A_i$; ordnet $x \in C$ die Einschränkung auf A_i zu,
P^i Bildmaß von P unter res_i.

Definitionen.

(4.1) $$w_x^i(h) := \sup_{\substack{s,t \in A_i \\ d_T(s,t) < h}} d_E(x(t), x(s)); \qquad i \in N,\ x \in C,\ h > 0$$

heißt Modulus von x auf A_i.

(4.2) $(P_n)_{n \in N}$, $P_n \in W(C)$, heißt straff, wenn zu jedem $\varepsilon > 0$ eine kompakte Menge $K \subset C$ existiert mit $P_n(K) > 1 - \varepsilon$ für alle $n \in N$.

(4.3) (X_t), $(X_t^n)_{n \in N}$ seien stetige Prozesse mit Parameterraum T und Zustandsraum E; P, P_n die wie in 3. zugeordneten W-Maße auf $C(T, E)$. Dann heißt (X_t^n) gegen (X_t) verteilungskonvergent, wenn P_n schwach gegen P konvergiert für $n \to \infty$, i.Z.

$$(X_t^n) \mapsto (X_t) \text{ genau dann, wenn } P_n \Rightarrow P.$$

Es gilt nun der wichtige, die schwache Konvergenz durch die endlich-dimensionalen Verteilungen charakterisierende

Satz 4.1. $(P_n)_{n \in N}$, $P_n \in W(C)$, *konvergiert genau dann schwach gegen* $P \in W(C)$, *wenn für alle* $I \in \mathfrak{E}$ *die endlich-dimensionalen Verteilungen* $(P_n^I)_{n \in N}$ *schwach gegen die endlich-dimensionalen Verteilungen* P^I *konvergieren und wenn* $(P_n)_{n \in N}$ *straff ist; i.Z.*

$$P_n \Rightarrow P \text{ genau dann, wenn } P_n^I \Rightarrow P^I, \quad I \in \mathfrak{E}, \ \wedge (P_n)_{n \in N} \text{ straff}.$$

Beweis. Man kann [3], S. 35, 54, 241, fast wörtlich übernehmen.

Entscheidend ist ja nur, daß C polnisch ist, um den Satz von Prohorov ([3, 10]) anwenden zu können.

Charakterisiert man nun die in der Straffheitsdefinition (4.2) auftretende kompakte Menge K mit Hilfe des Satzes von Arzelà-Ascoli-Bourbaki ([11]) durch spezielle Eigenschaften von C, so erhält man über sich so ergebende Straffheitsbedingungen aus Satz 4.1 weitere Kriterien für die schwache Konvergenz von W-Maßen auf C. Die genaue Durchführung ([7]) besteht im wesentlichen in der entsprechenden Verallgemeinerung von [13]. Wir geben deshalb hier nur die beiden wichtigsten Ergebnisse an.

Satz 4.2. *Die Folge $(P_n)_{n\in N}$ auf $(C, \mathfrak{B}_C)$ konvergiert genau dann schwach gegen $P\in W(C)$, wenn für alle $i\in N$ die Folge $(P_n^i)_{n\in N}$ auf $(C_i, \mathfrak{B}_{C_i})$ schwach gegen $P^i\in W(C_i)$ konvergiert.*

Bemerkung. Da die Abbildung res_i stetig ist, folgt „$P_n^i \Rightarrow P^i$" aus „$P_n \Rightarrow P$" bereits durch Theorem 5.1 von [3]. Interessant ist also vor allem, daß im Falle von W-Maßen auf $(C, \mathfrak{B}_C)$ auch die Umkehrung gilt.

Satz 4.3. *$(P_n)_{n\in N}$ konvergiert genau dann schwach gegen P, wenn gilt*

a) *die endlich-dimensionalen Verteilungen von P_n konvergieren schwach gegen die endlich-dimensionalen Verteilungen von P und*

b) $\lim_{h\to 0} \sup_n P_n\{x\in C\colon w_x^i(h)>\varepsilon\}=0$ *für alle* $\varepsilon>0$, $i\in N$.

Bemerkung. Das Auftreten des Modulus $w_x^i(h)$ rührt von der Bedingung der gleichgradigen Stetigkeit im Satz von Arzelà her.

