POLSKA AKADEMIA NAUK, INSTYTUT MATEMATYCZNY

DISSERTATIONES MATHEMATICAE

(ROZPRAWY MATEMATYCZNE)

CCCXXV

KRYSTYNA TWARDOWSKA

Approximation theorems of Wong–Zakai type for stochastic differential equations in infinite dimensions

WARSZAWA 1993

Krystyna Twardowska
Institute of Mathematics
Warsaw Technical University
Pl. Politechniki 1
00-661 Warszawa, Poland

Published by the Institute of Mathematics, Polish Academy of Sciences
Typeset in TEX at the Institute
Printed and bound by
drukarnia
herman & herman
02-240 Warszawa, ul. Jakobinów 23, tel: 846-79-66, tel/fax: 49-89-95

PRINTED IN POLAND

ISBN 83-85116-76-1 ISSN 0012-3862

CONTENTS

1. Introduction . . . 5
   1.1. The Wong–Zakai theorem and its generalizations . . . 5
   1.2. Approximation methods for stochastic differential equations . . . 7
   1.3. Extensions of the Wong–Zakai theorem and their applications . . . 9
2. Approximation theorem of Wong–Zakai type for functional stochastic differential equations . . . 10
   2.1. Introductory remarks . . . 10
   2.2. Definitions and notation . . . 10
   2.3. Description of the model . . . 11
   2.4. Approximation theorem . . . 15
   2.5. Examples . . . 24
3. An extension of the Wong–Zakai theorem to stochastic evolution equations in Hilbert spaces . . . 26
   3.1. Introductory remarks . . . 26
   3.2. Definitions and notation . . . 26
   3.3. Description of the model . . . 27
   3.4. The main theorem . . . 31
   3.5. Examples . . . 41
      3.5.1. Equations satisfying the assumptions of Theorem 3.4.1 . . . 41
      3.5.2. Stochastic delay equations . . . 43
      3.5.3. Stochastic wave equations . . . 45
4. Comparison of the results . . . 46
   4.1. Finite-dimensional case . . . 46
   4.2. Stochastic delay equations . . . 47
5. On the relation between the Itô and Stratonovich integrals in Hilbert spaces . . . 47
6. Conclusions . . . 49
   References . . . 50

1991 *Mathematics Subject Classification*: 34G20, 34K50, 35R15, 60H05, 60H10, 60H15, 60H30.
Received March 27, 1992; revised version July 13, 1992.

Abstract

Some generalizations of the approximation theorem of Wong–Zakai type for stochastic differential equations are examined. One of them deals with functional stochastic differential equations defined on some spaces of continuous functions. The second one concerns the situations when the state space and the Wiener process have values in some Hilbert spaces. The comparison of these results as well as some examples are also included. The correction terms computed here are then applied to the derivation of the relation between the Itô and Stratonovich integrals. Other important applications of the above theorems are indicated.

1. Introduction

1.1. The Wong–Zakai theorem and its generalizations. The theorem on the convergence of ordinary integrals to stochastic integrals was first proved by Wong and Zakai ([86], [87]) for a one-dimensional state space and one-dimensional Wiener process. The solution $x(t)$, $a \le t \le b$, to the stochastic differential equation

$$dx(t) = m(x(t), t)dt + \sigma(x(t), t)dw(t), \qquad x(a) = x_a\,, \tag{1.1.1}$$

is considered, where x_a is a random variable independent of $w(t) - w(a)$ and the functions m, σ satisfy the usual conditions guaranteeing the existence and uniqueness of the solution $x(t)$ ([86], [2], [40]). Let $x_n(t)$ be the solution of the ordinary differential equation

$$dx_n(t) = m(x_n(t), t)dt + \sigma(x_n(t), t)dw_n(t), \qquad x_n(a) = x_a\,, \tag{1.1.2}$$

for some regular approximations $w_n(t)$ of the Wiener process $w(t)$. Under suitable assumptions it is shown that $x_n(t)$ converges, as $n \to \infty$, to a process that does not satisfy the same equation (1.1.1), but it satisfies the equation

$$\begin{aligned} dy(t) &= m(y(t), t)dt + \frac{1}{2}\sigma(y(t), t)\frac{\partial\sigma(y(t), t)}{\partial y}\,dt + \sigma(y(t), t)dw(t),\\ y(a) &= x_a. \end{aligned} \tag{1.1.3}$$

The second term on the right hand side is the so-called "correction term". The reason for the difference between the two processes $x(t)$ and $y(t)$ is motivated by the approximate relationship $dw(t) \approx \sqrt{dt}$ (compare [2], [85], [86]).

