J. Math. Sci. Univ. Tokyo
8 (2001), 343–356.

On a Certain Metric on the Space of Pairs of a Random Variable and a Probability Measure

By Shigeo Kusuoka and Toshiyuki Nakayama

Abstract. The authors introduce new metrics on the space of pairs of a random variable and a probability measure over a Polish space, and study the properties of them. Finally, as an application, they discuss how the integrand for the martingale representation converges in the invariance principle.

1. Introduction

Let $p \in [1, \infty)$, and M be a Polish space. $\mathcal{P}(M)$ denotes the set of all probability measures on a Polish space M. For a separable metric space N, $\mathcal{X}_{M;N}$ denotes the set of all pairs (X, μ) for which X is a measurable map from M into N and $\mu \in \mathcal{P}(M)$. $\mathcal{X}^p_{M;N}$ denotes the set of $(X, \mu) \in \mathcal{X}_{M;N}$ such that $\int_M dis_N(X(x), y)^p \mu(dx) < \infty$ for all $y \in N$.

Let $Dis^{(p)}_{M;N} : \mathcal{X}^p_{M;N} \times \mathcal{X}^p_{M;N} \to [0, \infty)$ be given by

$$\begin{aligned} & Dis^{(p)}_{M;N}((X_1, \mu_1), (X_2, \mu_2)) \\ = \; & \inf\{(\int_{M\times M} ((dis_M(x_1, x_2) \wedge 1) \\ & \qquad + dis_N(X_1(x_1), X_2(x_2)))^p \nu(dx_1, dx_2))^{1/p}; \\ & \qquad \nu \in \mathcal{P}(M \times M),\ \nu \circ \pi_1^{-1} = \mu_1,\ \nu \circ \pi_2^{-1} = \mu_2\}. \end{aligned}$$

Here dis_M, dis_N are distance functions on M, N respectively, and $\pi_i : M \times M \to M$, $i = 1, 2$, are canonical projections given by $\pi_1(x_1, x_2) = x_1$, $\pi_2(x_1, x_2) = x_2$, $x_1, x_2 \in M$.

Definition. Let $(X_n, \mu_n), (X, \mu) \in \mathcal{X}^p_{M;N}, n \geq 1$. We say that $(X_n, \mu_n) \to (X, \mu)$ in $\mathcal{X}^p_{M;N}$ if $Dis^{(p)}_{M;N}((X_n, \mu_n), (X, \mu)) \to 0$.

1991 *Mathematics Subject Classification.* Preliminary 60B10; Secondary 60B5.

We will study some properties of this convergence. One of our main results is the following Skorohod's type theorem.

Theorem 1. *Suppose that N is an arcwise connected separable metric space. Let (X_n,μ_n), $(X,\mu) \in \mathcal{X}_{M;N}$, $n \geq 1$, and suppose that there are a probability space $(\Omega,\mathcal{F},P)$ and M-valued random variables Z_n, $n \geq 1$, and Z such that*
(1) $P \circ Z_n^{-1} = \mu_n, n \geq 1,$ *and* $P \circ Z^{-1} = \mu,$
and
(2) $Z_n \to Z$ *in probability.*
Then $(X_n,\mu_n) \to (X,\mu)$ *in* $\mathcal{X}^p_{M;N}$, *if and only if* $dis_N(X_n \circ Z_n, X \circ Z) \to 0$ *in* $L^p(\Omega,\mathcal{F},P)$.

Also we discuss a special case that μ_n is the probability law of random walk and μ is the Wiener measure and show a result Theorem 11 on Martingale representations. This result may be useful in mathematical finance by the following reason. We sometimes use discrete space-time models to approximate diffusion process models. There are many works which guarantee that such approximations are good concerning the option prices. But concerning hedging strategies, it is not well studied whether such approximations make sense. The convergence notion given here will give a certain base for this question. An application for the approximation of backward SDE will be given in Nakayama [3].

2. Basic Results

Let M be a Polish space and N be a separable metric space.

Let $C_b(M;N)$ denotes the set of continuous maps $f : M \to N$ such that

$$\sup_{x\in M} dis_N(f(x),y) < \infty, \qquad y \in N.$$

Let $\mu \in \mathcal{P}(M)$. For $p \in [1,\infty)$, let $\tilde{L}^p(M;N,d\mu)$ denotes the set of measurable maps $f : M \to N$ such that

$$\int_M dis_N(f(x),y)^p \mu(dx) < \infty, \qquad \text{for all } y \in N\ .$$

Then we have the following.

THEOREM 2. *Let $p \in [1,\infty)$, and N be an arcwise connected separable metric space. Then for any $f \in \tilde{L}^p(M;N,d\mu)$ and $\varepsilon > 0$, there is a $g \in C_b(M;N)$ such that*

$$(\int_M dis_N(f(x),g(x))^p \mu(dx))^{1/p} \le \varepsilon.$$

To prove Theorem 2, we need some preparations.

