Illinois Journal of Mathematics
Volume 51, Number 4, Winter 2007, Pages 1231–1242
S 0019-2082

ON THE CONVERGENCE OF MEASURABLE PROCESSES AND PREDICTION PROCESSES

HIDEATSU TSUKAHARA

Dedicated to the memory of Catherine Doléans-Dade and Frank Knight

Abstract. We study and characterize laws of measurable processes and their convergence with general state space and parameter set. Using those results, it is shown that convergence of the prediction processes implies that of the given processes. We also give a simple condition for convergence of the prediction processes when the given processes are Markovian.

1. Introduction

We consider $\mathbb{E}$-valued measurable processes on a σ-finite measure space $(\mathbb{T}, \mathscr{T}, \nu)$, where $\mathbb{E}$ is a metrizable Lusin space. They induce the laws on the space $\mathbb{M}_{\mathbb{E}}(\mathbb{T})$ of $\mathbb{E}$-valued measurable functions on $(\mathbb{T}, \mathscr{T}, \nu)$. In Section 2, we first study properties of probability measures on $\mathbb{M}_{\mathbb{E}}(\mathbb{T})$ and characterize them in terms of their finite-dimensional distributions. The notions of pseudo-path and pseudo-law (Dellacherie and Meyer [4]) are closely related, and it is proved that two measurable processes are almost equivalent if they induce the same law on $\mathbb{M}_{\mathbb{E}}(\mathbb{T})$. Furthermore, we remark that for measurable processes, only the finite-dimensional convergence on a set of full measure is sufficient for weak convergence in $\mathbb{M}_{\mathbb{E}}(\mathbb{T})$ and that its converse in a sense also holds. These are rather straightforward extensions of well-known results and detailed proofs of the unproved results in Section 2 may be found in Tsukahara [14].

In Section 3, we apply the results obtained in Section 2 to the prediction process (see Knight (1981, 1992)). The prediction process Z^z of a given measurable process X with law z on $\mathbb{M}_{\mathbb{E}}(\mathbb{R}_+)$ is the process consisting of the conditional distributions of the future of X given the past at each time $t \in \mathbb{R}_+$. Our interest is in their convergence in law; specifically, we give an alternative

Received March 31, 2006; received in final form July 13, 2006.
2000 *Mathematics Subject Classification.* Primary 60F17. Secondary 60G07, 60G25, 60J25.
The author would like to express his deep gratitude to Ditlev Monrad, Frank Knight and Catherine Doléans-Dade for their encouragement and helpful comments.

©2007 University of Illinois

proof of the fact that the convergence of the prediction processes is stronger than that of the given processes. In the special case when the given processes are all homogeneous Markov processes, we provide a simple condition on the resolvents sufficient for the convergence of the associated prediction processes.

2. Laws of measurable processes

Let $(\mathbb{T}, \mathscr{T}, \nu)$ be a σ-finite measure space and $(\mathbb{E}, d)$ a metric space. We set $\mathscr{E} = \mathscr{B}_d(\mathbb{E})$, the Borel σ-field on $\mathbb{E}$ generated by the d-open sets. Furthermore, we denote by $\mathbb{M}_{\mathbb{E}}(\mathbb{T}, \mathscr{T})$ the space of all $\mathscr{T}/\mathscr{E}$ measurable $\mathbb{E}$-valued functions on $\mathbb{T}$. We write $x \sim y$ for $x,\ y \in \mathbb{M}_{\mathbb{E}}(\mathbb{T}, \mathscr{T})$ if $\nu(t : x(t) \neq y(t)) = 0$. Let $\widetilde{\mathbb{M}}_{\mathbb{E}}(\mathbb{T}, \mathscr{T}) = \mathbb{M}_{\mathbb{E}}(\mathbb{T}, \mathscr{T})/\!\sim$ be the space of all equivalence classes of $\mathscr{T}/\mathscr{E}$ measurable functions.

Since $(\mathbb{T}, \mathscr{T}, \nu)$ is σ-finite, there exists a finite measure λ such that $\nu \ll \lambda$ and $\lambda(B) \leq \nu(B)$ for all $B \in \mathscr{T}$. Suppose now that $\mathbb{E}$ is separable. A sequence (w_n) converges in λ-measure to w if and only if for every $\varepsilon > 0$ and every $A \in \mathscr{F}$ with $\nu(A) < \infty$ we have $\lim_{n\to\infty} \nu(\{d(w_n(t), w(t)) \geq \varepsilon\} \cap A) = 0$. For $v, w \in \mathbb{M}_{\mathbb{E}}(\mathbb{T}, \mathscr{T})$,

$$\rho_\lambda(v, w) \triangleq \int_{\mathbb{T}} 1 \wedge d(v(t), w(t))\, \lambda(dt)$$

defines a pseudo-metric on $\mathbb{M}_{\mathbb{E}}(\mathbb{T}, \mathscr{T})$ and ρ_λ-convergence is convergence in λ-measure. We write $\mathbb{M}_{\mathbb{E}}(\lambda) = (\mathbb{M}_{\mathbb{E}}(\mathbb{T}, \mathscr{T}), \rho_\lambda)$ (pseudo-metric space). The corresponding metric space $\widetilde{\mathbb{M}}_{\mathbb{E}}(\lambda)$ is defined in an obvious way. Denote by τ_λ the topology induced by $\widetilde{\rho}_\lambda$. Dellacherie [3] showed that w_n converges to w in λ-measure if and only if $\int_A f(w_n(t))\, \lambda(dt) \to \int_A f(w(t))\, \lambda(dt)$ for every $f \in C_b(\mathbb{E})$ and every $A \in \mathscr{T}$. Thus the topology τ_λ does not depend on the metric d on $\mathbb{E}$. It is well known that if $(\mathbb{E}, d)$ is a separable metric space and if $\mathscr{T}$ is countably generated up to null sets, then $\mathbb{M}_{\mathbb{E}}(\lambda)$ is separable. If in addition $(\mathbb{E}, d)$ is complete, then $\mathbb{M}_{\mathbb{E}}(\lambda)$ is complete and separable (see, e.g., Kurtz [8]). When we discuss weak convergence of probability measures on $\widetilde{\mathbb{M}}_{\mathbb{E}}(\lambda)$, it is important, because of the Prohorov theorem, to find a compactness criterion in $\widetilde{\mathbb{M}}_{\mathbb{E}}(\lambda)$. For this type of result, see Kurtz [8] for the case $(\mathbb{T}, \mathscr{T}, \nu) = (\mathbb{R}_+, \mathscr{B}_+, m)$, where m is the Lebesgue measure, and Tsukahara [14] for the case where $\mathbb{T}$ has group structure.

