PROBABILITY
AND
MATHEMATICAL STATISTICS

Vol. 12, Fasc. 2 (1991), p. 319–334

ON THE APPROXIMATION THEOREM OF THE WONG–ZAKAI TYPE FOR THE FUNCTIONAL STOCHASTIC DIFFERENTIAL EQUATIONS

BY

KRYSTYNA TWARDOWSKA (WARSZAWA)

Abstract. In this paper we examine the generalization of the Wong–Zakai theorem for the nonlinear stochastic functional differential equations with values in the space R^d ($d \geqslant 1$). As the result of the piecewise linear approximation of the m-dimensional Wiener process we obtain an explicit formula for the limit of a sequence of solutions to the ordinary differential equations with a delay argument; this very limit is a solution to the stochastic differential equation with a delay argument with the additional term called the *Itô correction term.*

1. Introduction. The approximation Wong–Zakai theorem [12] was generalized, e.g., for the multidimensional case ([5], [11]), for the more general noises than the Wiener process ([7], [10]), for the infinite-dimensional case ([1]–[3]), but it did not include the stochastic differential equations with a delay argument.

What this paper contains is just such a generalization. We have been using the general theory of the functional stochastic differential equations ([6], [8], [9]). Mainly, we have based our argumentation on the tools of the approximation theorem in [5].

As an example we have considered the linear equation for the delay which is constant in time and for the noise being the one-dimensional Wiener process.

We wish to thank Professor J. Zabczyk for inspiring discussions on this subject and to Dr. A. L. Dawidowicz for several very constructive comments.

2. Definitions and notation. Let $t \in [0, T]$ and let $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$ be a complete probability space with $\mathscr{F}_t = (\mathscr{F}_t)_{t \in [0,T]}$ being the increasing family of sub-σ-algebras of the σ-algebra $\mathscr{F}$. We put $J = (-\infty, 0]$ and we introduce metric spaces $\mathscr{C}_- = C(J, R^d)$, $\mathscr{C}_1 = C((-\infty, T], R^d)$ and $\mathscr{C}_2^0 = C((-\infty, T], R^m) = \tilde{\Omega}$ of continuous functions. The space $\mathscr{C}_-$ is endowed with the metric

$$(f, g)_{\mathscr{C}_-} = \sum_{n=1}^{\infty} 2^{-n} \frac{\|f-g\|_n}{1+\|f-g\|_n} \quad \text{for } f, g \in \mathscr{C}_-, \ \|h\|_n = \max_{-n \leqslant t \leqslant 0} |h(t)|.$$

Similarly we define the metrics for $\mathscr{C}_1$ and $\mathscr{C}_2^0$ with $\|h\|_n = \max_{-n \leqslant t \leqslant T} |h(t)|$. Here d is a dimension of a state space and m is a dimension of the Wiener process; in the space $\mathscr{C}_2^0$ all functions are equal to zero at zero. We denote by $\mathscr{X}$ one of the above spaces.

Let $\mathscr{B}(\mathscr{X})$ denote the topological σ-algebra of the space $\mathscr{X}$. It is obvious that it is identical with σ-algebra generated by the family of all Borel cylinder sets in $\mathscr{X}$. So we construct the Wiener space $(\mathscr{C}_2^0, \mathscr{B}(\mathscr{C}_2^0), P^w)$, where P^w is a Wiener measure ([5], Chapter I). The coordinate process $B(t, w) = w(t)$, $w \in \mathscr{C}_2^0$, is an *m-dimensional Wiener process.*

The smallest Borel algebra that contains $\mathscr{B}_1, \mathscr{B}_2, \ldots$ is denoted by $\mathscr{B}_1 \cup \mathscr{B}_2 \cup \ldots$; $\mathscr{B}_{u,v}(X)$ denotes the smallest Borel σ-algebra for which a given stochastic process $X(t)$ is measurable for every $t \in [u, v]$; and $\mathscr{B}_{u,v}(dB)$ denotes the smallest Borel algebra for which $B(s) - B(t)$ is measurable for every (t, s) with $u \leqslant t \leqslant s \leqslant v$.

We introduce the condition

(A1) for every $t \in (-\infty, T]$ the algebra $\mathscr{B}_{-\infty,t}(X) \cup \mathscr{B}_{-\infty,t}(dB)$ is independent of $\mathscr{B}_{t,T}(dB)$

to give the meaning for the stochastic integrals in (2) below.

Let $B^n(t, w)$ be the following piecewise linear approximation of $B(t, w) = w(t)$:

$$B^{n,p}(t, w) = w^p(k/2^n) + 2^n(t - k/2^n)\big(w^p((k+1)/2^n) - w^p(k/2^n)\big) \tag{1}$$

for each $p = 1, \ldots, m$ and $kT/2^n \leqslant t < (k+1)T/2^n$ for $k = 0, 1, \ldots, 2^n - 1$.

We introduce the following notation and functions:

$$\delta = \frac{1}{2^n n}, \quad t_n^-(t) = \frac{[2^n t]}{2^n}, \quad t_n^+(t) = \frac{[2^n t]+1}{2^n}, \quad m(t) = \frac{t_n^-(t)}{2^{-n}},$$

where $[\cdot]$ denotes the integer part of the real number.

For further considerations a segment of trajectory must be defined. Let f be a function of $t \in (-\infty, T]$. For a fixed $t \in [0, T]$ we define a function f_t on $(-\infty, 0]$ by the formula

$$f_t(\theta) = f(t + \theta).$$

For the stochastic process $X(t, w)$ we define

$$X_t(\theta, w) = X(t + \theta, w), \qquad \theta \in J;$$

therefore $X_t(\cdot, w)$ denotes the segment of the trajectory $X(\cdot, w)$ on $(-\infty, t]$.

3. Description of a model. Now we consider $\tilde{\Omega}$, defined before by $\mathscr{C}_2^0 = C((-\infty, T], R^n)$. Let X be a continuous stochastic process $X(t, w)$: $(-\infty, T] \times \tilde{\Omega} \to R^d$, that is, X: $\tilde{\Omega} \to \mathscr{X} = \mathscr{C}_1$.

We take the fixed initial constant stochastic process

$$X^i(0 + \theta, w) = X_0^i(w) = X_0^{n,i}(w) = Y_0^i(w) \quad \text{ for } \theta \in J,\ i = 1, \ldots, d.$$

We introduce the following operators: $b\colon \mathscr{C}_- \to R^d$, $\sigma\colon \mathscr{C}_- \to L(R^m, R^d)$ ($L(R^m, R^d)$ is the Banach space of linear functions from R^m to R^d with the uniform operator norm $|\cdot|_L$). We assume

(A2) b and σ are continuous operators.