For each $n \ge 1$, let $\tilde{E}_n = [0,1] \times \{1,2,\dots,n\}$. We define a quasi-metric function $d : \tilde{E}_n \times \tilde{E}_n$ by

$$d((s,i),(t,j)) = \begin{cases} |t-s|, & i = j, \\ t+s, & i \ne j \end{cases}$$

We define an equivalence relation $\sim$ by $(s,i) \sim (t,j)$ if $d((s,i),(t,j)) = 0$, that is $(s,i) \sim (t,j)$ if $(s,i) = (t,j)$ or $s = t = 0$. Let $E_n = \tilde{E}_n/\sim$. Then (E_n,d) is a Poilsh space. Moreover, the map $F_n : [0,1] \times E_n \to E_n$ given by $F_n(s,(t,i)) = (st,i)$, $s,t \in [0,1]$, $i = 1,2,\dots,n$, is well-defined and continuous.

LEMMA 3. *Let M be a Polish space and ν be a finite measure on $(M,\mathcal{B}(M))$, where $\mathcal{B}(M)$ is the Borel algebra over M. Let $f : M \to E_n$ be a measurable map with $f(M) \subset \{(1,i); i = 1,2,\dots,n\}$. Then for any $\varepsilon > 0$, there is a continuous map $g : M \to E_n$ such that $\nu(f \ne g) \le \varepsilon$.*

PROOF. We prove the assertion by induction in n. If $n = 1$, the map f itself is continuous. Let us assume that the assertion holds for n.

Let $f : M \to E_{n+1}$ be a measurable space with $f(M) \subset \{(1,i); i = 1,2,\dots,n+1\}$. Note that E_n can be regarded as a metric subspace of E_{n+1}. Let $A = f^{-1}((1,n+1))$. Then there are compact sets K_0 and K_1 in M such that

$$K_0 \subset M \setminus A,\ \nu((M \setminus A) \setminus K_0) < \varepsilon/3, \qquad K_1 \subset A,\ \nu(A \setminus K_1) < \varepsilon/3.$$

Then we have a continuous function $\varphi : M \to [-1,1]$ such that $\varphi(x) = -1, x \in K_0$, and $\varphi(x) = 1, x \in K_1$. Let $M' = \varphi^{-1}([-1,0])$ and $f' : M' \to E_n$ be given by

$$f'(x) = \begin{cases} f(x) & \text{if } f(x) \in E_n, \\ (1,1) & \text{otherwise .} \end{cases}$$

Then by the assumption of induction, there is a continuous function $g' : M' \to E_n$ such that $\nu(\{x \in M';\ f'(x) \neq g'(x)\}) < \varepsilon/3$. Now let $g : M \to E_{n+1}$ be given by

$$g(x) = \begin{cases} (\varphi(x), n+1) & \text{if } \varphi(x) > 0, \\ F_n(-\varphi(x), g'(x)) & \text{if } \varphi(x) \leq 0. \end{cases}$$

Then one can easily check that g is continuous. Also, we see that

$$\{f = g\} \supset K_1 \cup (K_0 \cap \{f' = g'\}),$$

and so $\nu(f \neq g) \leq \varepsilon$. Thus the assertion holds for $n+1$. This completes the proof. $\square$

Now let us prove Theorem 2. Since N is separable, there is a dense countable subset $\{y_k\}_{k=1}^{\infty}$ in N. Let $B_n = \{y \in N;\ dis_N(y, y_n) < \varepsilon/4\}$, $n = 1, 2, \ldots$. Then we see that $\bigcup_{n=1}^{\infty} B_n = N$. Let $A_1 = f^{-1}(B_1)$ and $A_n = f^{-1}(B_n) \setminus (\cup_{k=1}^{n-1} f^{-1}(B_k))$, $n \geq 2$. Then $A_n, n = 1, 2, \ldots$ are mutually disjoint and $\bigcup_{n=1}^{\infty} A_n = M$. So there is an $n \geq 1$ such that

$$(\int_M (1 - \sum_{k=1}^{n} 1_{A_k}(x)) dis_N(f(x), y_1)^p \mu(dx))^{1/p} < \varepsilon/4.$$

Since N is arcwise connected, there is a continuous map such that $\varphi : E_n \to N$ such that $\varphi((0,1)) = y_1$ and $\varphi((1,k)) = y_k$, $k = 1, \ldots, n$. Let $a = \max_{z,z' \in E_n} dis_N(\varphi(z), \varphi(z'))$. Let $h : M \to E_n$ be given by

$$h(x) = \begin{cases} (1,k), & x \in A_k, k = 2, \ldots, n, \\ (1,1), & \text{otherwise} . \end{cases}$$

Then by Lemma 3, there is a continuous map $h' : M \to E_n$ such that $\mu(h \neq h') < (\varepsilon/(2a+1))^p$. Let $g = \varphi \circ h'$. Then $g \in C_b(M; N)$ and

$$dis_N(f(x), (\varphi \circ h)(x)) \leq (1 - \sum_{k=1}^{n} 1_{A_k}(x)) dis_N(f(x), y_1) + \varepsilon/4,$$

and

$$(\int_M dis_N((\varphi \circ h)(x), (\varphi \circ h')(x))^p \mu(dx))^{1/p} \leq a\mu(h \neq h')^{1/p} \leq \varepsilon/2.$$

These imply that

$$(\int_M dis_N(f(x), g(x))^p \mu(dx))^{1/p} \leq \varepsilon.$$

This completes the proof. $\square$

3. Properties of $Dis^{(p)}_{M;N}$

Now let M be a Polish space and N be a separable metric space. We easily see the following.