In the special case where $(\mathbb{T}, \mathscr{T}, \nu) = (\mathbb{R}_+, \mathscr{B}_+, m)$, where m is the Lebesgue measure, let us write $\mathbb{M}_{\mathbb{E}} = \mathbb{M}_{\mathbb{E}}(\mathbb{R}_+, \mathscr{B}_+, \lambda)$, where $\mathscr{B}_+ = \mathscr{B}(\mathbb{R}_+)$ and $\lambda(dt) = e^{-t}dt$. It is shown in Knight [7] that if E is a metrizable Lusin space, $\widetilde{\mathbb{M}}_{\mathbb{E}}$ is also a metrizable Lusin space. The following lemma, due to Knight [7] and Kurtz [8], gives us a way of picking a function from each equivalence class in a measurable fashion.

LEMMA 2.1. *There exists a* $\mathscr{B}(\widetilde{\mathbb{M}}_{\mathbb{E}}) \otimes \mathscr{B}(\mathbb{R}_+)/\mathscr{E}$ *measurable mapping* G *from* $\widetilde{\mathbb{M}}_{\mathbb{E}} \times \mathbb{R}_+$ *into* $\mathbb{E}$ *such that* $G(\widetilde{w}, \bullet) \in \widetilde{w}$ *for all* $\widetilde{w} \in \widetilde{\mathbb{M}}_{\mathbb{E}}$.

Next, we discuss probability measures on $\mathbb{M}_{\mathbb{E}}$. Let $X = (X_t)_{t\in\mathbb{T}}$ be a measurable process on $(\Omega, \mathscr{F}, P)$ with values in $(\mathbb{E}, \mathscr{E})$. We assume until the end of this section, unless otherwise stated, that $\mathbb{E}$ is a metrizable Lusin space. Since X is measurable, the paths $X_\bullet(\omega)$ belong to $\mathbb{M}_{\mathbb{E}}(\lambda)$ for each $\omega \in \Omega$. Let us denote by $\widetilde{X}(\omega)$ the equivalence class containing $X_\bullet(\omega)$. If $\mathscr{T}$ is countably generated up to ν-null sets, then it is easy to see that the mapping $\omega \mapsto \widetilde{X}(\omega)$ is $\mathscr{F}/\mathscr{B}(\widetilde{\mathbb{M}}_{\mathbb{E}}(\lambda))$ measurable. Thus the mapping $\omega \mapsto \widetilde{X}(\omega)$ induces a probability law on $\widetilde{\mathbb{M}}_{\mathbb{E}}(\lambda)$. Conversely, we have the following lemma, whose proof is inspired by Skorokhod [12].

LEMMA 2.2. *Suppose that $\mathscr{T}$ is countably generated up to ν-null sets. Then for any $\widetilde{\mathbb{M}}_{\mathbb{E}}(\lambda)$-valued random variable $\widetilde{X}$, there is a measurable process X such that $X_\bullet(\omega) \in \widetilde{X}(\omega)$.*

Proof. First let us assume $\mathbb{E} = [0,1]$ and consider $\mathbb{M}_{[0,1]} = \mathbb{M}_{[0,1]}(\lambda)$ as a subset of $\mathbb{L}^2(\mathbb{T}, \mathscr{T}, \lambda)$. Since $\mathscr{T}$ is countably generated, $\mathbb{L}^2(\mathbb{T}, \mathscr{T}, \lambda)$ is separable. Thus there exists a countable orthonormal basis $(\widetilde{\phi}_j)_{j\in\mathbb{N}}$ for $\mathbb{L}^2(\mathbb{T}, \mathscr{T}, \lambda)$. Pick a representative ϕ_j for each $\widetilde{\phi}_j$ and fix these. If $w \in \mathbb{M}_{[0,1]}$, then we have a representation

$$w(t) = \sum_{j=1}^{\infty} \langle w, \phi_j\rangle \phi_j(t),$$

where the limit is in $\mathbb{L}^2(\mathbb{T}, \mathscr{T}, \lambda)$, the equality holds λ-a.e. and $\langle w, \phi_j\rangle = \int_{\mathbb{T}} w(t)\phi_j(t)\lambda(dt)$. Note that the value of $\langle w, \phi_j\rangle$ is the same for any $w' \in \widetilde{w}$, so we may write $\langle \widetilde{w}, \phi_j\rangle$. Put

$$\gamma_n(\widetilde{w}, t) \triangleq \sum_{j=1}^{n} \langle \widetilde{w}, \phi_j\rangle \phi_j(t).$$

Clearly the mapping $(\widetilde{w}, t) \mapsto \gamma_n(\widetilde{w}, t)$ is $\mathscr{B}(\mathbb{M}_{[0,1]}) \otimes \mathscr{T}$ measurable. Define $n_k(\widetilde{w})$ to be the smallest positive integer n for which

$$\sup_{m>n} \lambda\left\{t \in \mathbb{T}\colon |\gamma_n(\widetilde{w}, t) - \gamma_m(\widetilde{w}, t)| > \frac{1}{k^2}\right\} \le \frac{1}{k^2}$$

holds. Then one can easily see that the mapping $\widetilde{w} \mapsto n_k(\widetilde{w})$ is $\mathscr{B}(\mathbb{M}_{[0,1]})$ measurable. Set

$$g(\widetilde{w}, t) \triangleq \begin{cases} \limsup_{k\to\infty} \gamma_{n_k(\widetilde{w})}(\widetilde{w}, t), & \text{if } \limsup_{k\to\infty} \gamma_{n_k(\widetilde{w})}(\widetilde{w}, t) \in [0,1], \\ 0, & \text{otherwise.} \end{cases}$$

It is evident from the construction that $g(\widetilde{w}, \bullet) \in \widetilde{w}$ and $(\widetilde{w}, t) \mapsto g(\widetilde{w}, t)$ is $\mathscr{B}(\mathbb{M}_{[0,1]}) \otimes \mathscr{T}$ measurable.