More precisely, we have

THEOREM 1.1.1 [86]. *Suppose that* $(\Omega, \mathfrak{F}, P)$ *is a probability space and*

(i) $m(x,t)$, $\sigma(x,t)$, $\partial\sigma(x,t)/\partial x$, $\partial\sigma(x,t)/\partial t$ *are continuous in* $-\infty < x < \infty$, $a \le t \le b$,

(ii) $m(x,t)$, $\sigma(x,t)$, $\sigma(x,t)\cdot(\partial\sigma(x,t)/\partial x)$ *satisfy the Lipschitz condition with a constant* $k > 0$,

(iii) $\sigma(x,t) \ge \beta > 0$ (*or* $-\sigma(x,t) \ge \beta > 0$) *and* $|\partial\sigma(x,t)/\partial t| \le k\sigma^2(x,t)$,

(iv) *for each* n, $w_n(t,\omega)$ *is of bounded variation, continuous and has a piecewise continuous derivative*,

(v) *for almost all* $\omega \in \Omega$ *there exist* $n_0(\omega)$, $k(\omega)$, *both finite, such that for all* $n > n_0$ *and all* t *in* $[a,b]$, $w_n(t,\omega) \le k(\omega)$,

(vi) $w_n(t,\omega)$ *converges to* $w(t,\omega)$ *in* $[a,b]$ *almost surely as* $n \to \infty$,

(vii) $x_n(t)$ *and* $y(t)$ *satisfy equations* (1.1.2) *and* (1.1.3), *respectively.*

Then $x_n(t)$ *converges to* $x(t)$ *in* $[a,b]$ *almost surely as* $n \to \infty$.

There have been many generalizations of the above theorem to the case of several variables (see [26], [27], [44], [67]). If we consider the functions $m : [a,b] \times \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : [a,b] \times \mathbb{R}^d \to \mathbb{R}^{d\times m}$ then the correction term has the form

$$\frac{1}{2}\sum_{p=1}^{m}\sum_{j=1}^{d}\frac{\partial\sigma^{ip}(y(t),t)}{\partial y^j}\sigma^{jp}(y(t),t) \qquad \text{for } i = 1,\ldots,d\,.$$

Let us recall one of the main results in this area:

THEOREM 1.1.2 [27]. *Let* $X \in \mathbb{R}^d$, $m \in C^1_{\mathrm{b}}(\mathbb{R}^d,\mathbb{R}^d)$, $\sigma \in C^2_{\mathrm{b}}(\mathbb{R}^d,\mathbb{R}^{d\times m})$ (*i.e., of class* C^1 *and* C^2 *with bounded derivatives, respectively*). *Suppose* $B_n(t,w)$ *is a regular approximation of the* m*-dimensional Wiener process* $w(t)$ *on a Wiener space* (W^r_0, P) *and the following equations are satisfied*:

$$x^i_n(t,w) = X^i(w) + \int_0^t m^i(x_n(s,w))\,ds + \sum_{p=1}^{m}\int_0^t \sigma^{ip}(x_n(s,w))\dot{B}^p_n(s,w)\,ds\,,$$

$$\begin{aligned} y^i(t,w) = X^i(w) &+ \int_0^t m^i(y(s,w))\,ds + \sum_{p=1}^{m}\int_0^t \sigma^{ip}(y(s,w))\,dw(s) \\ &+ \frac{1}{2}\sum_{p=1}^{m}\sum_{j=1}^{d}\int_0^t \frac{\partial\sigma^{ip}(y(t),t)}{\partial y^j}\sigma^{jp}(y(t),t)\,dt \end{aligned}$$

for $i = 1,\ldots,d$. *Then*, *for every* $T > 0$,

$$\lim_{n\to\infty} E[\sup_{0\le t\le T}|x_n(t,w) - y(t,w)|^2] = 0\,.$$

Further generalizations deal with problems with more general noises than the Wiener process. The result due to Protter [61] for continuous semimartingale differentials can be stated in a simplified form as

THEOREM 1.1.3. *Consider the equations*

$$\begin{aligned} dx_n(t) &= f(t,x_n(t),Z_n(t))\,dZ_n(t), \\ (*) \qquad dx(t) &= f(t,x(t),Z(t)) \circ dZ(t), \\ (**) \qquad dy(t) &= f(t,y(t),Z(t))dZ(t) + \tfrac{1}{2}\{f(\partial f/\partial y) \\ &\quad + (\partial f/\partial Z)\}(t,y(t),Z(t))d[Z^c,Z^c](t), \end{aligned}$$

where Z_n *are piecewise* C^1 *approximations of a continuous semimartingale* Z *and* $\circ$ *denotes the Stratonovich integral. Under suitable assumptions*, *if* Z_n *tends to*

Z, then x_n tends to x satisfying $(*)$ (*and hence satisfying* $(**)$ *by the well-known relation between the Itô and Stratonovich integrals*).

The Wong–Zakai theorem has extensions in two directions: more general driven processes are considered or coefficients are allowed to depend on the trajectories of the solutions. In the first case, semimartingales with jumps have been considered by Marcus [43] and Kushner [38]. Results of this type were also examined by Bally [4], Ferreyra [20], Gyöngy [22], Mackevičius [41] and Picard [57]. In the second direction, pioneering work was done in [87] by Wong and Zakai, and more recently by Doss [19], Koneczny [34] and also by Nakao and Yamato [52].

In the infinite-dimensional case some generalizations are known where the Wiener process is one-dimensional and the state space is infinite-dimensional ([1], [7], [15], [19], [23]).