PROPOSITION 4. *(1) $Dis^{(p)}_{M;N}((X_1,\mu_1),(X_2,\mu_2)) = 0$, if and only if $\mu_1 = \mu_2$ and $X_1(x) = X_2(x)$ $\mu_1 - a.s.x$, for $(X_1,\mu_1),(X_2,\mu_2) \in \mathcal{X}^p_{M;N}$.*
(2) $Dis^{(p)}_{M;N}((X_1,\mu_1),(X_2,\mu_2)) = Dis^{(p)}_{M;N}((X_2,\mu_2),(X_1,\mu_1))$
and

$$\begin{aligned} & Dis^{(p)}_{M;N}((X_1,\mu_1),(X_3,\mu_3)) \\ \le\ & Dis^{(p)}_{M;N}((X_1,\mu_1),(X_2,\mu_2)) + Dis^{(p)}_{M;N}((X_2,\mu_2),(X_3,\mu_3)) \end{aligned}$$

for any $(X_i,\mu_i) \in \mathcal{X}^p_{M;N}$, $i = 1,2,3$.

REMARK. Let us define an equivalence relation $\sim$ in $\mathcal{X}^p_{M;N}$ for which $(X_1,\mu_1) \sim (X_2,\mu_2)$ if $\mu_1 = \mu_2$ and $X_1(x) = X_2(x)$ $\mu_1 - a.s.x$. Then $(\mathcal{X}_{M;N}/\sim, Dis^{(p)}_{M;N})$ becomes a metric space, and our notion of convergence is the same as the associated one to this metric.

PROPOSITION 5. *Let $(X_n,\mu_n),(X,\mu) \in \mathcal{X}^p_{M;N}$, $n \ge 1$, and suppose that $(X_n,\mu_n) \to (X,\mu)$, $n \to \infty$, in $\mathcal{X}^p_{M;N}$. Then there are a probability space $(\Omega,\mathcal{F},P)$ and M-valued random variables Z_n, $n \ge 1$, and Z such that*
(1) $P \circ Z_n^{-1} = \mu_n, n \ge 1$, and $P \circ Z^{-1} = \mu$,
(2) $Z_n \to Z$, $n \to \infty$ in probability,
and
(3) $X_n \circ Z_n \to X \circ Z$, $n \to \infty$, in $L^p(\Omega,\mathcal{F},P)$.

PROOF. By the assumption, there are $\nu_n \in \mathcal{P}(M \times M)$, $n \ge 1$, such that $\nu_n \circ \pi_1^{-1} = \mu$, $\nu_n \circ \pi_2^{-1} = \mu_n$, $n \ge 1$, and

$$\int_{M\times M} (dis_M(x_1,x_2) \wedge 1 + dis_N(X(x_1),X_n(x_2)))^p \nu_n(dx_1,dx_2) \to 0.$$

Since M is Polish, there exist measurable maps $\rho_n : M \to \mathcal{P}(M)$, $n \ge 1$, such that $\nu_n(dx_1,dx_2) = \mu(dx_1)\rho_n(x_1)(dx_2)$. Let $\Omega = M^{\{0\}\cup\mathbf{N}}$, $\mathcal{F}$ be a Borel algebra of Ω and $P(dx) = \mu(dx_0) \otimes (\otimes_{n=1}^{\infty}\rho_n(x_0)(dx_n))$. Let $Z : \Omega \to M$,

$Z_n : \Omega \to M$, $n \geq 1$, be given by $Z(x) = x_0$, $Z_n(x) = x_n$, $n \geq 1, x = (x_0, x_1, \ldots)$. Then we have $P \circ Z^{-1} = \mu$, $P \circ Z_n^{-1} = \mu_n$, and

$$\begin{aligned} & E^P[((dis_M(Z, Z_n) \wedge 1) + dis_N(X \circ Z, X_n \circ Z_n))^p] \\ = & \int_{M \times M} ((dis_M(x_1, x_2) \wedge 1 + dis_N(X(x_1), X_n(x_2)))^p) \nu_n(dx_1, dx_2) \to 0. \end{aligned}$$

So, we have our assertion. □

PROPOSITION 6. *Let* $(X_n, \mu_n), (X, \mu) \in \mathcal{X}^p_{M;N}$, $n \geq 1$, *and suppose that there are a probability space* $(\Omega, \mathcal{F}, P)$ *and* M*-valued random variables* Z_n, $n \geq 1$, *and* Z *such that*
(1) $P \circ Z_n^{-1} = \mu_n$, $n \geq 1$, *and* $P \circ Z^{-1} = \mu$,
(2) $Z_n \to Z$, $n \to \infty$, *in probability,*
and
(3) $dis_N(X_n \circ Z_n, X \circ Z) \to 0$, $n \to \infty$, *in* $L^p(\Omega, \mathcal{F}, P)$.
Then $(X_n, \mu_n) \to (X, \mu)$ *in* $\mathcal{X}^p_{M;N}$.