As a measurable space, $(\mathbb{E}, \mathscr{E})$ is of course a measurable Lusin space. Thus, by Kuratowski's theorem (see Dellacherie and Meyer [4], III.20), there exists an isomorphism h on $(\mathbb{E},\mathscr{E})$ into $[0,1]$ with $h(\mathbb{E}) \in \mathscr{B}[0,1]$. Let $e_0 \in \mathbb{E}$ be arbitrary but fixed. Define

$$G(\widetilde{w},t) \triangleq \begin{cases} h^{-1}(g(\widetilde{h\circ w},t)), & \text{if } g(\widetilde{h\circ w},t) \in h(\mathbb{E}),\\ e_0, & \text{otherwise.}\end{cases}$$

Clearly one has $G(\widetilde{w},\bullet) \in \widetilde{w}$ and $(\widetilde{w},t) \mapsto G(\widetilde{w},t)$ is $\mathscr{B}(\mathbb{M}_{\mathbb{E}}(\lambda)) \otimes \mathscr{T}$ measurable. It then follows that for any $\widetilde{\mathbb{M}}_{\mathbb{E}}$-valued random variable $\widetilde{X}$, $X_t(\omega) \triangleq G(\widetilde{X}(\omega),t)$ defines a measurable process and $X_\bullet(\omega) \in \widetilde{X}(\omega)$. □

The notion of probability law on $\widetilde{\mathbb{M}}_{\mathbb{E}}(\lambda)$ induced by X may not be easily understood, but the following theorem clarifies its meaning in terms of the finite-dimensional distributions of the process. For processes X and Y and $S \subset \mathbb{T}$, we write $X \overset{\text{fd}(S)}{=} Y$ if $(X_{t_1},\dots,X_{t_k})$ and $(Y_{t_1},\dots,Y_{t_k})$ have the same law in $\mathbb{E}^k$ for all $t_i \in S$, $1 \le i \le k$, $k \in \mathbb{N}$.

THEOREM 2.3. *Let X and Y be $\mathbb{E}$-valued measurable processes, and suppose that $\mathscr{T}$ is countably generated up to ν-null sets. Then X and Y induce the same laws on $\widetilde{\mathbb{M}}_{\mathbb{E}}(\lambda)$ if and only if there exists an $S \in \mathscr{T}$ with $\lambda(S^{\complement}) = 0$ such that $X \overset{\text{fd}(S)}{=} Y$.*

Proof. First we prove the *if* part. Let $\mathscr{G} = \{g \in b(\mathscr{T}\otimes\mathscr{E})\colon \{g(t,\cdot)\colon t \in \mathbb{T}\}$ is uniformly equicontinuous on $\mathbb{E}\}$ and

$$\Phi_{B,g}(w) \triangleq \int_B g(t,w(t))\,\lambda(dt),\ w \in \mathbb{M}_{\mathbb{E}},\ B \in \mathscr{T} \text{ and } g \in \mathscr{G}.$$

We have $\Phi_{B,g}(w) \in C_b(\widetilde{\mathbb{M}}_{\mathbb{E}}))$. One can easily prove that the subalgebra $\mathscr{A}$ in $C_b(\widetilde{\mathbb{M}}_{\mathbb{E}})$ generated by 1 and the $\Phi_{B,g}$, $B \in \mathscr{T}$, $g \in \mathscr{G}$, separates measures on $\widetilde{\mathbb{M}}_{\mathbb{E}}$. Then by Fubini's theorem, for each $\Phi \in \mathscr{A}$, we have $\int \Phi(X)dP = \int \Phi(Y)dP$.

To show the *only if* part, we will use the function G constructed in the proof of Lemma 2.2. Denote as earlier by $\widetilde{X}(\omega)$ and $\widetilde{Y}(\omega)$ the equivalence classes containing $X_\bullet(\omega)$ and $Y_\bullet(\omega)$, respectively. Then for each ω, $G(\widetilde{X}(\omega),\bullet) = X_\bullet(\omega)$, λ-a.e., so by an application of Fubini's theorem there is an $S_1 \in \mathscr{T}$ with $\lambda(S_1^{\complement}) = 0$ such that $X_t(\omega) = G(\widetilde{X}(\omega),t)$, ω-a.s. for all $t \in S_1$. Similarly we can find such a set S_2 for Y. Put $S = S_1 \cap S_2$, so we have $\lambda(S^{\complement}) = 0$. And for all $t \in S$, $X_t(\omega) = G(\widetilde{X}(\omega),t)$, ω-a.s. and $Y_t(\omega) = G(\widetilde{Y}(\omega),t)$, ω-a.s. By the assumption, $\widetilde{X}$ and $\widetilde{Y}$ have the same law on $\mathbb{M}_{\mathbb{E}}$, so for any $t_1,\dots,t_m \in S$,

$$\Big(G(\widetilde{X}(\omega),t_1),\dots,G(\widetilde{X}(\omega),t_m)\Big) \overset{\mathscr{L}}{=} \Big(G(\widetilde{Y}(\omega),t_1),\dots,G(\widetilde{Y}(\omega),t_m)\Big).$$

It therefore follows that $(X_{t_1},\dots,X_{t_m})$ and $(Y_{t_1},\dots,Y_{t_m})$ have the same law on $\mathbb{E}^m$ for any $t_1,\dots,t_m \in S$. □

The above theorem connects the notion of probability measure on $\mathbb{M}_{\mathbb{E}}(\lambda)$ with that of almost equivalence. Let X and Y be two $\mathbb{E}$-valued measurable processes defined on possibly different probability spaces. Then X and Y are said to be *ν-almost equivalent* if for every finite system of pairs (ϕ_i, g_i), $1 \leq i \leq n$, where ϕ_i is a positive and integrable function on $\mathbb{T}$ and g_i is a bounded $\mathscr{E}/\mathscr{B}(\mathbb{R})$ measurable function on $\mathbb{E}$, the random vectors

$$\left(\int_{\mathbb{T}} g_1(X_t)\phi_1(t)\,\nu(dt), \dots, \int_{\mathbb{T}} g_n(X_t)\phi_n(t)\,\nu(dt)\right)$$

and

$$\left(\int_{\mathbb{T}} g_1(Y_t)\phi_1(t)\,\nu(dt), \dots, \int_{\mathbb{T}} g_n(Y_t)\phi_n(t)\,\nu(dt)\right)$$

have the same law on $\mathbb{R}^n$. It is then easy to see from the proof of Theorem 2.3 that X and Y induce the same law on $\mathbb{M}_{\mathbb{E}}(\lambda)$ if and only if X and Y are ν-almost equivalent.