In [1] the following result is stated:

THEOREM 1.1.4. *Let* $(\Omega, \mathfrak{F}, P)$ *be a probability space. Consider the stochastic problem*

$$du(t) = (A(t)u(t) + \tfrac{1}{2}B^2u(t))dt + Bu(t)dw(t) + f(t)dt, \qquad u(0) = u_0 \tag{1.1.4}$$

in a real separable Hilbert space H, where $w(t)$ is a one-dimensional Brownian motion. For each $t \in [0,T]$ we assume that $A(t)$ generates an analytic semigroup and B genarates a strongly continuous group. Let $f, f_n : [0,T] \times \Omega \to H$, $u_0 : \Omega \to H$ be given data. Then, under standard assumptions, $u(t)$ is the unique generalized solution of (1.1.4) *and it is the limit of the solutions of the approximating deterministic problems*

$$du_n(t) = A(t)u_n(t)dt + Bu_n(t)\dot{w}_n(t)dt + f_n(t)dt, \qquad u_n(0) = u_0, \tag{1.1.5}$$

obtained by approaching the white noise $dw(t)$ with a sequence of regular coloured noises $\dot{w}_n(t)$.

We observe that the correction term is there of the form $\frac{1}{2}B^2u(t)$. Some slight modifications of the above theorem are given in [7] and [15].

In [23] the noise process is multi(finite-)dimensional and the operators acting on the infinite-dimensional state space are unbounded but again linear. The correction term introduced there behaves like the Lie bracket of some linear operators.

The assumptions imposed on the operators A and B are such that the considered equations admit many meaningful physical applications.

These generalizations do not concern stochastic differential delay equations. The Wong–Zakai type approximation theorems for such equations were proved in [78] and [79].

1.2. Approximation methods for stochastic differential equations. As already mentioned, the correction term appears in the Wong–Zakai type approximation theorems. However, some types of approximation theorems for stochastic differential equations are known that do not give any correction term.

For example, the objective of paper [42] was to extend the Cauchy–Maruyama approximation method to delay stochastic differential equations based on semimartingales with spatial parameters. This procedure is also applicable to non-delay equations.

In [29], [48], [49], [55], [59] and [71], both pathwise and mean-square convergence of some approximation schemes to stochastic differential equations are examined. These schemes are based on the Euler, Milshteĭn, Monte Carlo and Runge–Kutta methods. Some standard Monte Carlo techniques with the unbiased estimation of the transition density of the solution process instead of the approximation of the individual trajectories are also used in [84].

In [53] an efficient approximation scheme is proposed for stochastic differential equations based on irregular samples taken at the passage times of the driven process through a series of thresholds. This approximation is asymptotically efficient with respect to the irregular samples.

A very important contribution to approximation methods for stochastic differential equations was made by Kushner in [37], [38]. In [39] Kushner and Yin consider a class of recursive stochastic algorithms in which parallel processing methods are used for the Monte Carlo optimization of systems. Weak convergence methods are applied to sequences of iterates that converge to the solution of either ordinary or stochastic differential equations.

Approximation using the Taylor expansion is considered by Greiner and Strittmatter [21] and Platen [60]. The successive approximation can be found in the papers of Kawabata [32] and Tudor [75].

Summing up, the described approximation methods have followed five directions (see [64], [70]): mean-square approximation ([12], [48], [58]), pathwise approximation [71], approximation of expectations of the solutions ([49], [70]), numerical computation of the Lyapunov exponents [69] and asymptotically efficient schemes for minimization of the normalized quadratic mean error ([11], [53]).

In [45] the UT condition is defined for a sequence $\{z^n\}_{n\in\mathbb{N}}$ of $\mathfrak{F}^n$-adapted semimartingales and an approximation of the noise in the stochastic differential equation is introduced. A stability result is proved. This theorem is of a different kind than the Wong–Zakai theorem. Although a noise approximation is considered, no correction term appears in the limit equation. This is due to the UT property. The piecewise linear approximations of the Wiener process used e.g. in [78], [79] do not satisfy the UT condition so the correction term does appear. On the other hand, the discrete time approximation of the Wiener process satisfies the UT condition and the result of [45] can be applied.

A thorough description of numerical problems in this area can be found in [3], [33], [64].

1.3. Extensions of the Wong–Zakai theorem and their applications. The approximation theorem of Wong–Zakai type contained in Chapter 2 and [79] includes the case of stochastic differential equations with delayed argument. We use the general theory of functional differential equations ([25], [50], [51]). Mainly, we base our arguments on the approximation theorems of [27] and [79].

The correction term computed in Chapter 2 is a value of a measure connected with a directional derivative of the drift term in the stochastic delay differential equation. The same correction term appears in [18] and [79]. In [18] it was applied to computing the relation between the Stratonovich and Itô integrals of functions with delayed argument.

As examples we consider some types of equations with delay constant in time and with noise being the one-dimensional Wiener process.

In Chapter 3 an extension of the Wong–Zakai theorem to stochastic evolution equations in Hilbert spaces with the Wiener process with values in another Hilbert space is examined. The approximands form a sequence of deterministic differential equations. In the limit equation we obtain an Itô correction term of the form $\widetilde{\mathrm{tr}}(QD\mathcal{B}\mathcal{B})$, understood in the sense described in §3.2 (see also [19] and [78]). Q denotes here the covariance of the Wiener process, $\mathcal{B}$ is an operator acting on the Wiener process and $D\mathcal{B}$ is its Fréchet derivative. The present theorem is a modification under some weaker assumptions of the result of [78].