PROOF. Let $\nu_n \in \mathcal{P}(M \times M)$ given by $\nu_n = P \circ (Z, Z_n)^{-1}$. Then we have

$$\begin{aligned} & Dis^{(p)}_{M;N}((X, \mu), (X_n, \mu_n))^p \\ \leq & \int_{M \times M} (dis_M(x_1, x_2) \wedge 1 + dis_N(X(x_1), X_n(x_2)))^p \nu_n(dx_1, dx_2) \\ = & E^P[(dis_M(Z, Z_n) \wedge 1 + dis_N(X \circ Z, X_n \circ Z_n))^p] \to 0. \end{aligned}$$

Thus we see that $(X_n, \mu_n) \to (X, \mu)$ in $\mathcal{X}^p_{M;N}$. □

LEMMA 7. *Suppose that* N *is an arcwise connected separable metric space. Let* (X_n, μ_n), $(X, \mu) \in \mathcal{X}^p_{M;N}$, $n \geq 1$. *Then* $(X_n, \mu_n) \to (X, \mu)$ *in* $\mathcal{X}^p_{M;N}$, *if and only if* $\mu_n \to \mu$ *weakly as* $n \to \infty$, *and*

$$\inf\{\limsup_{n \to \infty} E^{\mu_n}[dis_N(X_n, G)^p] + E^{\mu}[dis_N(X, G)^p];\ G \in C_b(M; N)\} = 0.$$

PROOF. (if part) Since $\mu_n \to \mu$ weakly, by Skorohod's theorem there are a probability space $(\Omega, \mathcal{F}, P)$ and M-valued random variables Z, Z_n,

$n \geq 1$, such that $P \circ Z_n^{-1} = \mu_n$, $n \geq 1$, $P \circ Z^{-1} = \mu$, and $Z_n \to Z$ in probability. For any $\varepsilon > 0$, there is a $G \in C_b(M; N)$ such that

$$\limsup_{n\to\infty} E^{\mu_n}[dis_N(X_n, G)^p]^{1/p} + E^{\mu}[dis_N(X, G)^p]^{1/p} < \varepsilon.$$

Then

$$\begin{aligned}
&\limsup_{n\to\infty} E^P[dis_N(X_n \circ Z_n, X \circ Z)^p]^{1/p} \\
\leq\ & E^P[dis_N(X \circ Z, G \circ Z)^p]^{1/p} + \limsup_{n\to\infty} E^P[dis_N(X_n \circ Z_n, G \circ Z_n)^p]^{1/p} \\
<\ & \varepsilon.
\end{aligned}$$

So we see that $dis_N(X_n \circ Z_n, X \circ Z) \to 0$ in $L^p(\Omega, \mathcal{F}, P)$. Thus we have our assertion from Proposition 6.
(only if part) Let $(\Omega, \mathcal{F}, P)$, $Z_n, n \geq 1$, and Z be as in Proposition 5. Then we see that $\mu_n \to \mu$ weakly. We see from Theorem 2 that for any $\varepsilon > 0$ there is a $G \in C_b(M; N)$ such that $E^{\mu}[dis_N(X, G)^p]^{1/p} < \varepsilon$. Then we have

$$\begin{aligned}
&\limsup_{n\to\infty} E^{\mu_n}[dis_N(X_n, G)^p]^{1/p} \\
=\ & \limsup_{n\to\infty} E^P[dis_N(X_n \circ Z_n, G \circ Z_n)^p]^{1/p} \\
=\ & E^P[dis_N(X \circ Z, G \circ Z)^p]^{1/p} < \varepsilon.
\end{aligned}$$

This implies our assertion. □

Now we are ready to prove Theorem 1.

Proof of Theorem 1. The ' if' part follows from Proposition 6. So it is sufficient to prove the 'only if' part. By Lemma 7 for any $\varepsilon > 0$, there is a $G \in C_b(M; N)$ such that

$$E^{\mu}[dis_N(X, G)^p]^{1/p} + \limsup_{n\to\infty} E^{\mu_n}[dis_N(X_n, G)^p]^{1/p} < \varepsilon.$$

Then we have

$$\begin{aligned}
&\limsup_{n\to\infty} E^P[dis_N(X \circ Z, X_n \circ Z_n)^p]^{1/p} \\
\leq\ & E^P[dis_N(X \circ Z, G \circ Z)^p]^{1/p} + \limsup_{n\to\infty} E^P[dis_N(X_n \circ Z_n, G \circ Z_n)^p]^{1/p} \\
& \qquad + \limsup_{n\to\infty} E^P[dis_N(G \circ Z, G \circ Z_n)^p]^{1/p} \\
<\ & \varepsilon.
\end{aligned}$$

Thus we have our assertion.