The concepts of pseudo-path and pseudo-law have been introduced in Dellacherie and Meyer [4], IV.35-45 for the case $\mathbb{T} = \mathbb{R}_+$, and they are closely related to the notions of equivalence classes and probability laws on $\mathbb{M}_{\mathbb{E}}(\lambda)$. They could be extended to a general parameter set $\mathbb{T}$ with a σ-field $\mathscr{T}$ and a σ-finite measure ν, although the choice of ν, which should play the canonical role, may not be evident.

Now we turn to the convergence in law of measurable processes. Let $X = (X_t)_{t\in\mathbb{T}}$ and $X^n = (X^n_t)_{t\in\mathbb{T}}$, $n \in \mathbb{N}$, be measurable processes on $(\Omega, \mathscr{F}, P)$ with state space $(\mathbb{E}, \mathscr{E})$ and parameter space $(\mathbb{T}, \mathscr{T}, \nu)$. It appears that we would need to show $X^n \xrightarrow{\text{fd}(S)} X$ for some $S \in \mathscr{T}$ with $\lambda(S^{\complement}) = 0$ and the tightness of $(X^n)_{n\in\mathbb{N}}$ in order to get $X^n \xrightarrow{\mathscr{L}} X$ in $\mathbb{M}_{\mathbb{E}}(\lambda)$. But the next theorem shows that in fact it suffices to prove $X^n \xrightarrow{\text{fd}(S)} X$ for some $S \in \mathscr{T}$ with $\lambda(S^{\complement}) = 0$. That is, tightness is unnecessary, although we must know that the limiting process X is measurable. One can prove the following theorem in a fashion similar to Cremers and Kadelka [2].

THEOREM 2.4. *Let $\mathbb{E}$ be a separable metrizable space and $\mathscr{T}$ be countably generated. Suppose that $(X^n)_{n\in\mathbb{N}}$ and X are $\mathbb{E}$-valued measurable processes on $(\Omega, \mathscr{F}, P)$. If $X^n \xrightarrow{\text{fd}(S)} X$ for some $S \in \mathscr{T}$ with $\lambda(S^{\complement}) = 0$, then $X^n \xrightarrow{\mathscr{L}} X$ in $\mathbb{M}_{\mathbb{E}}(\lambda)$.*

The converse of the above theorem in a sense also holds; the following results are straightforward extensions of Meyer and Zheng [9] and Sadi [11].

THEOREM 2.5. *Suppose that $\mathscr{T}$ is countably generated up to ν-null sets and that $\mathbb{E}$ is a metrizable Lusin space. Let $(X^n)_{n\in\mathbb{N}}$ and X be $\mathbb{E}$-valued measurable processes with parameter set $\mathbb{T}$ on some $(\Omega, \mathscr{F}, P)$. If $X^n \xrightarrow{\mathscr{L}} X$ in*

$\mathbb{M}_{\mathbb{E}}$, *then there exist an* $S \in \mathscr{T}$ *with* $\lambda(S^{\complement}) = 0$ *and a subsequence* (n') *such that* $X^{n'} \xrightarrow{\text{fd}(S)} X$.

Corollary 2.6. *Let* $(X^n)_{n\in\mathbb{N}}$ *and* X *be* $\mathbb{E}$*-valued measurable processes with parameter set* $\mathbb{T}$ *on some* $(\Omega, \mathscr{F}, P)$. *For* $X^n \xrightarrow{\mathscr{L}} X$ *in* $\mathbb{M}_{\mathbb{E}}$, *it is necessary and sufficient that for any subsequence* (n'), *there exist a further subsequence* (n'') *and an* $S \in \mathscr{T}$ *with* $\lambda(S^{\complement}) = 0$ *such that* $X^{n''} \xrightarrow{\text{fd}(S)} X$.

3. Convergence of the prediction processes

Let $X = (X_t)_{t\in\mathbb{R}_+}$ be a measurable process with values in $\mathbb{E}$. We assume that $\mathbb{E}$ is a metrizable Lusin space with $\mathscr{E} = \mathscr{B}(\mathbb{E})$, and let $\mathbb{M}_{\mathbb{E}} = \mathbb{M}_{\mathbb{E}}(\mathbb{R}_+, \mathscr{B}_+, m)$, where $\mathscr{B}_+ = \mathscr{B}(\mathbb{R}_+)$ and m is the Lebesgue measure. Setting $\lambda(dt) = e^{-t}dt$, we give $\mathbb{M}_{\mathbb{E}}$ the topology of convergence in λ-measure.

We define the *pseudo-path* filtration $\mathscr{F}'_t$ by

$$\mathscr{F}'_t = \sigma\left(\int_0^s f(w(u))\,du; s < t,\ f \in b\mathscr{E}\right),$$

and set $\mathscr{F}' \triangleq \mathscr{F}'_\infty = \bigvee_{t>0} \mathscr{F}'_t$. It is obvious that $\mathscr{F}' = \mathscr{B}(\mathbb{M}_{\mathbb{E}})$. Moreover, the shift operator θ_t on $\mathbb{M}_{\mathbb{E}}$ is defined by $\theta_t w(s) = w(t+s)$ for $s, t \in \mathbb{R}_+$ and is $\mathscr{F}'_{t+s}/\mathscr{F}'_s$ measurable. For the state space of the prediction process, let $\Pi \triangleq \mathscr{P}(\mathbb{M}_{\mathbb{E}})$, the set of probability measures on $(\mathbb{M}_{\mathbb{E}}, \mathscr{F}')$ endowed with the topology of weak convergence. This topology is called the *prediction topology* in Knight [7], with which Π becomes a metrizable Lusin space. We set $\mathscr{G} = \mathscr{B}(\Pi)$. A generic element of Π is usually denoted by z, and we sometimes write P^z for z; it is actually redundant but intuitively helpful.