The additional definition of b and σ is necessary. Namely, we usually assume that the stochastic differential equations of the type considered here are satisfied for $t \geqslant 0$ only. However, in the proof of our theorem and in the definition of a correction term, our equations are to be satisfied also for $t < 0$. Therefore we define $b(t, \varphi) = \sigma(t, \varphi) = 0$ for $t < 0$ or $b_1(t, \varphi) = b(\varphi)\chi(t)$, $\sigma_1(t, \varphi) = \sigma(\varphi)\chi(t)$, where $\chi(t) = 1$ for $t \geqslant 0$ and $\chi(t) = 0$ for $t < 0$, if the operators b and σ do not depend apparently on t. This is possible because our initial function is constant on $(-\infty, 0]$. In such a sense the operators b and σ are understood in this paper. The fact that the above assumptions are necessary can be observed just in formula (**) in Section 4 in the proof of our theorem.

Now we introduce the operators

$$\tilde{b}\colon \mathscr{C}_- \to \mathscr{C}_- \quad \text{and} \quad \tilde{\sigma}\colon \mathscr{C}_- \to C(J, L(R^m, R^d)),$$

where

$$\tilde{b}\colon \mathscr{C}_- \ni g \to \big(J \ni \tau \to b(g(\cdot + \tau)) \in R^d\big),$$

$$\tilde{\sigma}\colon \mathscr{C}_- \ni g \to \big(J \ni \tau \to \sigma(g(\cdot + \tau)) \in L(R^m, R^d)\big),$$

that is, for $\tau < 0$ and a shift transformation $S_\tau\colon J \ni \vartheta \to \vartheta + \tau$,

$$[\tilde{b}(g)](\tau) = b(g \circ S_\tau) = b(g(\cdot + \tau)), \quad [\tilde{\sigma}(g)](\tau) = \sigma(g \circ S_\tau) = \sigma(g(\cdot + \tau)).$$

Remark. 1. From the above-given definitions we see that the initial process has to be constant.

Remark 2. This construction explains why we consider $(-\infty, 0]$ to be a domain of the initial function. In fact, we shall deal only with a part of this function on $[-r, 0]$ for a fixed real number $r > 0$. If we considered the segment $[-r, 0]$ to be a domain, it would make impossible to define correctly the functions $\tilde{b}$ and $\tilde{\sigma}$.

We consider the following stochastic differential equation with a delay argument for every $i = 1, \ldots, d$:

$$X^i(t, w) = X_0^i(w) + \int_0^t b^i(X_s(\cdot, w))ds + \sum_{p=1}^m \int_0^t \sigma^{ip}(X_s(\cdot, w))dw^p(s). \tag{2}$$

By replacing the Wiener process by B^n in equation (2) we obtain the following approximations of (2):

$$X^{n,i}(t, w) = X_0^{n,i}(w) + \int_0^t b^i(X_s^n(\cdot, w))ds + \sum_{p=1}^m \int_0^t \sigma^{ip}(X_s^n(\cdot, w))\dot{B}^{n,p}(s, w)ds. \tag{2^n}$$

For further considerations we formulate another stochastic differential equation, i.e.,

$$(3)\qquad Y^i(t, w) = Y_0^i(w) + \int_0^t b^i(Y_s(\cdot, w))ds + \sum_{p=1}^{m} \int_0^t \sigma^{ip}(Y_s(\cdot, w))dw^p(s)$$
$$+\frac{1}{2}\sum_{p=1}^{m}\sum_{j=1}^{d}\int_0^t \tilde{D}_j\sigma^{ip}(Y_s(\cdot, w))\sigma^{jp}(Y_s(\cdot, w))ds$$

for every $i = 1, \ldots, d$. Further, $D\sigma^{ip}$ is the Fréchet derivative from $\mathscr{C}_-$ to $L(\mathscr{C}_-, R)$ (necessary assumptions are given below) while $\tilde{D}_j\sigma^{ip}(Y_s(\cdot, w)) = \mu_{s,w,Y}^{ipj}(\{0\})$ is the j-th coordinate of a measure $\mu = \mu_{s,w,Y}^{ip}$ on $\mathscr{C}_-$ such that

$$\mu(\Phi) = \sum_{j=1}^{d} \int_{-\infty}^{0} \Phi_j(v)\mu^j(dv).$$

We have $\mu(A) = \mu(A \cap (-\infty, 0)) + \mu(A \cap \{0\}) = \tilde{\mu}(A) + \mu(\{0\})\delta_0(A)$, where δ_0 is a Dirac measure, $A \in \mathscr{B}((-\infty, 0))$. It is obvious that

$$D\sigma^{ip}(g)(\Phi) = \sum_{j=1}^{d} \int_{-\infty}^{0} \Phi_j(v)\mu_{s,w,g}^{ipj}(dv)$$

is a direction derivative. We notice that for a smooth function $h(\cdot)$ we have $\int_{-\infty}^{0} h(v)\delta_0(dv) = h(0)$. We introduce a function

$$\tilde{A}_t^{jpn}\colon\ \tau \in J \to \sigma^{jp}(X_{t+\tau}^n(\cdot, w))\dot{B}^{n,p}(t+\tau, w) \in R.$$

Remark 3. We may take in (3) the integral $\int_0^t$ because σ^{ip} has no sense for $s \leqslant r$ but we have there

$$\frac{dX^n(t+\tau, w)}{dt} = 0$$

(see (**) in Section 4).

We put $\Psi(t, w) = b(X_t(w))$ and $\Phi(t, w) = \sigma(X_t(w))$. The second integral in (2) is the Itô integral ([5], [8]).

Let us introduce the following conditions:

(A3) The initial stochastic process X_0 is $\mathscr{F}_0$-measurable and

$$P(|X_0(w)| < \infty) = 1, \quad \text{where } |X_0(w)| = \sum_{i=1}^{d} |X_0^i(w)|,$$

$\mathscr{B}_{-\infty,0}(X_0)$ is independent of $\mathscr{B}_{0,T}(B)$.

(A4) For every $\varphi, \psi \in \mathscr{C}_-$ the following Lipschitz condition is fulfilled:

$$|b(\varphi)-b(\psi)|^2 + |\sigma(\varphi)-\sigma(\psi)|_L^2 \leqslant L^1 \int_{-\infty}^{0} |\varphi(\theta)-\psi(\theta)|^2 dK\theta + L^2|\varphi(0)-\psi(0)|^2,$$

where $K(\theta)$ is a certain bounded measure on J, L^1 and L^2 are some constants.