Let us remind again that if $\mu_n \to \mu$ weakly, then by Skorohod's theorem there are a probability space (Ω, F, P) and measurable maps $Z_n : \Omega \to M$, $n \geq 1$, and $Z : \Omega \to M$ such that $P \circ Z_n^{-1} = \mu_n$, $n \geq 1$, and $P \circ Z^{-1} = \mu$ and that $Z_n \to Z$, $n \to \infty$, $P - a.s.$

The following is easy consequence of Theorem 1 and Skorohod's theorem.

Corollary 8. *Suppose that $\mu, \mu_n \in \mathcal{P}(M)$, $n \geq 1$.*
(1) *Let N be an arcwise connected Polish space and M be a Polish space, and let $X_n : M \to N$, $X : M \to N$, $X_n^{(k)} : M \to N$ and $X^{(k)} : M \to N$, $n, k \geq 1$, be measurable maps. If $(X_n^{(k)}, \mu_n) \to (X^{(k)}, \mu)$ in $\mathcal{X}^p_{M;N}$, $k = 1, 2, \ldots,$ and if*

$$\lim_{k\to\infty} \sup \sup_n E^{\mu_n}[dis_N(X_n, X_n^{(k)})^p] = 0$$

and

$$\lim_{k\to\infty} \sup E^{\mu}[dis_N(X, X^{(k)})^p] = 0,$$

then $(X_n, \mu_n) \to (X, \mu)$ in $\mathcal{X}^p_{M;N}$.
(2) *Let N_1 and N_2 be arcwise connected Polish spaces, $X_n^{(i)} : M \to N_i$, $n \geq 1$, and $X^{(i)} : M \to N_i$, $i = 1, 2$, be measurable maps. If $(X_n^{(i)}, \mu_n) \to (X^{(i)}, \mu)$ in $\mathcal{X}^p_{M;N_i}$, $i = 1, 2$, then $((X_n^{(1)}, X_n^{(2)}), \mu_n) \to ((X^{(1)}, X^{(2)}), \mu)$ in $\mathcal{X}^p_{M;N_1 \times N_2}$.*
(3) *Let E be a real Banch space. $X_n^{(i)} : M \to E$, $n \geq 1$, and $X^{(i)} : M \to E$, $i = 1, 2$, be measurable maps. If $(X_n^{(i)}, \mu_n) \to (X^{(i)}, \mu)$ in $\mathcal{X}^p_{M;E}$, $i = 1, 2$, and if $a_n \to a$, and $b_n \to b$ in $\mathbf{R}$, then $(a_n X_n^{(1)} + b_n X_n^{(2)}, \mu_n) \to (aX^{(1)} + bX^{(2)}, \mu)$ in $\mathcal{X}^p_{M;E}$.*

4. Application

Let $d \geq 1$, and $\{v_0, v_1, \ldots, v_d\}$ be a subset of $\mathbf{R}^d$ in a general position, i,e, $v_1 - v_0, \ldots, v_d - v_0$ are linearly independent. Let us assume that $p_0, p_1, \ldots, p_d \in (0, 1)$ exist and satisfy

$$\sum_{i=0}^d p_i = 1, \qquad \sum_{i=0}^d p_i v_i = 0, \qquad \sum_{i=0}^d p_i v_i \otimes v_i = I_d,$$

where I_d is the $d \times d$ identity matrix.

Let $(\Omega_n, \mathcal{F}_n, P_n)$, $n \geq 1$ be probability spaces, and let $Z_k^{(n)}$, $k = 1, 2, \dots, n$ be independent identically distributed $\mathbf{R}^d$-valued random variables defined on $(\Omega_n, \mathcal{F}_n, P_n)$ such that

$$P_n(Z_k^{(n)} = v_i) = p_i, \qquad i = 0, 1, \dots, d, \quad k = 1, \dots, n.$$

Let

$$W_n(t) = \frac{1}{\sqrt{n}} \sum_{k=1}^{[nt]} Z_k^{(n)}, \qquad t \in [0,1].$$

Now let us think of the space $D^d = D([0,1]; \mathbf{R}^d)$. Let $X : [0,1] \times D^d \to \mathbf{R}^d$ be defined by $X(t, w) = w(t)$, $t \in [0,1]$, $w \in D^d$. We define a filtration $\{\mathbf{F}_t\}_{t \in [0,1]}$ by

$$\mathcal{F}_t = \bigcap_{u > t} \sigma\{X(s, \cdot);\ s \leq u \wedge 1\}.$$

Let μ_n, $n \geq 1$, denote a probability measure $P_n \circ W_n^{-1}$ on D^d.