We would like to mention at this moment that stochastic evolution equations in Hilbert spaces are a general model for many different types of ordinary and partial stochastic differential equations.

In Chapter 4 we compare the results of Chapters 2 and 3. We observe in §4.1 that the general form of the correction term in Hilbert spaces is an extension of the known correction formula in the finite-dimensional case. The operator $\widetilde{\mathrm{tr}}$ becomes the trace of a matrix. This extension also includes the case of stochastic delay equations after their transformation to stochastic evolution equations in a Hilbert space with an appropriate strongly continuous semigroup. Finally (see §4.2), the correction term for the stochastic delay equations computed in Chapter 2 is compared with the correction term computed in Chapter 3 for the more general case of stochastic delay equations in Hilbert spaces. Although two different models for stochastic delay equations are assumed, this interesting comparison turns out to be possible.

The relations between different types of stochastic integrals have been examined for years (see for example [2], [66]). In [18] the Stratonovich integrals of functions with delayed argument are considered. A relation between this integral and the Itô stochastic integral is shown. The correction term computed there is the same as the term occurring in the approximation theorem of Wong–Zakai type in [79]. In Chapter 5 we examine the relation between the Itô and Stratonovich integrals in Hilbert spaces. That is, the Wiener process and the nonlinear operators under the integral sign have values in some Hilbert spaces. We follow the same idea as in [80].

Finally, Chapter 6 indicates some directions of further research as well as possible applications to some new kinds of problems.

2. Approximation theorem of Wong–Zakai type for functional stochastic differential equations

2.1. Introductory remarks. In this chapter we study the generalization of the Wong–Zakai theorem to nonlinear stochastic functional differential equations with values in the space $\mathbb{R}^d$ ($d \geq 1$) (see [79]). By piecewise linear approximation of the m-dimensional Wiener process we obtain an explicit formula for the limit of a sequence of solutions to certain ordinary differential equations with delayed argument; this very limit is the solution to the stochastic differential equation with delayed argument with an additional term called the Itô correction term. Moreover, we give some examples of the application of the theorem.

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

$$(f,g)_{\mathfrak{C}_-} = \sum_{n=1}^{\infty} 2^{-n} \frac{\|f-g\|_n}{1 + \|f-g\|_n}$$

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

Let $\mathfrak{B}(\mathfrak{X})$ denote the topological σ-algebra of the space $\mathfrak{X}$. It is obvious that it is identical with the σ-algebra generated by the family of all Borel cylinder sets in $\mathfrak{X}$. So we construct the Wiener space $(\mathfrak{C}_2^0, \mathfrak{B}(\mathfrak{C}_2^0), P^{\mathrm{w}})$, where P^{w} is the Wiener measure ([27], Chapter I). The coordinate process $B(t,w) = w(t)$, $w \in \mathfrak{C}_2^0$, is the m-dimensional Wiener process.

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

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

$$B^{n,p}(t,w) = w^p\left(\frac{k}{2^n}\right) + 2^n\left(t - \frac{k}{2^n}\right)\left(w^p\left(\frac{k+1}{2^n}\right) - w^p\left(\frac{k}{2^n}\right)\right) \tag{2.2.1}$$

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

We introduce the following notation and functions

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

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

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

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

is called the *segment* of the trajectory of f on $(-\infty, t]$.

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

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

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

2.3. Description of the model. Now we consider $\widetilde{\Omega} = \mathfrak{C}_2^0$. Let X be a continuous stochastic process $X(t, w) : (-\infty, T] \times \widetilde{\Omega} \to \mathbb{R}^d$, that is, $X : \widetilde{\Omega} \to \mathfrak{X} = \mathfrak{C}_1$.

We take some fixed initial constant stochastic processes $X^i(0 + \theta, w) = X_0^i(w) = X_0^{n,i}(w) = Y_0^i(w)$ for $\theta \in J$, $i = 1, \ldots, d$. We also consider operators $b : \mathfrak{C}_- \to \mathbb{R}^d$, $\sigma : \mathfrak{C}_- \to L(\mathbb{R}^m, \mathbb{R}^d)$ ($L(\mathbb{R}^m, \mathbb{R}^d)$ is the Banach space of linear functions from $\mathbb{R}^m$ to $\mathbb{R}^d$ with the uniform operator norm $|\cdot|_L$).

We introduce the condition

($\widetilde{\mathrm{A}}$1) for every $t \in (-\infty, T]$ the algebra $\mathfrak{B}_{-\infty,t}(X) \cup \mathfrak{B}_{-\infty,t}(dB)$ is independent of $\mathfrak{B}_{t,T}(dB)$

to give a meaning to the stochastic integrals in (2.3.1) below.

We assume

($\widetilde{\mathrm{A}}$2) b and σ are continuous operators.