(A5) For every $\varphi \in \mathscr{C}_-$ the following growth condition is fulfilled:

$$(b(\varphi))^2 + (\sigma(\varphi))_L^2 \leqslant L^1 \int_{-\infty}^{0} (1+\varphi^2(\theta))dK\theta + L^2(1+\varphi^2(0)).$$

(A6) We have

$$P\Big(\int_0^T |b(X_s)|ds < \infty\Big) = 1, \quad P\Big(\int_0^T \sigma^2(X_s)ds < \infty\Big) = 1.$$

(A7) $b^i, \sigma^{ip} \in \mathscr{C}_b^1(\mathscr{C}_-)$ for every $i = 1, \ldots, d$, $p = 1, \ldots, m$, where $\mathscr{C}_b^1$ is the space of bounded mappings with continuous first derivative, that is, for any numbers $A > 0$ and $\varepsilon > 0$ there exist numbers $B > 0$ and $\delta > 0$ such that $\|X_s^1 - X_s^2\|_{-B}^0 < \delta$ implies

$$\Big|\int_{-\infty}^0 \Phi(v)\mu_1(dv) - \int_{-\infty}^0 \Phi(v)\mu_2(dv)\Big| < \|\Phi\|_{-A}^0\varepsilon.$$

DEFINITION 1. We say that the d-dimensional continuous stochastic process X: $(-\infty, T] \times \mathscr{C}_2^0 \to R^d$ is a *strong solution to equation* (2) *with a given process* $w(t)$ if conditions (A1), (A2), (A6) are satisfied and equation (2) is valid with probability 1 for all $t \in (-\infty, T]$.

The uniqueness of strong solutions is understood in the sense of trajectories, that is, if for every two strong solutions X and $\tilde{X}$ to equation (2) defined on the same probability space we have

$$P\big(\sup_{t \in (-\infty, T]} |X(t, w) - \tilde{X}(t, w)| > 0\big) = 0.$$

DEFINITION 2. We recall that the absolutely continuous stochastic process X^n: $(-\infty, T] \times \mathscr{C}_2^0 \to R^d$ is a *solution to equation* (2^n) if conditions (A2), (A3) are satisfied and equation (2^n) is valid with probability 1 for all $t \in (-\infty, T]$.

Let us notice that conditions (A2)–(A7) ensure the existence and uniqueness of the strong solution Y to equation (3). Indeed (see [6], Sections 5 and 7), under conditions (A2)–(A5) there exists a strong solution to equation (2). The uniqueness may be derived from the proof of Theorem 11, Section 10 in [6], for the multidimensional case with an additional remark that measurability is a consequence of continuous dependence of solutions on the initial condition. Now we consider the term

$$(*) \qquad b^i(Y_t(\cdot, w)) + \tilde{D}_j\sigma^{ip}(Y_t(\cdot, w))\sigma^{jp}(Y_t(\cdot, w))$$

in equation (3). Since $\tilde{D}_j\sigma^{ip}$ is a measure, we have

$$|\tilde{D}_j\sigma^{ip}(\varphi)\sigma^{jp}(\varphi)| \leqslant \bar{C}|\sigma^{jp}(\varphi)|,$$

$$|\tilde{D}_j\sigma^{ip}(\varphi)\sigma^{jp}(\varphi) - \tilde{D}_j\sigma^{ip}(\psi)\sigma^{jp}(\psi)| \leqslant \bar{C}|\sigma^{jp}(\varphi) - \sigma^{jp}(\psi)|,$$

where $\bar{C}$ is a constant. Thus conditions (A4) and (A5) are fulfilled for the term ($*$). It is obvious that other conditions are also fulfilled and equation (3) has also exactly one strong solution.

Moreover, for every $n \in N$, under condition (A4) there exists exactly one solution to the ordinary differential equation (2^n) with a delay argument (see [4] and [6]).

The following limit Z_t^n in the space $\mathscr{C}_-$ is understood in the norm sense:

$$Z_t^n(\cdot) = \lim_{h\to 0} \frac{X_{t+h}^n(\cdot)-X_t^n(\cdot)}{h},$$

that is

$$\max_{-\infty<\theta\leq 0} |h^{-1}(X_{t+h}^n(\theta)-X_t^n(\theta))-Z_t^n(\theta)| \to 0 \text{ as } h\to 0, \quad \theta\in J$$

(we choose an appropriate norm on $[-r, 0]$ from the family of seminorms). We have

$$Z_t^n(\theta) = \frac{d}{dt}X_t^n(\theta) = \dot{X}^n(t+\theta), \quad \theta\in J.$$

Putting $u = t+\theta$ we have

$$\dot{X}^n(t+\cdot)\colon\ \theta \to \frac{dX^n(u)}{du}, \quad -\infty < u \leq t.$$

Moreover, it is obvious that

$$\frac{dX_t(\theta)}{dt} = \frac{dX_t(\theta)}{d\theta}$$

because $X_t(\theta) = X(t+\theta)$.

If we understand t as a variable we have

$$\dot{X}_\cdot^n\colon\ t\in[0, T] \to \dot{X}_t^n \in \mathscr{C}_-.$$

4. The Approximation Theorem. We shall prove the following

THEOREM 1. *Let the conditions* (A2)–(A5) *and* (A7) *be fulfilled. Let $B^n(t, w)$ be the approximation of type* (1) *of the Wiener process. We assume that X^n and Y are solutions to* (2^n) *and* (3), *respectively, with a constant initial stochastic process. Then conditions* (A1) *and* (A6) *are satisfied and, for every $T > 0$,*

$$\lim_{n\to\infty} \sup_{0\leq t\leq T} \mathrm{E}[|X^n(t, w)-Y(t, w)|^2] = 0.$$

Proof. The assumptions of the theorem ensure the existence and uniqueness of the solutions to equations (2^n) and (3). For every $i = 1, \ldots, d$ we write the subtraction of equations (2^n) and (3):

$$X^{n,i}(t, w)-Y^i(t, w) = H_1(t)+H_2(t)+H_3+H_4(t),$$

where

$$\begin{aligned} H_1(t) = {} & \sum_{p=1}^{m} \int_{t_n^-}^{t} \sigma^{ip}(X_s^n(\cdot, w))\dot{B}^{n,p}(s, w)ds - \sum_{p=1}^{m} \int_{t_n^-}^{t} \sigma^{ip}(Y_s(\cdot, w))dw^p(s) \\ & -\frac{1}{2}\sum_{p=1}^{m}\sum_{j=1}^{d}\int_{t_n^-}^{t} \tilde{D}_j\sigma^{ip}(Y_s(\cdot, w))\sigma^{jp}(Y_s(\cdot, w))ds \\ = {} & H_{11}(t) - \sum_{p=1}^{m} H_{12}^p(t) - H_{13}(t), \end{aligned}$$

$$H_2(t) = \sum_{p=1}^{m} \int_{1/2^n}^{t_n^-} \sigma^{ip}(X_s^n(\cdot, w))\dot{B}^{n,p}(s, w)ds - \sum_{p=1}^{m} \int_{1/2^n}^{t_n^-} \sigma^{ip}(Y_s(\cdot, w))dw^p(s)$$

$$-\frac{1}{2}\sum_{p=1}^{m}\sum_{j=1}^{d}\int_{1/2^n}^{t_n^-} \tilde{D}_j\sigma^{ip}(Y_s(\cdot, w))\sigma^{jp}(Y_s(\cdot, w))ds,$$

$$H_3 = \sum_{p=1}^{m} \int_{0}^{1/2^n} \sigma^{ip}(X_s^n(\cdot, w))\dot{B}^{n,p}(s, w)ds - \sum_{p=1}^{m} \int_{0}^{1/2^n} \sigma^{ip}(Y_s(\cdot, w))dw^p(s)$$

$$-\frac{1}{2}\sum_{p=1}^{m}\sum_{j=1}^{d}\int_{0}^{1/2^n} \tilde{D}_j\sigma^{ip}(Y_s(\cdot, w))\sigma^{jp}(Y_s(\cdot, w))ds,$$

$$H_4(t) = \int_0^t b^i(X_s^n(\cdot, w))ds - \int_0^t b^i(Y_s(\cdot, w))ds.$$

Further, c_l, $l = 0, 1, \ldots, 25$, denote some positive constants.