Let $\mathcal{L}_n$, $n \geq 1$, be the set of bounded function $f : D^d \times [0,1] \to \mathbf{R}^d$ such that

$$f(w)(t) = f(w)(\frac{k}{n}), \qquad \frac{k-1}{n} < t \leq \frac{k}{n},\ k = 1, \dots, n,$$

and $f(\cdot)(t) : D^d \to \mathbf{R}^d$ is continuous and $\mathcal{F}_{(k-1)/n}$-measurable for $(k-1)/n < t \leq k/n$, $k = 1, \dots, n$. Then we have the following (e.g. Nakayama [3]).

Proposition 9. *Let $n \geq 1$.*
(1) *If $X = c + \int_{(0,1]} f(w)(t) dw(t)$, $c \in \mathbf{R}$, $f \in \mathcal{L}_n$, then*

$$E^{\mu_n}[X^2] = c^2 + E^{\mu_n}[\int_{(0,1]} |f(w)(t)|^2 dt].$$

(2) *For any $X \in L^2(D^d, \mathcal{F}_1, d\mu_n)$ there exist $c_n(X) \in \mathbf{R}$ and $F_n(X) \in \mathcal{L}_n$, such that*

$$X = c_n(X) + \int_{(0,1]} (F_n(X)(w))(t) dw(t).$$

Let μ be a standard Wiener measure on D^d. Let $\mathcal{L}$ be the set of measurable maps $f : D^d \to L^2([0,1]; \mathbf{R}^d, dt)$ such that $(t,w) \in [0,1] \times D^d \to f(w)(t) \in \mathbf{R}^d$ is progressively measurable and

$$E^{\mu}[\int_0^1 |f(w)(t)|^2 dt] < \infty.$$

Then we have the following naturally by Ito's representation theorem.

PROPOSITION 10. (1) *If* $X = c + \int_{(0,1]} f(w)(t)dw(t)$, $c \in \mathbf{R}$, $f \in \mathcal{L}$, *then*

$$E^{\mu}[X^2] = c^2 + E^{\mu}[\int_{(0,1]} |f(w)(t)|^2 dt].$$

(2) *For any* $X \in L^2(D^d, \mathcal{F}_1, d\mu)$ *there exist* $c(X) \in \mathbf{R}$ *and* $F(X) \in \mathcal{L}$, *such that*

$$X = c(X) + \int_0^1 (F(X)(w))(t)dw(t).$$

Then Donsker's invariance principle (c.f.[1]) implies that $\mu_n \to \mu$, weakly as $n \to \infty$. Our main purpose in this section is to prove the following.

THEOREM 11. (1) *If* $(X_n, \mu_n) \to (X, \mu)$, *in* $\mathcal{X}^2_{D^d;\mathbf{R}}$, *then* $((c_n(X_n), F_n(X_n)), \mu_n) \to ((c(X), F(X)), \mu)$ *in* $\mathcal{X}^2_{D^d;\mathbf{R}\times L^2([0,1];\mathbf{R}^d,dt)}$, *and* $(\{E^{\mu_n}[X_n|\mathcal{F}_t]\}_{t\in[0,1]}, \mu_n) \to (\{E^{\mu}[X|\mathcal{F}_t]\}_{t\in[0,1]}, \mu)$ *in* $\mathcal{X}^2_{D^d;D}$.
(2) *If* $f_n \in \mathcal{L}_n$, $n \geq 1$, *and* $f \in \mathcal{L}$, *and if* $(f_n, \mu_n) \to (f, \mu)$ *in* $\mathcal{X}^2_{D^d;L^2([0,1];\mathbf{R}^d,dt)}$, *then* $(\int_{(0,1]} f_n(w)(t)dw(t), \mu_n) \to (\int_0^1 f(w)(t)dw(t), \mu)$, $n \to \infty$, *in* $\mathcal{X}^2_{D^d;\mathbf{R}}$.

We make some preparations before proving this theorem.

PROPOSITION 12. *Let* $m \geq 1$ *and* $f \in \mathcal{L}_m$. *Then* $(\int_{(0,1]} f(w)(t)dw(t), \mu_n) \to (\int_0^1 f(w)(t)dw(t), \mu)$, $n \to \infty$, *in* $\mathcal{X}^2_{D^d;\mathbf{R}}$. *Also,* $(\{\int_{(0,t]} f(w)(s)dw(s)\}_{t\in[0,1]}, \mu_n) \to (\{\int_{(0,t]} f(w)(s)dw(s)\}_{t\in[0,1]}, \mu)$, $n \to \infty$, *in* $\mathcal{X}^2_{D^d;D}$.