According to Corollary 2.5 of Knight [7], the prediction process $Z^z = (Z^z_t)_{t\in\mathbb{R}_+}$ for $z \in \Pi$ is the process with values in $(\Pi, \mathscr{G})$ that is P^z-a.s. uniquely determined by the following two requirements:

(1) $Z^z_r(A) = P^z(\,\theta_r^{-1}A \mid \mathscr{F}'_{r+}\,)$, $r \in \mathbb{Q}_+$, $A \in \mathscr{F}'$;

(2) Z^z_t is càdlàg for the prediction topology on Π defined above.

Thus the prediction process is defined for the law $z \in \Pi$ induced by X rather than for the process X itself. In terms of the generalized coordinate process $\widetilde{X}$ on $\mathbb{M}_{\mathbb{E}}$ defined by $\widetilde{X}_t(w) \triangleq G(\widetilde{w}, t)$, where $\widetilde{w}$ is the equivalence class containing w and G is the function in Lemma 2.1, (i) may then be written as

$$Z^z_r(A) = P^z(\,\widetilde{X}_{t+\bullet} \in A \mid \mathscr{F}'_{r+}\,).$$

3.1. Convergence of the given processes. We shall now show that the convergence of the prediction processes implies that of the given processes. Let X^n, $n \in \mathbb{N}$, and X be measurable processes with values in $\mathbb{E}$ and with laws z_n and z, and let Z^n and Z be the prediction processes of X^n and X, respectively. Since $\mathbb{M}_{\mathbb{E}}$ is a metrizable Lusin space, so is Π (Dellacherie and Meyer [4], III.60). Then the path space $\mathbb{M}_\Pi = \mathbb{M}_\Pi(\mathbb{R}_+, \mathscr{B}_+, \lambda)$ of the prediction process is also a metrizable Lusin space.

THEOREM 3.1. *If $(Z^n)_{n\in\mathbb{N}}$ converges in law to Z in $\mathbb{M}_\Pi$, then $(X^n)_{n\in\mathbb{N}}$ converges in law to X in $\mathbb{M}_\mathbb{E}$.*

Proof. By Corollary 2.6, for any subsequence (n'), there is a further subsequence (n'') and an $S \subset \mathbb{R}_+$ with $\lambda(S^\complement) = 0$ such that $Z^{n''} \xrightarrow{\mathrm{fd}(S)} Z$. Also for any $\phi \in C_b(\mathbb{M}_\mathbb{E})$, the mapping $\mu \mapsto \mu(\phi) = \int \phi\, d\mu$ from Π into $\mathbb{R}$ is continuous. It thus follows that for any $t \in S$ and any $\phi \in C_b(\mathbb{M}_\mathbb{E})$,

$$E^{z_{n''}}[\, \phi(\widetilde{X}_{t+\bullet}) \mid \mathscr{F}'_{t+}\,] \xrightarrow{\mathscr{L}} E^{z}[\, \phi(\widetilde{X}_{t+\bullet}) \mid \mathscr{F}'_{t+}\,] \text{ in } \mathbb{R}.$$

The sequence $(E^{z_{n''}}[\, \phi(\widetilde{X}_{t+\bullet}) \mid \mathscr{F}'_{t+}\,])$ is bounded, so it is uniformly integrable. Hence we get $E^{z_{n''}}[\phi(\widetilde{X}_{t+\bullet})] \to E^{z}[\phi(\widetilde{X}_{t+\bullet})]$ for any $t \in S$ and any $\phi \in C_b(\mathbb{M}_\mathbb{E})$. In other words, $X^{n''}_{t+\bullet} \xrightarrow{\mathscr{L}} X_{t+\bullet}$ in $\mathbb{M}_\mathbb{E}$ for each $t \in S$. Using the translation operator θ_t, we may write this as $\theta_t X^{n''} \xrightarrow{\mathscr{L}} \theta_t X$ in $\mathbb{M}_\mathbb{E}$ for each $t \in S$.

It is clear that θ_t is continuous since if $w_n \to w$ in λ-measure, then $w_n(t + \bullet) \to w(t+\bullet)$ in λ-measure. Set $\mathscr{A} = \{\theta_t^{-1} G \colon G \text{ open in } \mathbb{M}_\mathbb{E},\ t \in S\}$. Then it is straightforward to verify that $\mathscr{A}$ is a base for the topology for $\mathbb{M}_\mathbb{E}$ and that $\mathscr{A}$ is closed under finite unions. It hence follows that the family $\{\theta_t\}_{t\in S}$ satisfies the conditions of Pollard's theorem (Pollard(1977)). Consequently we obtain $X^{n''} \xrightarrow{\mathscr{L}} X$ in $\mathbb{M}_\mathbb{E}$. We have shown that for any subsequence (n'), there exists a further subsequence (n'') for which $X^{n''} \xrightarrow{\mathscr{L}} X$ in $\mathbb{M}_\mathbb{E}$, which obviously implies that the sequence (X^n) of the given processes converges in law to X in $\mathbb{M}_\mathbb{E}$. □

REMARK 3.2. The assertion of Theorem 3.1 is in fact equivalent to that of Lemma 2.21 (1) of Knight [7]. Our proof here is different from his, and the point of our proof is that the result can be shown without using the Markov property of the prediction process; only the defining property of the prediction process is necessary.

3.2. Markovian case. In this subsection, we assume that $\mathbb{E}$ is Polish. Let $p(t, x, B)$, $t \in \mathbb{R}_+$, $x \in \mathbb{E}$, $B \in \mathscr{E}$, be a Markov transition function which satisfies

(3.1) $(t, x) \mapsto p(t, x, B)$ is $\mathscr{B}(0,\infty) \otimes \mathscr{E}$ measurable for each $B \in \mathscr{E}$;
(3.2) $\{x \mapsto p(t, x, B) \colon t > 0,\ B \in \mathscr{E}\}$ separates points of $\mathbb{E}$.