Seien nun $(X_t^n)_{n\in N}$, (X_t) auf $(\Omega, \mathfrak{A}, p)$ erklärte stetige stochastische Prozesse mit Parameter aus T und Werten in E. P_n, P seien die wie in 3. zugeordneten W-Maße auf $C(T, E)$. Dann gilt

Satz 4.4. *$(X_t^n)_{n\in N}$ ist verteilungskonvergent gegen (X_t) genau dann, wenn gilt*

a) *die endlich-dimensionalen Verteilungen von (X_t^n) konvergieren schwach gegen die endlich-dimensionalen Verteilungen von (X_t).*

b) $\lim_{h\to 0} \sup_n p\{\omega\colon w^i_{X_t^n(\omega)}(h)>\varepsilon\}=0$.

Beweis. $P_n\{x\in C\colon w_x^i(h)>\varepsilon\}=p\{\omega\colon w^i_{X_t^n(\omega)}(h)>\varepsilon\}$.

Bemerkung. Im Falle $E=R$ wird $C(R,T)$ zu einem Banachraum. Die Konvergenzsätze benützen jedoch in keiner Weise diese Eigenschaft, da sie ja auch gelten wenn E polnisch ist. Deshalb sei darauf hingewiesen, daß sich im Falle $E=R$ ein Kriterium in [1] für $T=[0,1]$, das die Linearität von $C[0,1]$ benutzt, ohne Schwierigkeit auf polnische Parametermengen T verallgemeinern läßt.

Zum Schluß bringen wir noch ein für die Praxis handlicheres Kriterium für den Fall $T=R_+$.

Satz 4.5. *Konvergieren die endlich-dimensionalen Verteilungen $(P_n^I)_{n\in N}$ der Folge $(X_t^n)_{n\in N}$ schwach gegen die endlich-dimensionalen Verteilungen P^I von (X_t) und gilt weiterhin*

(4.4) *Für jedes $i\geqq 1$ existieren positive, reelle Konstanten a_i, b_i, c_i, so daß für alle $\varepsilon>0$ gilt:*

$$p\{\omega\colon d_E(X_t^n(\omega), X_n^n(\omega))>\varepsilon\}\leqq \frac{c_i\,|t-s|^{b_i+1}}{\varepsilon^{a_i}}, \qquad t,s\in[0,i],$$

dann folgt:

$$P_n \Rightarrow P$$

bzw.

$$(X_t^n) \mapsto (X_t).$$

Korollar. *Bedingung* (4.4) *ist erfüllt, falls gilt*

(4.5) *Für alle* $i \geqq 1$ *existieren Konstanten* $a_i, b_i, c_i > 0$ *mit*

$$E\{d_E(X_t^n, X_s^n)^{a_i}\} \leqq c_i |t-s|^{1+b_i}, \qquad t, s \in [0, i].$$

Bemerkung. Das Korollar verallgemeinert ein Ergebnis in [6] für den Fall $E=R$, $T=[0,1]$.

Beweis. Wir zeigen, daß aus (4.4) die Bedingung b) von Satz 4.4 folgt; daraus folgt dann die Behauptung. Zunächst gilt:

$$\begin{aligned} A &:= \{\omega: w^i_{X_t^n(\omega)}(h) > \varepsilon\} = \{\omega: \sup_{\substack{|t-s| \leqq h \\ t, s \in [0, i]}} d_E(X_t^n(\omega), X_s^n(\omega)) > \varepsilon\} \\ &= \{\omega: \sup_{\substack{|t-s| \leqq h \\ t, s \in [0, i] \cap S}} d_E(X_t^n(\omega), X_s^n(\omega)) > \varepsilon\} \end{aligned}$$

mit $S := \{m \cdot 2^{-n}; m, n = 0, 1, 2, \ldots\}$. Die letzte Gleichung gilt wegen der Stetigkeit der Pfade und da S in R_+ dicht ist. $i \geqq 1$, $\varepsilon > 0$, $\delta > 0$ seien vorgegeben. Wir wählen m_0 so, daß

(4.6) $$m_0 > 1,$$

(4.7) $$i \cdot c_i \cdot \sum_{m=m_0}^{\infty} \frac{m^{2a_i}}{2^{m b_i}} < \delta,$$

(4.8) $$\sum_{m=m_0}^{\infty} \frac{1}{m^2} < \varepsilon/2$$

und setzen $h := 2^{-m_0}$.