Now we introduce the operators $\widetilde{b} : \mathfrak{C}_- \to \mathfrak{C}_-$ and $\widetilde{\sigma} : \mathfrak{C}_- \to C(J, L(\mathbb{R}^m, \mathbb{R}^d))$, where

$$\begin{aligned}
\widetilde{b} &: \mathfrak{C}_- \ni g \to (J \ni \tau \to b(g(\cdot + \tau)) \in \mathbb{R}^d),\\
\widetilde{\sigma} &: \mathfrak{C}_- \ni g \to (J \ni \tau \to \sigma(g(\cdot + \tau)) \in L(\mathbb{R}^m, \mathbb{R}^d)),
\end{aligned}$$

that is, using the shift transformation $S_\tau : J \ni \vartheta \to \vartheta + \tau$ for $\tau < 0$,

$$\begin{aligned}
[\widetilde{b}(g)](\tau) &= b(g \circ S_\tau) = b(g(\cdot + \tau)),\\
[\widetilde{\sigma}(g)](\tau) &= \sigma(g \circ S_\tau) = \sigma(g(\cdot + \tau)).
\end{aligned}$$

Remark 2.3.1. This construction explains why we take $(-\infty, 0]$ to be the domain of the initial function. If we took the interval $[-r, 0]$, $r > 0$, to be the domain it would be impossible to define $\widetilde{b}$ and $\widetilde{\sigma}$ correctly (that is, for $X_t = Y_t$ it could happen that $\widetilde{\sigma}(X_t) \neq \widetilde{\sigma}(Y_t)$).

We consider the following stochastic differential equation with delayed argument:

$$X^i(t,w) = X^i_0(w) + \int_0^t b^i(X_s(\cdot,w))\,ds + \sum_{p=1}^m \int_0^t \sigma^{ip}(X_s(\cdot,w))\,dw^p(s) \tag{2.3.1}$$

for $i = 1,\ldots,d$. By replacing the Wiener process by B^n we obtain the following approximations of (2.3.1):

$$\begin{aligned}X^{n,i}(t,w) = X^{n,i}_0(w) &+ \int_0^t b^i(X^n_s(\cdot,w))\,ds \\ &+ \sum_{p=1}^m \int_0^t \sigma^{ip}(X^n_s(\cdot,w))\dot{B}^{n,p}(s,w)\,ds\,.\end{aligned} \tag{2.3.2^n}$$

We also introduce another stochastic differential equation:

$$\begin{aligned}Y^i(t,w) = Y^i_0(w) &+ \int_0^t b^i(Y_s(\cdot,w))\,ds + \sum_{p=1}^m \int_0^t \sigma^{ip}(Y_s(\cdot,w))\,dw^p(s) \\ &+ \frac{1}{2}\sum_{p=1}^m \sum_{j=1}^d \int_0^t \widetilde{D}_j\sigma^{ip}(Y_s(\cdot,w))\sigma^{jp}(Y_s(\cdot,w))\,ds\end{aligned} \tag{2.3.3}$$

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

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

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

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

is a directional derivative. We shall also use the property of the Dirac measure that for a smooth function $h(\cdot)$ we have $\int_{-\infty}^0 h(v)\,\delta_0(dv) = h(0)$. We introduce a function

$$\widetilde{A}^{jpn}_t : J \ni \tau \to \sigma^{jp}(X^n_{t+\tau}(\cdot,w))\dot{B}^{n,p}(t+\tau,w) \in \mathbb{R}\,.$$

We put $\Psi(t,w) = b(X_t(w))$ and $\Phi(t,w) = \sigma(X_t(w))$. The second integral in (2.3.1) is the Itô integral ([27], [40]).

Let us introduce the following conditions:

($\widetilde{\mathrm{A}}3$) The initial process X_0 is $\mathfrak{F}_0$-measurable and $P(|X_0(w)| < \infty) = 1$, where $|X_0(w)| = \sum_{i=1}^d |X_0^i(w)|$, $\mathfrak{B}_{-\infty,0}(X_0)$ is independent of $\mathfrak{B}_{0,T}(B)$.

($\widetilde{\mathrm{A}}4$) For every $\varphi, \psi \in \mathfrak{C}_-$ the following Lipschitz condition is satisfied:

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

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

($\widetilde{\mathrm{A}}5$) For every $\varphi \in \mathfrak{C}_-$ the following growth condition is satisfied:

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

where $\varphi^2(0) = \sum_{i=1}^d \varphi_i^2(0)$.

($\widetilde{\mathrm{A}}6$) We have

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

($\widetilde{\mathrm{A}}7$) $b^i, \sigma^{ip} \in C_{\mathrm{b}}^1(\mathfrak{C}_-)$, for every $i = 1, \ldots, d$, $p = 1, \ldots, m$, where C_{b}^1 denotes the space of bounded mappings with continuous first derivative, that is, for every $A > 0$ and $\varepsilon > 0$ there exist $B > 0$ and $\delta > 0$ such that $\|X_s^1 - X_s^2\|_{[-B,0]} < \delta$ implies

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

($\| \cdot \|_{[-B,0]}$ denotes the usual supremum norm on $[-B, 0]$).