From Lemma 2.8 in Knight [7], for each $x \in \mathbb{E}$, there is a measurable Markov process $X = (X_t)_{t\in\mathbb{R}_+}$ with finite-dimensional distributions determined by

$p(t,x,B)$. Namely, for $0 \le t_1 < \cdots < t_k$, we have

$$(3.3)\quad P^x(X_{t_1} \in B_1, \ldots, X_{t_k} \in B_k) = \int_{B_{k-1}} \cdots \int_{B_1} p(t_1, x, dx_1) \cdots p(t_{k-1} - t_{k-2}, x_{k-2}, dx_{k-1}) \cdot \cdot\, p(t_k - t_{k-1}, x_{k-1}, B_k).$$

Note that we do not assume $p(0,x,\bullet) = \delta_x(\bullet)$. Thus the process X may not start at x under P^x. We call the above process X the measurable Markov process having the P^x-law with transition function $p(t,x,B)$. This process induces a law on $\mathbb{M}_{\mathbb{E}}$, which we denote by $\varphi(x)$. This is uniquely determined by $p(t,x,B)$ (see Knight [7], p. 53). We look at φ as a mapping from $\mathbb{E}$ into Π. Lemma 2.9 of Knight [7] shows that φ is $\mathscr{E}/\mathscr{B}(\Pi)$ measurable. Furthermore, Theorem 2.36 of Knight [7] states that $\sigma(x \mapsto p(t,x,B)\colon t > 0,\ B \in \mathscr{E}) = \sigma(x \mapsto R_\lambda f(x)\colon \lambda > 0,\ f \in b\mathscr{E})$, where R_λ is the resolvent defined by

$$R_\lambda f(x) \triangleq \int_0^\infty e^{-\lambda t} T_t f(x)\, dt$$

and $T_t f(x) \triangleq \int p(t,x,dy) f(y)$ is the semigroup associated with $p(t,x,B)$. Thus (3.2) amounts to assuming that $\{x \mapsto R_\lambda f(x)\colon \lambda > 0,\ f \in b\mathscr{E}\}$ separates points of $\mathbb{E}$. It is clear that φ is 1-1. The key result is Theorem 2.10 of Knight [7], which says that for each $x \in \mathbb{E}$, we have

$$P^{\varphi(x)}\left[\varphi(\widetilde{X}_t) = Z_t^{\varphi(x)} \text{ for a.e. } t\right] = 1;$$

that is, the process $\varphi(X) = (\varphi(X_t))_{t \in \mathbb{R}_+}$ and $Z^{\varphi(x)}$ induce the same law on $\mathbb{M}_\Pi$.

Now consider a sequence of Markov transition functions $(p_n(t,x,B))_{n \in \mathbb{N}}$ and $p(t,x,B)$ satisfying (3.1) and (3.2) above, and denote by X^n, $n \in \mathbb{N}$, and X the measurable Markov processes having the P^x-law with $p_n(t,x,B)$, $n \in \mathbb{N}$, and $p(t,x,B)$, respectively. Let $\varphi^n(x)$ and $\varphi(x)$ be the laws on $\mathbb{M}_{\mathbb{E}}$ induced by X^n and X with (3.3), as defined above, and put $Z^n = Z^{\varphi^n(x)}$ and $Z = Z^{\varphi(x)}$. Our problem is to find under which conditions on $(p_n(t,x,B))$, $X^n \xrightarrow{\mathscr{L}} X$ in $\mathbb{M}_{\mathbb{E}}$ implies $Z^n \xrightarrow{\mathscr{L}} Z$ in $\mathbb{M}_\Pi$ (note that the dependence on x is suppressed here). More precisely, if $X^n \xrightarrow{\mathscr{L}} X$ in $\mathbb{M}_{\mathbb{E}}$ for each $x \in \mathbb{E}$, then what additional conditions are necessary for Z^n to converge in law to Z in $\mathbb{M}_\Pi$ for each x? The assumption amounts to $\varphi^n(x) \to \varphi(x)$ for each $x \in \mathbb{E}$. From the above observation, we know that $\varphi^n(X^n)$ and Z^n induce the same distribution on $\mathbb{M}_\Pi$, so what we need is $\varphi^n(X^n) \xrightarrow{\mathscr{L}} \varphi(X)$ in $\mathbb{M}_\Pi$ for each $x \in \mathbb{E}$. Thus the problem is reduced to a familiar one of preservation of convergence in law under mappings. This is discussed in Section 5 of Billingsley [1] and a necessary and sufficient condition was obtained in Topsøe [13]. Here we use a simple condition, which is an easy consequence of H. Rubin's theorem (Billingsley [1], Theorem 5.5).

PROPOSITION 3.3. *Let X^n and X be S-valued random variables with S a separable metric space, and let h_n and h be measurable mappings from S into a metric space S'. If h is continuous and if h_n converges to h uniformly on compact sets, then $X^n \xrightarrow{\mathscr{L}} X$ implies $h_n(X^n) \xrightarrow{\mathscr{L}} h(X)$.*

To apply this proposition to our problem, let us define $\Phi\colon \mathbb{M}_{\mathbb{E}} \to \mathbb{M}_{\Pi}$ by $\Phi(w)(t) = \varphi(w(t))$, $w \in \mathbb{M}_{\mathbb{E}}$, and similarly define Φ^n. The processes $(\varphi^n(X_t^n))_{t\in\mathbb{R}_+}$, $n \in \mathbb{N}$, are then written as $\Phi^n(X^n)$, $n \in \mathbb{N}$. It is clear that if φ is continuous, so is Φ. Denote the metrics on Π and $\mathbb{M}_{\Pi}$ by d' and ρ' so that we have

$$\rho'\left(\Phi^n(x), \Phi(x)\right) = \int_0^\infty 1 \wedge d'\left(\varphi^n(w(t)), \varphi(w(t))\right)\, \lambda(dt).$$