Wir definieren $B \in \mathfrak{A}$ durch

$$B := \bigcup_{m=m_0}^{\infty} \bigcap_{k \leqq i 2^m} \left\{\omega: d_E\left(X^n_{\frac{k+1}{2^m}}(\omega), X^n_{\frac{k}{2^m}}(\omega)\right) > \frac{1}{m^2}\right\}.$$

Wegen (4.4) und (4.7) können wir abschätzen:

$$\begin{aligned} p(B) &\leqq \sum_{m=m_0}^{\infty} \sum_k \left\{\omega: d_E(\;,\;) > \frac{1}{m^2}\right\} \\ &\leqq \sum_{m=m_0}^{\infty} i \cdot 2^m c_i m^{2a_i} 2^{-m(1+b_i)} \\ &= \sum_{m=m_0}^{\infty} i \cdot c_i m^{2a_i} \cdot 2^{-m b_i} \\ &< \delta. \end{aligned}$$

Wegen (4.6) und (4.7) kann man die Abschätzung aus [2] (S. 299) $\left(\text{mit } \eta_\nu := \frac{1}{\nu^2}\right)$ anwenden und erhält:

$$\bigcup_{\substack{|s-t| \leqq 2^{-m_0} \\ s,t \in S \cap [0,i]}} \{\omega : d_E(X_s^n(\omega), X_t^n(\omega)) > \varepsilon\} \subset B.$$

Daraus folgt auch $A \subset B$ und $p(A) < p(B) < \delta$, also

$$\sup_n p(A) < \delta,$$

da die Abschätzung von n unabhängig war.

Beweis des Korollars.

$$p\{\omega : d_E(\ , \ ) > \varepsilon\} \leqq \frac{E\, d_E(\ , \ )^{a_i}}{\varepsilon^{a_i}} \leqq \frac{c_i |t-s|^{1+b_i}}{\varepsilon^{a_i}}.$$

Die 1. Ungleichung folgt aus der Tschebyscheffschen Ungleichung, die zweite aus (4.5).

Bemerkung. Das Korollar erweist sich sehr nützlich, um die schwache Konvergenz numerisch realisierbarer Näherungsprozesse gegen Diffusionsprozesse zu beweisen ([7]).

Literatur

1. Acosta, A. D. de: Existence and Convergence of Probability Measures in Banach Spaces. Trans. Amer. Math. Soc. **152**, 273–298 (1970)
2. Bauer, H.: Wahrscheinlichkeitstheorie und Grundzüge der Maßtheorie. Berlin: Walter de Gruyter & Co. 1968
3. Billingsley, P.: Convergence of Probability Measures. New York: John Wiley & Sons 1968
4. Bourbaki, N.: Livre III, Topologie générale; Chapitre 10, Espaces fonctionels. Paris: Hermann 1961
5. Dudley, R. M.: Weak convergence of probabilities on nonseparable metric spaces and empirical measures on Euclidean Spaces. Trans. Amer. Math. Soc. **87**, 109–126 (1966)
6. Gikhman, I. I., Skorohod, A. V.: Introduction to the Theory of Random Processes. London: W. B. Saunders Comp. 1969
7. Fahrmeir, L.: Schwache Konvergenz von Wahrscheinlichkeitsmaßen und ein hybrides Monte-Carlo-Verfahren für lineare Randwertaufgaben. Dissertation TU München (1972)
8. Neuhaus, G.: Stetigkeitseigenschaften stochastischer Prozesse mit Parameter aus einem pseudometrischen Raum. Z. Wahrscheinlichkeitstheorie verw. Gebiete **23**, 275–288 (1972)
9. Neuhaus, G.: On weak convergence of stochastic processes with multidimensional time parameter. Ann. Math. Statist **42**, 1285–1295 (1971)
10. Parthasarathy, K. R.: Probability Measures on Metric Spaces. New York: Academic Press 1967
11. Schubert, H.: Topologie. Stuttgart: Teubner 1964
12. Wichura, M. J.: Inequalities with applications to the weak convergence of random processes with multidimensional time parameters. Ann. Math. Statist. **40**, 681–687 (1969)
13. Whitt, W.: Weak convergence of Probability Measures on the Function Space $C[0, \infty)$. Ann. Math. Statist. **41** (1970)

Ludwig Fahrmeir
Institut für Angewandte Mathematik
der Technischen Universität
D-8000 München 2, Arcisstraße 21
Bundesrepublik Deutschland

(Eingegangen am 28. Februar 1973)