Definition 2.3.1. We say that a d-dimensional continuous stochastic process $X : (-\infty, T] \times \mathfrak{C}_2^0 \to \mathbb{R}^d$ is a *strong solution* to equation (2.3.1) for a given process $w(t)$ if conditions ($\widetilde{\mathrm{A}}1$), ($\widetilde{\mathrm{A}}2$), ($\widetilde{\mathrm{A}}6$) are satisfied and equation (2.3.1) is valid with probability 1 for all $t \in (-\infty, T]$.

The uniqueness of strong solutions is understood in the sense of trajectories: for any two strong solutions X and $\widetilde{X}$ to equation (2.3.1) defined on the same probability space we have

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

Definition 2.3.2. An absolutely continuous stochastic process $X^n : (-\infty, T] \times \mathfrak{C}_2^0 \to \mathbb{R}^d$ is a *solution* to equation ($2.3.2^n$) if conditions ($\widetilde{\mathrm{A}}2$), ($\widetilde{\mathrm{A}}3$) are satisfied and equation ($2.3.2^n$) is valid with probability 1 for all $t \in (-\infty, T]$.

Notice that conditions ($\widetilde{\text{A}}2$)–($\widetilde{\text{A}}7$) ensure the existence and uniqueness of the strong solution Y to equation (2.3.3). Indeed (see [28], Sections 5 and 7), under conditions ($\widetilde{\text{A}}2$)–($\widetilde{\text{A}}5$) there exists a strong solution to equation (2.3.1). The uniqueness may be derived from the proof of Theorem 11, Section 10 of [28], 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

$$b^i(Y_t(\cdot,w)) + \widetilde{D}_j\sigma^{ip}(Y_t(\cdot,w))\sigma^{jp}(Y_t(\cdot,w)) \tag{2.3.4}$$

in equation (2.3.3). Since $\widetilde{D}_j\sigma^{ip}$ is a measure, we have

$$|\widetilde{D}_j\sigma^{ip}(\varphi)\sigma^{jp}(\varphi)| \le \overline{\mathfrak{C}}|\sigma^{jp}(\varphi)|,$$
$$|\widetilde{D}_j\sigma^{ip}(\varphi)\sigma^{jp}(\varphi) - \widetilde{D}_j\sigma^{ip}(\psi)\sigma^{jp}(\psi)| \le \overline{\mathfrak{C}}|\sigma^{jp}(\varphi) - \sigma^{jp}(\psi)|\,,$$

where $\overline{\mathfrak{C}}$ is a constant. Thus, conditions ($\widetilde{\text{A}}4$) and ($\widetilde{\text{A}}5$) are satisfied for the term (2.3.4). It is obvious that the other conditions are also satisfied and equation (2.3.3) also has exactly one strong solution.

Moreover, for every $n \in \mathbb{N}$, under condition ($\widetilde{\text{A}}4$) there exists exactly one solution to the ordinary differential equation ($2.3.2^n$) with delayed argument (see [25] and [28]).

The following limit Z_t^n is understood in the sense of the locally convex topology in $\mathfrak{C}_-$:

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

that is,

$$\max_{-r\le\theta\le 0} \left| \frac{1}{h}(X_{t+h}^n(\theta) - X_t^n(\theta)) - Z_t^n(\theta) \right| \to 0 \qquad \text{as } h \to 0\,,$$

for every $r > 0$.

We have

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

Putting $u = t+\theta$ we have

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

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 view t as a variable we have

$$\dot{X}_{\boldsymbol{\cdot}}^n : [0,T] \ni t \to \dot{X}_t^n \in \mathfrak{C}_-\,.$$

2.4. Approximation theorem. We have proved in [79] the following

THEOREM 2.4.1. *Let conditions* $(\widetilde{\mathrm{A}}2)$–$(\widetilde{\mathrm{A}}5)$ *and* $(\widetilde{\mathrm{A}}7)$ *be satisfied. Let* $B^n(t,w)$ *be the approximation of type* (2.2.1) *of the Wiener process. We assume that* X^n *and* Y *are solutions to* $(2.3.2^n)$ *and* (2.3.3), *respectively*, *with a constant initial stochastic process. Then*, *conditions* $(\widetilde{\mathrm{A}}1)$ *and* $(\widetilde{\mathrm{A}}6)$ *are satisfied and*, *for every* $T>0$,

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

P r o o f. The assumptions of the theorem ensure the existence and uniqueness of the solutions to equations $(2.3.2^n)$ and (2.3.3). For every $i=1,\ldots,d$ we subtract equations $(2.3.2^n)$ and (2.3.3):