From (2^n) we have

$$|X^{n,i}(t, w) - X^{n,i}(s, w)| \leqslant c_0\Big(\sum_{p=1}^{m}\int_s^t |\dot{B}^{n,p}(u, w)|du + (t-s)\Big). \tag{4}$$

From the boundedness of σ we obtain

$$\mathrm{E}[\sup_{0\leqslant t\leqslant T} |H_{11}(t)|^2] = \mathrm{E}[\sup_{0\leqslant t\leqslant T} |\sum_{p=1}^{m}\int_{t_n^-}^{t} \sigma^{ip}(X_s^n(\cdot, w))\dot{B}^{n,p}(s, w)ds|^2]$$

$$\leqslant \mathrm{E}[\sup_{0\leqslant t\leqslant T} |\sum_{p=1}^{m}\big(2^{-n}\sup_s \sigma^{ip}(X_s^n(\cdot, w))\big)^2\big(\int_{t_n^-}^{t_n^+} \dot{B}^{n,p}(s, w)^2 ds\big)|]$$

$$\leqslant c_1\cdot 2^{-n}\mathrm{E}[\sup_{0\leqslant t\leqslant T} \big(\int_{t_n^-}^{t_n^+} |\dot{B}^{n,p}(s, w)|ds\big)^2]$$

and we estimate

$$\mathrm{E}[\sup_{0\leqslant t\leqslant T} \big(\int_{t_n^-}^{t_n^+} |\dot{B}^{n,p}(s, w)|ds\big)^2] \leqslant \mathrm{E}[\sup_k \big(\int_{k/2^n}^{(k+1)/2^n} |\dot{B}^{n,p}(s,w)|ds\big)^4]^{1/2}$$

$$\leqslant \mathrm{E}[\sum_{k=1}^{m(T)}\big(\int_{k/2^n}^{(k+1)/2^n} |\dot{B}^{n,p}(s, w)|ds\big)^4]^{1/2} = \Big[\sum_{k=1}^{m(T)}\mathrm{E}[\int_{k/2^n}^{(k+1)/2^n} |\dot{B}^{n,p}(s, w)|ds]^4\Big]^{1/2}$$

$$= \left[\sum_{k=1}^{m(T)}\mathrm{E}\left[B^{n,p}\left(\frac{k+1}{2^n}, w\right) - B^{n,p}\left(\frac{k}{2^n}, w\right)\right]^4\right]^{1/2}$$

$$= \left[\sum_{k=1}^{m(T)}\mathrm{E}\left[w^p\left(\frac{k+1}{2^n}\right) - w^p\left(\frac{k}{2^n}\right)\right]^4\right]^{1/2}$$

$$\leqslant \big(m(T)3(1/2^n)^2\big)^{1/2} \leqslant c_2\big(2^n(1/2^n)^2\big)^{1/2} = c_2(1/2^n)^{1/2}.$$

Therefore

$$\mathrm{E}[\sup_{0\leqslant t\leqslant T} |H_{11}(t)|^2] \leqslant c_3(1/2^n)^{1/2} \to 0 \quad \text{as } n\to\infty.$$

Further we estimate

$$\mathrm{E}[\sup_{0\leqslant t\leqslant T} |H_{12}^p(t)|^2] = \mathrm{E}\Big[\sup_{0\leqslant t\leqslant T} \Big|\int_{t_n^-}^{t} \sigma^{ip}(Y_{s_n^-}(\cdot, w))dw^p(s)$$

$$+\int_{t_n^-}^{t} \big(\sigma^{ip}(Y_s(\cdot, w))-\sigma^{ip}(Y_{s_n^-}(\cdot, w))\big)dw^p(s)\Big|^2\Big]$$

$$\leqslant \mathrm{E}[\sup_{0\leqslant t\leqslant T} |\int_{t_n^-}^{t} \sigma^{ip}(Y_{s_n^-}(\cdot, w))dw^p(s)|^2]$$

$$+\mathrm{E}\Big[\sup_{0\leqslant k\leqslant m(T)} \sup_{0\leqslant t\leqslant 1/2^n} \Big|\int_{k/2^n}^{k/2^n+t} \big(\sigma^{ip}(Y_s(\cdot, w))-\sigma^{ip}(Y_{s_n^-}(\cdot, w))\big)dw^p(s)\Big|^2\Big]$$

$$= \hat{H}_1(t)+\hat{H}_2(t).$$

From the Hölder inequality, the well-known inequality for the Itô integrals, the assumptions on σ^{ip} and (7.32) in [5] we have

$$\hat{H}_1(t) \leqslant \mathrm{E}[\sup_{0\leqslant t\leqslant T} \int_{t_n^-}^{t} \sigma^{ip}(Y_{s_n^-}(\cdot, w))^2 ds] \leqslant c_4(\int_{t_n^-}^{t} ds)^2$$

$$\leqslant c_4(1/2^n)^2 \to 0 \quad \text{as } n\to\infty.$$

Let w' be the Wiener process translated in time, i.e.,

$$w'(t) = w(t+k/2^n)-w(k/2^n).$$

It is obvious that the process $Y_{k/2^n+t}$ may be considered as the solution to (3) after replacing w by w' and Y_0^i by $Y_{k/2^n}^i$ defined by the formula

$$Y_{k/2^n}(\theta) = Y(k/2^n+\theta).$$

Let $\mathscr{B}$ be the smallest σ-algebra such that $Y_{k/2^n}$ is a stochastic process with respect to it. Let Y_t be a solution to (3) with the initial condition $Y_0^\xi = \xi$. Let E′ denote the expectation and conditional expectation standing for the integration with respect to w'. Since the increments of the Wiener process are stationary, we may replace the computing of E of the original process by the computing of E′ of the translated process. Therefore, using (A4) we have

$$\hat{H}_2(t) \leqslant \sum_{k=0}^{m(T)} \mathrm{E}'\Big\{\mathrm{E}\Big[\sup_{0\leqslant t\leqslant 1/2^n} \Big(\int_0^t \big(\sigma^{ip}(Y_{k/2^n+s}(\cdot, w))-\sigma^{ip}(Y_{k/2^n}(\cdot, w))\big)dw^p(s)\Big)^2\Big]\Big|\mathscr{B}\Big\}$$