PROOF. Note that if $n > 2m$ for $\mu_n - a.s.w$

$$\int_{(0,t]} f(w)(s)dw(s)$$

$$
\begin{aligned}
&= \sum_{k=1}^{n} f(w)(\frac{k}{n})(w((\frac{k}{n}) \wedge t) - w((\frac{k-1}{n}) \wedge t)) \\
&= \sum_{j=1}^{m} \sum_{k=1}^{n} 1_{((j-1)/m, j/m)}(k/n) f(w)(\frac{j}{m})(w((\frac{k}{n}) \wedge t) - w((\frac{k-1}{n}) \wedge t)). \\
&= \sum_{j=1}^{m} f(w)(\frac{j}{m})(w((\frac{j}{m}) \wedge t) - w((\frac{j-1}{m}) \wedge t)) \qquad \mu_n - a.s..
\end{aligned}
$$

Here we use the fact that

$$
\begin{aligned}
&\sum_{k=1}^{n} 1_{((j-1)/m, j/m)}(k/n)(w((\frac{k}{n}) \wedge t) - w((\frac{k-1}{n}) \wedge t))] \\
= \ & w((\frac{j}{m}) \wedge t) - w(\frac{j-1}{m} \wedge t) \qquad \mu_n - a.s.w.
\end{aligned}
$$

For $N > n$, let $G_N : D^d \to D$ be given by

$$
G_N(w)(t) = \sum_{j=1}^{m} f(w)(\frac{j}{m}) N \int_0^{1/N} (w((\frac{j}{m} + s) \wedge t) - w((\frac{j-1}{m} + s) \wedge t)) ds.
$$

Then G_N is a bounded continuous map and

$$
\begin{aligned}
&|G_N(w)(t) - \sum_{j=1}^{m} f(w)(\frac{j}{m})(w((\frac{j}{m}) \wedge t) - w((\frac{j-1}{m}) \wedge t))| \\
\leq \ & 2m(\sup_{t,w} |f(w)(t)|)(\sup_{t\in[0,1], s\in[0,1/N]} |w((t+s) \wedge 1) - w(t)|).
\end{aligned}
$$

So we see that

$$
\begin{aligned}
&\sup_{t\in[0,1]} |G_N(w)(t) - \int_{(0,t]} f(w)(s) dw(s))| \\
\leq \ & 2m((\sup_{t,w} |f(w)(t)|)(\sup_{t\in[0,1], s\in[0,1/N]} |w((t+s) \wedge 1) - w(t)|) \\
&\quad + n^{-1/2} \max\{|v_0|, \ldots, |v_d|\})
\end{aligned}
$$

for $\mu_n - a.s.w.$ This implies that

$$
\begin{aligned}
&\lim_{N\to\infty} (\limsup_{n\to\infty} E^{\mu_n}[\sup_{t\in[0,1]} (\int_{(0,t]} f(w)(s) dw(s) - G_N(w)(t))^2] \\
&\qquad + E^{\mu}[\sup_{t\in[0,1]} (\int_{(0,t]} f(w)(s) dw(s) - G_N(w)(t))^2]) = 0.
\end{aligned}
$$

Therefore by Lemma 7 we have our assertion. $\square$

Also, we have the following.

PROPOSITION 13. *Let* $m \geq 1$, $f \in \mathcal{L}_m$ *and* $X_n = \int_{(0,1]} f(w)(t)dw(t)$, $\mu_n - a.s.w.$ *Then*

$$E^{\mu_n}[\int_0^1 |F_n(X_n)(w)(t) - f(w)(t)|^2 dt] \to 0, \qquad n \to \infty$$

PROOF. Let $n > m$. Then for $k = 1, \ldots, n$, we see that

$$F_n(X_n)(w)(t) = f(w)(\frac{k}{n}), \qquad t \in (\frac{k-1}{n}, \frac{k}{n}], \qquad \mu_n - a.s.w.$$

So we have our assertion. $\square$

Let $\mathcal{L}' = \bigcup_{m=1}^{\infty} \mathcal{L}_m$. The following is well-known.

PROPOSITION 14. *For any* $f \in \mathcal{L}$ *and* $\varepsilon > 0$, *there is a* $g \in \mathcal{L}'$ *such that*

$$E^{\mu}[\int_0^1 |f(w)(t) - g(w)(t)|^2 dt] < \varepsilon.$$

Now let us prove Theorem 11. First we prove the assetion (1).

Suppose that $(X_n, \mu_n) \to (X, \mu)$, in $\mathcal{X}^2_{D^d;\mathbf{R}}$. Then by Proposition 14 there are $g_k \in \mathcal{L}'$, $k \geq 1$, such that

$$E^{\mu}[\int_0^1 |F(X)(w)(t) - g_k(w)(t)|^2 dt] \to 0, \qquad k \to \infty.$$

Let $X_n^{(k)} = c(X) + \int_{(0,1]} g_k(w)(t)dw(t) \in L^2(D^d, d\mu_n)$, $n \geq 1$, and $X^{(k)} = c(X) + \int_0^1 g_k(w)(t)dw(t) \in L^2(D^d, d\mu)$. Then we have

$$E^{\mu}[|X - X^{(k)}|^2] \to 0, \qquad k \to \infty$$

By Proposition 13, we see that $((c_n(X_n^{(k)}), F_n(X_n^{(k)})), \mu_n) \to ((c(X^{(k)}), F(X^{(k)})), \mu)$ in $\mathcal{X}^2_{D^d,\mathbf{R}\times L^2([0,1];\mathbf{R}^d,dt)}$. By Proposition 12, we see that $(\{E^{\mu_n}[X_n^{(k)}|\mathcal{F}_t]\}_{t\in[0,1]}, \mu_n) \to (\{E^{\mu}[X^{(k)}|\mathcal{F}_t]\}_{t\in[0,1]}, \mu)$, $n \to \infty$, in $\mathcal{X}^2_{D^d;D}$.