Let Γ be a compact subset on $\mathbb{M}_{\mathbb{E}}$ and let $\varepsilon > 0$ be given. Choose $T > 0$ satisfying $\lambda(T, \infty) < \varepsilon$. It follows from Kurtz [8], Theorem 4.1 (this is the only place where the Polish assumption is used), that we can find a compact $K \subset \mathbb{E}$ such that $\sup_{w\in\Gamma} \lambda(t \le T\colon w(t) \notin K) \le \varepsilon$. Now assume for the moment that $\varphi^n \to \varphi$ uniformly on compact sets. Then

$$\begin{aligned}\rho'\left(\Phi^n(w), \Phi(w)\right) &\le \int_{[0,T]\cap\{w(t)\in K\}} 1 \wedge d'\left(\varphi^n(w(t)), \varphi(w(t))\right)\, \lambda(dt) \\ &\qquad + \int_{[0,T]\cap\{w(t)\notin K\}} 1 \wedge d'\left(\varphi^n(w(t)), \varphi(w(t))\right)\, \lambda(dt) + \varepsilon \\ &\le \int_{[0,T]\cap\{w(t)\in K\}} 1 \wedge d'\left(\varphi^n(w(t)), \varphi(w(t))\right)\, \lambda(dt) + 2\varepsilon.\end{aligned}$$

The integral converges to 0 uniformly in $w \in \Gamma$ by the bounded convergence theorem. Hence, as $n \to \infty$,

$$\sup_{w\in\Gamma} \rho'\left(\Phi^n(w), \Phi(w)\right) \to 0.$$

In view of Proposition 3.2, $Z^n \xrightarrow{\mathscr{L}} Z$ will follow.

We would like to express the assumed compact convergence of φ^n to φ in terms of resolvents R_λ^n and R_λ of X^n and X, respectively. First, by Lemma 2.15 of Knight [7], φ is continuous if and only if $R_\lambda f$ is continuous on $\mathbb{E}$ for $f \in C_b(\mathbb{E})$. Compact convergence of φ^n to φ means

$$\sup_{x\in K} \left| E^{\varphi^n(x)}(g) - E^{\varphi(x)}(g) \right| \to 0$$

for each compact $K \subset \mathbb{E}$ and each $g \in C_b(\mathbb{M}_{\mathbb{E}})$. We may replace g by a member of a convergence determining class. We use the following class:

$$\left\{ \prod_{k=1}^{m} \int_0^\infty e^{-\lambda_k s} f_{j_k}(w(s))\, ds \colon m \in \mathbb{N},\ \lambda_k \in \mathbb{Q}_+,\ f_{j_k} \in \{f_j\},\ 1 \le k \le m \right\},$$

where (f_j) is dense in $C(\widehat{\mathbb{E}}) \cap \{f \colon \widehat{\mathbb{E}} \to [0,1]\}$. So we need to find a condition for

$$E^{\varphi^n(x)}\left[\prod_{k=1}^{m} \int_0^\infty e^{-\lambda_k t} f_{j_k}(\widetilde{X}_t)\, dt\right] \to E^{\varphi(x)}\left[\prod_{k=1}^{m} \int_0^\infty e^{-\lambda_k t} f_{j_k}(\widetilde{X}_t)\, dt\right] \tag{3.4}$$

uniformly in $x \in K$ for a compact K and $m \in \mathbb{N}$, $\lambda_k \in \mathbb{Q}_+$ and $f_{j_k} \in (f_j)$. Let us look at the case $m = 1$. The left-hand side is

$$E^{\varphi^n(x)}\left[\int_0^\infty e^{-\lambda t} f_{j_k}(\widetilde{X}_t)\, dt\right] = R^n_\lambda f_j(x).$$

Hence we need the uniform convergence on compact sets of the resolvents, that is,

$$R^n_\lambda f(x) \to R_\lambda f(x) \quad \text{uniformly in } x \in K \tag{3.5}$$

for each λ and $f \in C(\widehat{\mathbb{E}})$. For a general m, we use the argument given in the proof of Lemma 2.15 of Knight [7]. Write the left-hand side of (3.4) as

$$E^{\varphi^n(x)}\left[\int_0^\infty \cdots \int_0^\infty e^{-\sum_{k=1}^m \lambda_k s_k} f_{j_1}(\widetilde{X}_{s_1}) \cdots f_{j_m}(\widetilde{X}_{s_m})\, ds_1 \cdots ds_m\right].$$

Express this multiple integral as a sum of $m!$ integrals corresponding to the $m!$ possible orderings of $s_1, \ldots, s_m$. Then it is enough to look at, for instance, the case $s_1 < \cdots < s_m$:

$$E^{\varphi^n(x)}\left[\int_0^\infty \int_{s_1}^\infty \cdots \int_{s_{m-1}}^\infty e^{-\sum_{k=1}^m \lambda_k s_k} f_{j_1}(\widetilde{X}_{s_1}) \cdots f_{j_m}(\widetilde{X}_{s_m})\, ds_1 \cdots ds_m\right]. \tag{3.6}$$

Using the Markov property, this is equal to

$$\begin{aligned}
&\int_0^\infty e^{-\lambda_1 s_1} E^{\varphi^n(x)}\Bigg[f_{j_1}(\widetilde{X}_{s_1}) E^{\varphi^n(x)}\Bigg(\int_{s_1}^\infty e^{-\lambda_2 s_2} f_{j_2}(\widetilde{X}_{s_2}) \\
&\qquad \int_{s_2}^\infty \cdots \int_{s_{m-1}}^\infty e^{-\lambda_m s_m} f_{j_m}(\widetilde{X}_{s_m})\, ds_m \cdots ds_2 \Bigg| \mathscr{F}'_{s_1+}\Bigg)\Bigg] ds_1 \\
&= \int_0^\infty e^{-\lambda_1 s_1} E^{\varphi^n(x)}\Bigg[f_{j_1}(\widetilde{X}_{s_1}) E^{\varphi^n(x)}\Bigg(\int_0^\infty e^{-\lambda_2 (t_2+s_1)} f_{j_2}(\widetilde{X}_{t_2+s_1}) \\
&\qquad \int_{t_2}^\infty \cdots \int_{t_{m-1}}^\infty e^{-\lambda_m (t_m+s_1)} f_{j_m}(\widetilde{X}_{t_m+s_1})\, dt_m \cdots dt_2 \Bigg| \mathscr{F}'_{s_1+}\Bigg)\Bigg]
\end{aligned}$$