$$\leqslant c_5 \sum_{k=0}^{m(T)} \mathrm{E}'\Big\{\mathrm{E}\Big[\int_0^{1/2^n} \big(\sigma^{ip}(Y_s(\cdot, w))-\sigma^{ip}(\xi)\big)^2 ds\Big]\Big|\mathscr{B}\Big\}$$

$$\leqslant c_6\big(m(T)+1\big)\sum_{j=1}^{d} \mathrm{E}'\{\mathrm{E}[\int_0^{1/2^n}(L^1\int_{-\infty}^{0}|Y_s^\xi(u)-Y_{k/2^n}(u)|^2 dK(u)$$

$$+L^2|Y^\xi(s)-Y(k/2^n)|^2)ds]\,|\mathscr{B}\},$$

where $\xi(\theta)=Y(k/2^n+\theta)$ and, as in (7.57) of [5], we get

$$\hat{H}_2(t)\leqslant c_7\frac{1}{2^n}\to 0 \quad \text{as } n\to\infty.$$

Therefore, we obtain

$$|Y^i(s)-Y^i(k/2^n)|^2=\Big|\int_0^s b^i\big(Y_u^\xi(\cdot,w)\big)du+\sum_{p=1}^{m}\int_0^s\sigma^{ip}\big(Y_u^\xi(\cdot,w)\big)dw^p(u)\Big|^2.$$

From (A7) we have

$$\mathrm{E}[\sup_{0\leqslant t\leqslant T}|H_{13}(t)|^2]=\tfrac{1}{2}\mathrm{E}[\sup_{0\leqslant t\leqslant T}\Big|\sum_{p=1}^{m}\sum_{j=1}^{d}\int_{t_n^-}^{t}\tilde{D}_j\sigma^{ip}\big(Y_s(\cdot,w)\big)\sigma^{jp}\big(Y_s(\cdot,w)\big)\Big|^2 ds]$$

$$\leqslant c_8\Big|\int_{t_n^-}^{t}ds\Big|^2\leqslant c_9(1/2^n)^2\to 0 \quad \text{as } n\to\infty.$$

To estimate $H_2(t)$ we shall first present the integral $\int_{t_n^+}^{t_n^-}$ as a sum $\sum_{k=1}^{m(t)-1}\int_{k/2^n}^{(k+1)/2^n}$. At first let us observe that for $-\infty<u\leqslant 0$

$$\begin{aligned}X^{n,i}(t+u,w)&=X^{n,i}(0)+\int_0^{t+u}b^i\big(X_s^n(\cdot,w)\big)ds\\&\quad+\sum_{p=1}^{m}\int_0^{t+u}\sigma^{ip}\big(X_s^n(\cdot,w)\big)\dot{B}^{n,p}(s,w)ds\\&=X^{n,i}(0)+\int_{-u}^{t}b^i\big(X_{s+u}^n(\cdot,w)\big)ds\\&\quad+\sum_{p=1}^{m}\int_{-u}^{t}\sigma^{ip}\big(X_{s+u}^n(\cdot,w)\big)\dot{B}^{n,p}(s+u,w)ds\\&=X^{n,i}(u)+X^{n,i}(0)-X^{n,i}(0)+\int_0^t b^i\big(X_{s+u}^n(\cdot,w)\big)ds\\&\quad+\sum_{p=1}^{m}\int_0^t\sigma^{ip}\big(X_{s+u}^n(\cdot,w)\big)\dot{B}^{n,p}(s+u,w)ds,\end{aligned}$$

where we used the assumption that the initial process is constant. Therefore, for $-\infty<\tau\leqslant 0$ we have

$$(**)\quad \frac{dX^{n,i}(t+\cdot,w)}{dt}(\tau)=b^i\big(X_{t+\tau}^n(\cdot,w)\big)+\sum_{p=1}^{m}\sigma^{ip}\big(X_{t+\tau}^n(\cdot,w)\big)\dot{B}^{n,p}(t+\tau,w).$$

Now we take into consideration

$$
\begin{aligned}
H = & \int_{k/2^n}^{(k+1)/2^n} \sigma^{ip}\big(X_s^n(\cdot, w)\big)\dot{B}^{n,p}(s, w)ds = \sigma^{ip}\big(X_s^n(\cdot, w)\big)B^{n,p}(s, w)\Big|_{k/2^n}^{(k+1)/2^n} \\
& + \sigma^{ip}\big(X_{k/2^n}^n(\cdot, w)\big)B^{n,p}\left(\frac{k+1}{2^n}, w\right) - \sigma^{ip}\big(X_{k/2^n}^n(\cdot, w)\big)B^{n,p}\left(\frac{k+1}{2^n}, w\right) \\
& - \int_{k/2^n}^{(k+1)/2^n} D\sigma^{ip}\big(X_s^n(\cdot, w)\big)\frac{dX^n(s+\cdot, w)}{ds}B^{n,p}(s, w)ds \\
= & \ \sigma^{ip}\big(X_{k/2^n}^n(\cdot, w)\big)\left(B^{n,p}\left(\frac{k+1}{2^n}, w\right) - B^{n,p}\left(\frac{k}{2^n}, w\right)\right) \\
& + \Big(\sigma^{ip}\big(X_{(k+1)/2^n}^n(\cdot, w)\big) - \sigma^{ip}\big(X_{k/2^n}^n(\cdot, w)\big)\Big)B^{n,p}\left(\frac{k+1}{2^n}, w\right) \\
& - \int_{k/2^n}^{(k+1)/2^n} D\sigma^{ip}\big(X_s^n(\cdot, w)\big)\frac{dX^n(s+\cdot, w)}{ds}B^{n,p}(s, w)ds \\
= & \ J_1(k) + \int_{k/2^n}^{(k+1)/2^n} D\sigma^{ip}\big(X_s^n(\cdot, w)\big)\Big(\tilde{b}\big(X_s^n(\cdot, w)\big) \\
& + \sum_{p=1}^{m} \tilde{A}_s^{pn}\Big)\left(B^{n,p}\left(\frac{k+1}{2^n}, w\right) - B^{n,p}(s, w)\right)ds \\
= & \ J_1(k) + \sum_{j=1}^{d} \int_{k/2^n}^{(k+1)/2^n} \int_{-\infty}^{0} b^j\big(X_{s+v}^n(\cdot, w)\big)\mu_{s,w,X}^{ipj}(dv) \\
& \times \left(B^{n,p}\left(\frac{k+1}{2^n}, w\right) - B^{n,p}(s, w)\right)ds \\
& + \sum_{j=1}^{d} \int_{k/2^n}^{(k+1)/2^n} \int_{-\infty}^{0} \sum_{p=1}^{m} \sigma^{jp}\big(X_{s+v}^n(\cdot, w)\big)\dot{B}^{n,p}(s+v, w)\tilde{\mu}_{s,w,X}^{ipj}(dv) \\
& \times \left(B^{n,p}\left(\frac{k+1}{2^n}, w\right) - B^{n,p}(s, w)\right)ds \\
& + \sum_{j=1}^{d} \int_{k/2^n}^{(k+1)/2^n} \sum_{p=1}^{m} \tilde{D}_j\sigma^{ip}\big(X_s^n(\cdot, w)\big)\sigma^{jp}\big(X_s^n(\cdot, w)\big)\dot{B}^{n,p}(s, w) \\
& \times \left(B^{n,p}\left(\frac{k+1}{2^n}, w\right) - B^{n,p}(s, w)\right)ds \\
= & \ J_1(k) + \sum_{r=1}^{3} \sum_{j=1}^{d} J_{r2}^j(k).
\end{aligned}
$$