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

where

$$\begin{aligned}
H_1(t)&=\sum_{p=1}^{m}\int_{t_n^-}^{t}\sigma^{ip}(X_s^n(\cdot,w))\dot B^{n,p}(s,w)\,ds-\sum_{p=1}^{m}\int_{t_n^-}^{t}\sigma^{ip}(Y_s(\cdot,w))\,dw^p(s)\\
&\quad-\frac12\sum_{p=1}^{m}\sum_{j=1}^{d}\int_{t_n^-}^{t}\widetilde D_j\sigma^{ip}(Y_s(\cdot,w))\sigma^{jp}(Y_s(\cdot,w))\,ds\\
&=H_{11}(t)-\sum_{p=1}^{m}H_{12}^p(t)-H_{13}(t),\\
H_2(t)&=\sum_{p=1}^{m}\int_{1/2^n}^{t_n^-}\sigma^{ip}(X_s^n(\cdot,w))\dot B^{n,p}(s,w)\,ds-\sum_{p=1}^{m}\int_{1/2^n}^{t_n^-}\sigma^{ip}(Y_s(\cdot,w))\,dw^p(s)\\
&\quad-\frac12\sum_{p=1}^{m}\sum_{j=1}^{d}\int_{1/2^n}^{t_n^-}\widetilde D_j\sigma^{ip}(Y_s(\cdot,w))\sigma^{jp}(Y_s(\cdot,w))\,ds,\\
H_3&=\sum_{p=1}^{m}\int_{0}^{1/2^n}\sigma^{ip}(X_s^n(\cdot,w))\dot B^{n,p}(s,w)\,ds-\sum_{p=1}^{m}\int_{0}^{1/2^n}\sigma^{ip}(Y_s(\cdot,w))\,dw^p(s)\\
&\quad-\frac12\sum_{p=1}^{m}\sum_{j=1}^{d}\int_{0}^{1/2^n}\widetilde D_j\sigma^{ip}(Y_s(\cdot,w))\sigma^{jp}(Y_s(\cdot,w))\,ds,\\
H_4(t)&=\int_0^t b^i(X_s^n(\cdot,w))\,ds-\int_0^t b^i(Y_s(\cdot,w))\,ds.
\end{aligned}$$

Below, c_l, $l=0,1,\ldots,21$, denote some positive constants.

From $(2.3.2^n)$ we have

$$|X^{n,i}(t,w) - X^{n,i}(s,w)| \le c_0\Big(\sum_{p=1}^{m}\int_s^t |\dot{B}^{n,p}(u,w)|\,du + (t-s)\Big). \tag{2.4.1}$$

From the boundedness of σ we obtain

$$\begin{aligned}
&E[\sup_{0\le t<T}|H_{11}(t)|^2]\\
&\quad\le E\Big[\sup_{0\le t\le T}\Big|\sum_{p=1}^{m}\Big(\frac{1}{2^n}\sup_s \sigma^{ip}(X^n_s(\cdot,w))\Big)^2\Big(\int_{t_n^-}^{t_n^+}(\dot{B}^{n,p}(s,w))^2\,ds\Big)\Big|\Big]\\
&\quad\le c_1\frac{1}{2^n}E\Big[\sup_{0\le t\le T}\Big(\int_{t_n^-}^{t_n^+}|\dot{B}^{n,p}(s,w)|\,ds\Big)^2\Big]\\
&\quad\le c_1\frac{1}{2^n}E\Big[\sup_k\Big(\int_{k/2^n}^{(k+1)/2^n}|\dot{B}^{n,p}(s,w)|\,ds\Big)^4\Big]^{1/2}\\
&\quad= c_1\frac{1}{2^n}\Big(\sum_{k=1}^{m(T)}E\Big[B^{n,p}\Big(\frac{k+1}{2^n},w\Big)-B^{n,p}\Big(\frac{k}{2^n},w\Big)\Big]^4\Big)^{1/2}\\
&\quad\le c_1\frac{1}{2^n}\bigg(m(T)3\Big(\frac{1}{2^n}\Big)^2\bigg)^{1/2}\le c_2\Big(\frac{1}{2^n}\Big)^{3/2}.
\end{aligned}$$

Therefore

$$E[\sup_{0\le t\le T}|H_{11}(t)|^2]\to 0 \qquad \text{as } n\to\infty\,.$$

Further, we estimate

$$\begin{aligned}
E[\sup_{0\le t\le T}|H^p_{12}(t)|^2] &= E\Big[\sup_{0\le t\le T}\Big|\int_{t_n^-}^{t}\sigma^{ip}(Y_{s_n^-}(\cdot,w))\,dw^p(s)\\
&\qquad+\int_{t_n^-}^{t}(\sigma^{ip}(Y_s(\cdot,w))-\sigma^{ip}(Y_{s_n^-}(\cdot,w)))\,dw^p(s)\Big|^2\Big]\\
&\le E\Big[\sup_{0\le t\le T}\Big|\int_{t_n^-}^{t}\sigma^{ip}(Y_{s_n^-}(\cdot,w))\,dw^p(s)\Big|^2\Big]\\
&\qquad+E\Big[\sup_{0\le k\le m(T)}\sup_{0\le t\le 1/2^n}\Big|\int_{k/2^n}^{k/2^n+t}(\sigma^{ip}(Y_s(\cdot,w))\\
&\qquad-\sigma^{ip}(Y_{s_n^-}(\cdot,w)))\,dw^p(s)\Big|^2\Big]=\widehat{H}_1(t)+\widehat{H}_2(t)\,.
\end{aligned}$$

From the Hölder inequality, from the well-known inequality for the Itô integrals, the assumptions on σ^{ip} and (7.32) in [27], we have

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

$$\le c_3\Big(\int_{t_n^-}^{t} ds\Big)^2 \le c_4\left(\frac{1}{2^n}\right)^2 \to 0 \qquad \text{as } n \to \infty\,.$$