By Lemma 7, Corollary 8(3) and Proposition 9, we have

$$\begin{aligned} & |c_n(X_n) - c_n(X_n^{(k)})|^2 + E^{\mu_n}[\int_0^1 |F_n(X_n) - F_n(X_n^{(k)})|^2 dt] \\ = \quad & E^{\mu_n}[|X_n - X_n^{(k)}|^2] \to E^{\mu}[|X - X^{(k)}|^2], \qquad n \to \infty \end{aligned}$$

and

$$\begin{aligned} & E^{\mu_n}[\sup_{t\in[0,1]} |E^{\mu_n}[X_n|\mathcal{F}_t] - E^{\mu_n}[X_n^{(k)}|\mathcal{F}_t]|^2] \\ \leq \quad & 4E^{\mu_n}[|X_n - X_n^{(k)}|^2] \to 4E^{\mu}[|X - X^{(k)}|^2], \qquad n \to \infty. \end{aligned}$$

So by Corollary 8(1) we see that $((c_n(X_n), F_n(X_n)), \mu_n) \to ((c(X), F(X)), \mu)$ in $\mathcal{X}^2_{D^d,\mathbf{R}\times L^2([0,1];\mathbf{R}^d,dt)}$. This completes the proof of the assertion (1).

Now suppose that $f_n \in \mathcal{L}_n$, $n \geq 1$, and $f \in \mathcal{L}$, and that $(f_n, \mu_n) \to (f, \mu)$ $\in \mathcal{X}^2_{D^d;L^2([0,1];\mathbf{R}^d,dt)}$. Then by Proposition 14 there are $h_k \in \mathcal{L}'$, $k \geq 1$, such that

$$E^{\mu}[\int_0^1 |f(w)(t) - h_k(w)(t)|^2 dt] \to 0, \qquad k \to \infty.$$

Let $Y_n^{(k)} = \int_{(0,1]} h_k(w)(t)dw(t) \in L^2(D^d, \mu_n)$, and $Y^{(k)} = \int_0^1 h_k(w)(t)dw(t)$ $\in L^2(D^d, \mu)$. Then we have

$$E^{\mu}[|\int_0^1 f(w)(t)dw(t) - Y^{(k)}|^2] \to 0, \qquad k \to \infty.$$

Also, by Proposition 12 we see that $(Y_n^{(k)}, \mu_n) \to (Y^{(k)}, \mu)$ in $\mathcal{X}^2_{D^d;\mathbf{R}}$, and $(F_n(Y_n^{(k)}), \mu_n) \to (h_k, \mu)$ in $\mathcal{X}^2_{D^d;L^2([0,1];\mathbf{R}^d,dt)}$. Therefore we have

$$\begin{aligned} & E^{\mu_n}[|\int_{(0,1]} f_n(w)(t)dw(t) - Y_n^{(k)}|^2] \\ = \quad & E^{\mu_n}[\int_0^1 |f_n(w)(t) - F_n(Y_n^{(k)})(t)|^2 dt] \\ \to \quad & E^{\mu}[\int_0^1 |f(w)(t) - h_k(t)|^2 dt], \qquad n \to \infty. \end{aligned}$$

So by Corollary 8(1) and Proposition 12 we see that $(\int_{(0,1]} f_n(w)(t)dw(t), \mu_n)$ $\to (\int_0^1 f(w)(t)dw(t), \mu)$ in $\mathcal{X}^2_{D^d;\mathbf{R}}$. This completes the proof of Thorerm 11.

References

[1] Bilingsley, P., *Convergence of Probability Measures*, Wiley, New York, 1968.
[2] Kusuoka, S., A remark on American Securities, in "Ito's Stochastic Calculus and Probability Theory" ed. N. Ikeda, S. Watanabe, M. Fukushima and K. Kunita, pp. 213–231, Springer Berlin, 1996.
[3] Nakayama, T., Approximation of BSDE's by stochastic difference equations, Preprint.

(Received October 2, 2000)

Shigeo KUSUOKA
Graduate School of Mathematical Sciences
The University of Tokyo
Meguro-ku, Komaba 3-8-1
Tokyo, 153-8914 Japan

Toshiyuki NAKAYAMA
The Bank of Tokyo-Mitsubishi
Chiyoda-ku, Marunouchi 2-7-1
Tokyo, 100-8388 Japan