As in [5], we write

$$
\begin{aligned}
J_1(k) = {} & \sigma^{ip}(X^n_{k/2^n-\delta}(\cdot, w))\left(w^p\left(\frac{k+1}{2^n}\right)-w^p\left(\frac{k}{2^n}\right)\right)+\big(\sigma^{ip}(X^n_{k/2^n}(\cdot, w)) \\
& -\sigma^{ip}(X^n_{k/2^n-\delta}(\cdot, w))\big)\left(B^{n,p}\left(\frac{k+1}{2^n}, w\right)-B^{n,p}\left(\frac{k}{2^n}, w\right)\right) \\
& +\sigma^{ip}(X^n_{k/2^n-\delta}(\cdot, w))\left(B^{n,p}\left(\frac{k+1}{2^n}, w\right)-w^p\left(\frac{k+1}{2^n}\right)\right) \\
& +\sigma^{ip}(X^n_{k/2^n-\delta}(\cdot, w))\left(w^p\left(\frac{k}{2^n}\right)-B^{n,p}\left(\frac{k}{2^n}, w\right)\right) \\
= {} & J_{11}(k)+\ldots+J_{14}(k)
\end{aligned}
$$

and

$$
\begin{aligned}
H_2(t) = {} & \sum_{p=1}^{m}\Bigg(\Big(\sum_{k=1}^{m(t)-1} J_{11}(k)-\int_{1/2^n}^{t_n^-}\sigma^{ip}(Y_s(\cdot, w))dw^p(s)\Big) \\
& +\sum_{k=1}^{m(t)-1} J_{12}(k)+\sum_{k=1}^{m(t)-1} J_{13}(k)+\sum_{k=1}^{m(t)-1} J_{14}(k) \\
& +\sum_{j=1}^{d}\Big(\sum_{k=1}^{m(t)-1}\sum_{r=1}^{3} J^j_{r2}(k)-\frac{1}{2}\int_{1/2^n}^{t_n^-}\tilde{D}_j\sigma^{ip}(Y_s(\cdot, w))\sigma^{jp}(Y_s(\cdot, w))ds\Big)\Bigg) \\
= {} & \sum_{p=1}^{m}\big(I_1(t)+\ldots+I_5(t)\big).
\end{aligned}
$$

Now we estimate each term among $I_1, \ldots, I_5$, successively. For every $t_1 \in [0, T]$, using the martingale inequality, using (4), (44) and (7.2) in [5], we obtain

$$
\begin{aligned}
\mathrm{E}[\sup_{0\leqslant t\leqslant t_1}|I_1(t)|^2] &= \mathrm{E}\Big[\sup_{0\leqslant t\leqslant t_1}\Big|\int_{1/2^n}^{t_n^-}\big(\sigma^{ip}(X^n_{s_n^- -\delta}(\cdot, w))-\sigma^{ip}(Y_s(\cdot, w))\big)dw^p(s)\Big|^2\Big] \\
&\leqslant c_{10}\Big(\int_0^{t_1}\mathrm{E}[|Y(s, w)-X^n(s, w)|^2]ds+\int_0^{t_1}\Big(\int_{-\infty}^{0}\mathrm{E}[|Y_{s_1}(\cdot, w) \\
&\qquad -X^n_{s_1}(\cdot, w)|^2]dK(s_1)\Big)ds+\frac{n}{2^n}\Big).
\end{aligned}
$$

Now from (4) and the Hölder inequality we get

$$
\begin{aligned}
\mathrm{E}[\sup_{0\leqslant t\leqslant t_1}|I_2(t)|^2] \leqslant \mathrm{E}\Bigg[&\sum_{k=1}^{m(T)-1}\big(\sigma^{ip}(X^n_{k/2^n}(\cdot, w))-\sigma^{ip}(X^n_{k/2^n-\delta}(\cdot, w))\big)^2 \\
&\times\sum_{k=1}^{m(T)-1}\left(B^{n,p}\left(\frac{k+1}{2^n}, w\right)-B^{n,p}\left(\frac{k}{2^n}, w\right)\right)^2\Bigg]
\end{aligned}
$$

$$\leqslant c_{11}\Big(m(T)\sum_{k=1}^{m(T)-1}\mathrm{E}\Big[L^1\Big(\int_{-\infty}^{0}|X^n_{k/2^n}(\theta)-X^n_{k/2^n-\delta}(\theta)|\,dK\theta$$

$$+L^2\Big|X\Big(\frac{k}{2^n},w\Big)-X^n\Big(\frac{k}{2^n}-\delta,w\Big)\Big|\Big)^4\Big]m(T)\sum_{k=1}^{m(T)-1}\mathrm{E}\Big[\Big|B^{n,p}\Big(\frac{k+1}{2^n},w\Big)$$

$$-B^{n,p}\Big(\frac{k}{2^n},w\Big)\Big|^4\Big]\Big)^{1/2}\leqslant c_{12}\Big(m(T)\sum_{k=1}^{m(T)-1}\mathrm{E}\Big[\Big|X^n\Big(\frac{k}{2^n},w\Big)-X^n\Big(\frac{k}{2^n}-\delta,w\Big)\Big|^4\Big]$$

$$\times m(T)\sum_{k=1}^{m(T)-1}\mathrm{E}\Big[B^{n,p}\Big(\frac{k+1}{2^n},w\Big)-B^{n,p}\Big(\frac{k}{2^n},w\Big)\Big]^4\Big)^{1/2}$$

$$\leqslant c_{13}\Big(m(T)^2\Big(\frac{1}{n}\frac{1}{2^n}\Big)^2 m(T)^2\Big(\Big(\frac{1}{2^n}\Big)^2\Big)\Big)^{1/2}\leqslant c_{14}\frac{1}{n}\to 0\quad \text{as } n\to\infty.$$

Let

$$\eta_l(w)=\sum_{k=1}^{l}\big(\sigma^{ip}(X^n_{k/2^n-\delta}(\cdot,w))B^{n,p}(0,\theta_{(k+1)/2^n}w)\big),$$

where $\theta_t w(s)=w(t+s)-w(t)$. It is obvious that η_l is an $\mathscr{F}_l$-martingale for $\mathscr{F}_l=\mathscr{B}_{(l+2)/2^n}$.