$$= \int_0^\infty e^{-\lambda_1 s_1} E^{\varphi^n(x)} \Bigg[f_{j_1}(\widetilde{X}_{s_1}) e^{-(\lambda_2+\cdots+\lambda_m)s_1} E^{\varphi^n(\widetilde{X}_{s_1})} \bigg(\int_0^\infty e^{-\lambda_2 t_2} f_{j_2}(\widetilde{X}_{t_2}) \int_{t_2}^\infty \cdots \int_{t_{m-1}}^\infty e^{-\lambda_m t_m} f_{j_m}(\widetilde{X}_{t_m})\, dt_m \cdots dt_2 \bigg) \Bigg] ds_1$$

$$= E^{\varphi^n(x)} \left[\int_0^\infty e^{-\overline{\lambda} s_1} f_{j_1}(\widetilde{X}_{s_1}) g^n(\widetilde{X}_{s_1})\, ds_1 \right],$$

where $\overline{\lambda} \triangleq \lambda_1 + \cdots + \lambda_m$ and

$$g^n(x) \triangleq E^{\varphi^n(x)} \left[\int_0^\infty e^{-\lambda_2 t_2} f_{j_2}(\widetilde{X}_{t_2}) \int_{t_2}^\infty \cdots \int_{t_{m-1}}^\infty e^{-\lambda_m t_m} f_{j_m}(\widetilde{X}_{t_m})\, dt_m \cdots dt_2 \right].$$

Note that g^n is of the form (3.6) with $m-1$ in place of m. Thus if we assume that (3.4) holds for $m-1$ as the induction hypothesis, $g^n(s)$ will converge to $g(x)$, defined similarly, uniformly in $x \in K$. Writing $h^n(x) = f_{j_1}(x) g^n(x)$, the above expectation is equal to $R^n_{\overline{\lambda}} h^n(x)$. Assuming the induction hypothesis, $h^n(x)$ converges to $h(x) = f_{j_1}(x) g(x)$ uniformly in $x \in K$. So the condition we need is the following:

> $R^n_\lambda h^n(x) \to R_\lambda h(x)$ uniformly in $x \in K$ for each $\lambda > 0$, whenever $h^n \to h$ uniformly on compact sets.

As is seen by the above argument, the sequence (h^n) may be restricted to be uniformly bounded and we may assume that h is continuous and bounded.

We have therefore obtained the following theorem.

THEOREM 3.4. *In the setting of this subsection, suppose that $R^n_\lambda h^n$ converges to $R_\lambda h$ uniformly on compact sets for each $\lambda > 0$ whenever a uniformly bounded sequence (h^n) converges to a continuous bounded h uniformly on compact sets, and that $R_\lambda f$ is continuous for $f \in C_b(\mathbb{E})$. Then Z^n converges in law to Z in $\mathbb{M}_\Pi$ for each $x \in \mathbb{E}$.*

REFERENCES

[1] P. Billingsley, *Convergence of probability measures*, John Wiley & Sons Inc., New York, 1968. MR 0233396 (38 #1718)

[2] H. Cremers and D. Kadelka, *On weak convergence of integral functionals of stochastic processes with applications to processes taking paths in L_p^E*, Stochastic Process. Appl. **21** (1986), 305–317. MR 833957 (87h:60071)

[3] C. Dellacherie, *Convergence en probabilité et topologie de Baxter-Chacon*, Séminaire de Probabilités, XII (Univ. Strasbourg, Strasbourg, 1976/1977), Lecture Notes in Math., vol. 649, Springer, Berlin, 1978, p. p. 424. MR 520016 (81m:60010)

[4] C. Dellacherie and P.-A. Meyer, *Probabilités et potentiel*, Hermann, Paris, 1975, Chapitres I à IV, Édition entièrement refondue, Publications de l'Institut de Mathématique de l'Université de Strasbourg, No. XV, Actualités Scientifiques et Industrielles, No. 1372. MR 0488194 (58 #7757)

[5] F. B. Knight, *A predictive view of continuous time processes*, Ann. Probability **3** (1975), 573–596. MR 0383513 (52 #4394)

[6] ______, *Essays on the prediction process*, Institute of Mathematical Statistics Lecture Notes-Monograph Series, 1, Institute of Mathematical Statistics, Hayward, Calif., 1981. MR 719690 (84m:60044)

[7] ______, *Foundations of the prediction process*, Oxford Studies in Probability, vol. 1, The Clarendon Press Oxford University Press, New York, 1992, Oxford Science Publications. MR 1168699 (94b:60047)

[8] T. G. Kurtz, *Random time changes and convergence in distribution under the Meyer-Zheng conditions*, Ann. Probab. **19** (1991), 1010–1034. MR 1112405 (92m:60033)

[9] P.-A. Meyer and W. A. Zheng, *Tightness criteria for laws of semimartingales*, Ann. Inst. H. Poincaré Probab. Statist. **20** (1984), 353–372. MR 771895 (86c:60008)

[10] D. Pollard, *Induced weak convergence and random measures*, Z. Wahrsch. Verw. Gebiete **37** (1976/77), 321–328. MR 571672 (81j:60008)

[11] H. Sadi, *Une condition nécessaire et suffisante pour la convergence en pseudo-loi des processus*, Séminaire de Probabilités, XXII, Lecture Notes in Math., vol. 1321, Springer, Berlin, 1988, pp. 434–437. MR 960538 (89i:60010)

[12] A. V. Skorohod, *Measurability of stochastic processes*, Teor. Veroyatnost. i Primenen. **25** (1980), 140–142. MR 560064 (82c:60062)

[13] F. Topsøe, *Preservation of weak convergence under mappings*, Ann. Math. Statist **38** (1967), 1661–1665. MR 0219097 (36 #2180)

[14] H. Tsukahara, *Weak convergence and the prediction process*, Ph.D Thesis, Department of Statistics, University of Illinois at Urbana-Champaign, 1996.

Hideatsu Tsukahara, Department of Economics, Seijo University, 6–1–20 Seijo, Setagaya-ku, Tokyo 157–8511, Japan

E-mail address: tsukahar@seijo.ac.jp