Since $I_3(t)$ as well as $I_4(t)$ written as $\eta_l(w)$ are $\mathscr{F}_l$-martingales, from the martingale inequality we have

$$\mathrm{E}[\sup_{0\leqslant t\leqslant T}|I_3(t)|^2]\leqslant c_{15}n^{-1}\to 0\quad \text{as } n\to\infty,$$

$$\mathrm{E}[\sup_{0\leqslant t\leqslant T}|I_4(t)|^2]\leqslant c_{16}n^{-1}\to 0\quad \text{as } n\to\infty.$$

We write $I_5(t)=\sum_{j=1}^{d}J_5^j(t)$, where

$$J_5^j(t)=\sum_{k=1}^{m(t)-1}\sum_{r=1}^{3}J^j_{r2}(k)-\frac{1}{2}\sum_{k=1}^{m(t)-1}\int_{k/2^n}^{(k+1)/2^n}\tilde{D}_j\sigma^{ip}(Y_s(\cdot,w))\sigma^{jp}(Y_s(\cdot,w))\,ds$$

$$=\sum_{k=1}^{m(t)-1}\int_{k/2^n}^{(k+1)/2^n}\big(\tilde{D}_j\sigma^{ip}(X^n_s(\cdot,w))\sigma^{jp}(X^n_s(\cdot,w))$$

$$-\tilde{D}_j\sigma^{ip}(X^n_{k/2^n}(\cdot,w))\sigma^{jp}(X^n_{k/2^n}(\cdot,w))\big)\dot{B}^{n,p}(s,w)\Big(B^{n,p}\Big(\frac{k+1}{2^n},w\Big)$$

$$-B^{n,p}(s,w)\Big)ds+\sum_{k=1}^{m(t)-1}\int_{k/2^n}^{(k+1)/2^n}\int_{-\infty}^{0}b^j(X^n_{s+v}(\cdot,w))\,\mu^{ipj}_{s,w,X}(dv)$$

$$\times\Big(B^{n,p}\Big(\frac{k+1}{2^n},w\Big)-B^{n,p}(s,w)\Big)ds$$

$$+\sum_{k=1}^{m(t)-1}\int_{k/2^n}^{(k+1)/2^n}\Big(\tilde{D}_j\sigma^{ip}(X^n_{k/2^n}(\cdot,w))\sigma^{jp}(X^n_{k/2^n}(\cdot,w))$$

$$\times\left(\dot{B}^{n,p}(s,w)\left(B^{n,p}\left(\frac{k+1}{2^n},w\right)-B^{n,p}(s,w)\right)-c_p(2^{-n},n)\right)\Big)ds$$

$$+\frac{1}{2}\sum_{k=1}^{m(t)-1}\int_{k/2^n}^{(k+1)/2^n}\big(\tilde{D}_j\sigma^{ip}(X^n_{k/2^n}(\cdot,w))\sigma^{jp}(X^n_{k/2^n}(\cdot,w))$$

$$-\tilde{D}_j\sigma^{ip}(Y_s(\cdot,w))\sigma^{jp}(Y_s(\cdot,w))\big)ds$$

$$+\sum_{k=1}^{m(t)-1}\int_{k/2^n}^{(k+1)/2^n}\frac{1}{2^n}\tilde{D}_j\sigma^{ip}(X^n_{k/2^n}(\cdot,w))\sigma^{jp}(X^n_{k/2^n}(\cdot,w))$$

$$\times\left(c_p\left(\frac{1}{2^n},n\right)-\frac{1}{2}\right)+\sum_{k=1}^{m(t)-1}J^i_{22}(k)$$

$$=I_{51}(t)+\ldots+I_{56}(t),$$

where

$$c_j(t,n)=t^{-1}\mathrm{E}\Big[\int_0^t\dot{B}^{n,j}(s,w)\big(B^{n,j}(t,w)-B^{n,j}(s,w)\big)ds\Big],\qquad \lim_{n\to\infty}c_j(1/2^n,n)=\tfrac{1}{2}$$

(see [5], Lemma 7.1).

Since X^n is uniformly continuous on the finite interval, X^n_s is continuous as a function of the variable s with the functional values and we may estimate (analogously as in [5])

$$\mathrm{E}\big[\sup_{0\leqslant t\leqslant T}|I_{51}(t)|^2\big]\leqslant c_{17}\frac{n}{2^n}\to 0\quad \text{as } n\to\infty,$$

$$\mathrm{E}\big[\sup_{0\leqslant t\leqslant T}|I_{52}(t)|^2\big]\leqslant c_{18}\frac{n^3}{2^n}\to 0\quad \text{as } n\to\infty,$$

$$\mathrm{E}\big[\sup_{0\leqslant t\leqslant T}|I_{53}(t)|^2\big]\to 0\quad \text{as } n\to\infty,$$

$$\mathrm{E}\big[\sup_{0\leqslant t\leqslant T}|I_{55}(t)|^2\big]\leqslant c_{19}\left(c_p\left(\frac{1}{2^n},n\right)-\frac{1}{2}\right)^2\to 0\quad \text{as } n\to\infty,$$

$$\mathrm{E}\big[\sup_{0\leqslant t\leqslant T}|I_{56}(t)|^2\big]\to 0\quad \text{as } n\to\infty.$$

Further, using (A4) and (7.68) of [5], for every $t_1 \in [0, T]$ we obtain

$$\mathrm{E}[\sup_{0\leqslant t\leqslant t_1} |I_{54}(t)|^2] \leqslant c_{20}\Big(\int_0^{t_1} \mathrm{E}[|Y(s, w)-X^n(s_n^-, w)|^2]ds$$

$$+\int_0^{t_1}\Big(\int_{-\infty}^0 \mathrm{E}[|Y_{s_1}(\cdot, w)-X^n_{s\bar{1},n}(\cdot, w)|^2]dK(s_1)\Big)ds\Big)$$

$$\leqslant c_{21}\Big(\int_0^{t_1} \mathrm{E}[|Y(s, w)-X^n(s, w)|^2]ds+\int_0^{t_1} [\mathrm{E}|X^n(s, w)-X^n(s_n^-, w)|^2]ds$$

$$+\int_0^{t_1}\Big(\int_{-\infty}^0 \mathrm{E}[|Y(s_1, w)-X^n_{s_1}(\cdot, w)|^2]dK(s_1)\Big)ds$$

$$+\int_0^{t_1}\Big(\int_{-\infty}^0 \mathrm{E}[|X^n_{s_1}(\cdot, w)-X^n_{s\bar{1},n}(\cdot, w)|^2]dK(s_1)\Big)ds\Big)$$

$$\leqslant c_{22}\left(\int_0^{t_1} \mathrm{E}[|Y(s, w)-X^n(s, w)|^2]ds\right.$$

$$\left.+\int_0^{t_1}\Big(\int_{-\infty}^0 \mathrm{E}[|Y_{s_1}(\cdot, w)-X^n_{s_1}(\cdot, w)|^2]dK(s_1)\Big)ds+\frac{2n}{2^n}\right).$$

It is obvious that $|H_3| \leqslant \sup_{0\leqslant t\leqslant T} |H_1(t)|$; hence

$$\mathrm{E}[|H_3|^2] \to 0 \quad \text{as } n\to\infty.$$

For every $t_1 \in [0, T]$, using the Hölder inequality and (A4) we obtain

$$\mathrm{E}[\sup_{0\leqslant t\leqslant t_1} |H_4(t)|^2] \leqslant c_{23}\Big(\int_0^{t_1} \mathrm{E}[|Y(s, w)-X^n(s, w)|^2]ds$$

$$+\int_0^{t_1}\Big(\int_{-\infty}^0 \mathrm{E}[|Y_{s_1}(\cdot, w)-X^n_{s_1}(\cdot, w)|^2]dK(s_1)\Big)ds\Big).$$

Now we shall use the generalized Gronwall lemma ([8], Lemma 4.13).

LEMMA. *Let k_0, k_1, k_2 be nonnegative constants, let $u(t)$ be a bounded function for every $t \in (-\infty, T]$ and $v(t)$ be a nonnegative integrable function. We assume that $K(s)$ is a nondecreasing nonnegative right-continuous function such that $0 \leqslant K(s) \leqslant 1$ and that*

$$u(t) \leqslant k_0+k_1\int_0^t v(s)u(s)ds+k_2\int_0^t v(s)\Big(\int_{-\infty}^t u(s_1)dK(s_1)\Big)ds.$$

Then

$$u(t) \leqslant k_0 \exp\Big((k_1+k_2)\int_0^t v(s)ds\Big).$$

We use this lemma for

$$u(t) = \mathrm{E}\big[\sup_{0\leqslant s\leqslant t} |Y(t, w) - X^n(t, w)|^2\big] \quad \text{and} \quad v(t) = 1.$$

From this lemma and from the uniform convergence of $o(1)$ for $t_1 \in [0, T]$, using all the above estimations, we obtain

$$\begin{aligned}\mathrm{E}\big[\sup_{0\leqslant t\leqslant t_1} |X^n(t, w) - Y(t, w)|^2\big] &\leqslant o(1) + c_{24}\Big(\int_0^{t_1} \mathrm{E}\big[\sup_{0\leqslant s\leqslant t} |X^n(s, w) - Y(s, w)|^2\big] ds \\ &\quad + \int_0^{t_1}\Big(\int_{-\infty}^{0} \mathrm{E}\big[\sup_{0\leqslant s\leqslant t} |X^n_{s_1}(\cdot, w) - Y_{s_1}(\cdot, w)|^2\big] dK(s_1)\Big) ds\Big)\end{aligned}$$

and in consequence we have

$$\mathrm{E}\big[\sup_{0\leqslant t\leqslant t_1} |X^n(t, w) - Y(t, w)|^2\big] \leqslant o(1)\exp(c_{25}t) \to 0 \quad \text{as } n \to \infty.$$

This completes the proof.

5. Example. We consider the following equation:

$$dX(t) = B(X_t)dt + \Sigma(X_t)dw(t), \quad X_0(\theta, w) = \eta(\theta, w) \text{ for } \theta \in J,$$

where for some constants $b_0, b_1, \sigma_0, \sigma_1$ we define

$$B, \Sigma\colon \mathscr{C}_- \to R, \quad B(\varphi) = b_0(\varphi(0)) + b_1(\varphi(-r)),$$

$$\Sigma(\varphi) = \sigma_0(\varphi(0)) + \sigma_1(\varphi(-r)).$$

We note that $\varphi(0) = X_t(0) = X(t)$, $\varphi(-r) = X_t(-r) = X(t-r)$ and

$$dX(t) = \big(b_0(X(t)) + b_1(X(t-r))\big)dt + \big(\sigma_0(X(t)) + \sigma_1(X(t-r))\big)dw(t),$$

$$X_0 = \eta.$$

We take the piecewise linear approximation $B_n(t, w)$ of $B(t, w) = w(t)$ given by (1), and then equation (3) is of the form

$$\begin{aligned}dY(t) &= \big(b_0(Y(t)) + b_1(Y(t-r))\big)dt + \big(\sigma_0(Y(t)) + \sigma_1(Y(t-r))\big)dw(t) \\ &\quad + \tfrac{1}{2}\sigma_0\big(\sigma_0(Y(t)) + \sigma_1(Y(t-r))\big)dt, \quad Y_0 = \eta,\end{aligned}$$

because $\sigma_0(X(t))$ is the only term for which the support of the measure contains zero. Therefore $\mu(\{0\}) = \sigma_0$.

REFERENCES

[1] P. Acquistapace and B. Terreni, *An approach to Itô linear equations in Hilbert spaces by approximation of white noise with coloured noise*, Stoch. Anal. Appl. 2 (1984), pp. 131–186.

[2] G. Da Prato, *Stochastic differential equations with noncontinuous coefficients in Hilbert spaces*, Rend. Sem. Mat. Univ. Politc. Torino, Numero speciale (1982), pp. 73–85.

[3] H. Doss, *Liens entre équations différentielles stochastiques et ordinaires*, Ann. Inst. H. Poincaré 13.2 (1977), pp. 99–125.

[4] J. Hale, *Theory of Functional Differential Equations*, Springer-Verlag, New York–Heidelberg–Berlin 1977.

[5] N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*, North-Holland Publ. Co., Amsterdam 1981.

[6] K. Itô and M. Nisio, *On stationary solutions of a stochastic differential equation*, J. Math. Kyoto Univ. 4.1 (1964), pp. 1–75.

[7] F. Koneczny, *On Wong–Zakai approximation of stochastic differential equations*, J. Multivariate Anal. 13 (1983), pp. 605–611.

[8] R. Liptser and A. Shiryayev, *Statistics of Random Processes*, Springer-Verlag, New York–Heidelberg–Berlin 1977.

[9] S. E. A. Mohammed, *Stochastic Functional Differential Equations*, Pitman Publishing Inc., Marshfield 1984.

[10] S. Nakao and J. Yamato, *Approximation theorem on stochastic differential equations*, Proc. Intern. Symp. SDE, Kyoto 1976.

[11] D. W. Strook and S. R. S. Varadhan, *On the support of diffusion processes with applications to the strong maximum principle*, Proc. 6-th Berkeley Symp. on Math. Stat. and Prob. 3 (1972), pp. 333–359.

[12] E. Wong and M. Zakai, *On the convergence of ordinary integrals to stochastic integral*, Ann. Math. Statist. 36 (1965), pp. 1560–1564.

Institute of Mathematics
Warsaw Technical University
pl. Politechniki 1
00-661 Warszawa, Poland

Received on 15.12.1987;
revised version on 6.3.1992