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 (2.3.3) after replacing w by w' and Y_0 by $Y_{k/2^n}$ defined by the formula $Y_{k/2^n}(\theta) = Y(k/2^n + \theta)$. Let $\mathfrak{B}_k$ be the smallest σ-algebra such that $Y_{k/2^n}$ is a stochastic process with respect to it. Let Y_t^{ξ} be the solution to (2.3.3) with the initial condition $Y_0^{\xi} = \xi$. Let E' denote the expectation and conditional expectation with respect to w'. Since the increments of the Wiener process are stationary, we may replace computing the expectation of the original process by computing E' for the translated process. Therefore, using ($\widetilde{\text{A}}$4) we have

$$\begin{aligned}
\widehat{H}_2(t) \le{}& \sum_{k=0}^{m(T)} E'\Big\{E\Big[\sup_{0\le t\le k/2^n}\Big(\int_0^t (\sigma^{ip}(Y_{k/2^n+s}(\cdot, w)) \\
&- \sigma^{ip}(Y_{k/2^n}(\cdot, w)))\, dw^p(s)\Big)^2\Big] \mid \mathfrak{B}_k\Big\} \\
\le{}& c_5 \sum_{k=0}^{m(T)} E'\Big\{E\Big[\int_0^{1/2^n} (\sigma^{ip}(Y_s(\cdot, w)) - \sigma^{ip}(\xi))^2 ds\Big] \mid \mathfrak{B}_k\Big\} \\
\le{}& c_6(m(T)+1)\sum_{i=1}^{d} E'\Big\{E\Big[\int_0^{1/2^n}\Big(L^1\int_{-\infty}^{0} |Y_s^{i\xi}(u) - Y_{k/2^n}^{i}(u)|^2\, dK(u) \\
&+ L^2\Big|Y^{i\xi}(s) - Y^i\Big(\frac{k}{2^n}\Big)\Big|^2\Big)\, ds\Big] \mid \mathfrak{B}_k\Big\},
\end{aligned}$$

where $\xi(\theta) = Y(k/2^n + \theta)$ and, as in (7.57) of [27], we get

$$\widehat{H}_2(t) \le c_7\frac{1}{2^n} \to 0 \qquad \text{as } n \to \infty\,.$$

From ($\widetilde{\text{A}}$7) we have

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

$$\le c_8\Big|\int_{t_n^-}^{t} ds\Big|^2 \le c_9\left(\frac{1}{2^n}\right)^2 \to 0 \qquad \text{as } n \to \infty\,.$$

To estimate $H_2(t)$ we first represent the integral $\int_{t_n^+}^{t_n^-}$ as $\sum_{k=1}^{m(t)-1}\int_{k/2^n}^{(k+1)/2^n}$.

Observe that for $-\infty < u \le 0$,

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

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

$$\frac{dX^{n,i}(t+\cdot,w)}{dt}(\tau) = b^i(X_{t+\tau}^n(\cdot,w)) + \sum_{p=1}^m \sigma^{ip}(X_{t+\tau}^n(\cdot,w))\dot{B}^{n,p}(t+\tau,w)\,. \tag{2.4.2}$$

Now consider

$$\begin{aligned}
H &= \int_{k/2^n}^{(k+1)/2^n} \sigma^{ip}(X_s^n(\cdot,w))\dot{B}^{n,p}(s,w)\,ds \\
&= \sigma^{ip}(X_{k/2^n}^n(\cdot,w))\bigg(B^{n,p}\Big(\frac{k+1}{2^n},w\Big) - B^{n,p}\Big(\frac{k}{2^n},w\Big)\bigg) \\
&\quad + (\sigma^{ip}(X_{(k+1)/2^n}^n(\cdot,w)) - \sigma^{ip}(X_{k/2^n}^n(\cdot,w)))B^{n,p}\Big(\frac{k+1}{2^n},w\Big) \\
&\quad - \int_{k/2^n}^{(k+1)/2^n} D\sigma^{ip}(X_s^n(\cdot,w))\frac{dX^n(s+\cdot,w)}{ds}B^{n,p}(s,w)\,ds \\
&= J_1(k) + \int_{k/2^n}^{(k+1)/2^n} D\sigma^{ip}(X_s^n(\cdot,w)) \\
&\quad \times \Big(\widetilde{b}(X_s^n(\cdot,w)) + \sum_{p=1}^m \widetilde{A}_s^{pn}\Big)\bigg(B^{n,p}\Big(\frac{k+1}{2^n},w\Big) - B^{n,p}(s,w)\bigg)\,ds\,.
\end{aligned}$$

It follows that

$$
\begin{aligned}
H &= J_1(k) + \sum_{j=1}^{d} \int\limits_{k/2^n}^{(k+1)/2^n} \int\limits_{-\infty}^{0} b^j(X^n_{s+v}(\cdot,w))\, \mu^{ipj}_{s,w,X}(dv) \\
&\quad \times \left(B^{n,p}\left(\frac{k+1}{2^n}, w\right) - B^{n,p}(s,w) \right) ds \\
&\quad + \sum_{j=1}^{d} \int\limits_{k/2^n}^{(k+1)/2^n} \int\limits_{-\infty}^{0} \sum_{p=1}^{m} \sigma^{jp}(X^n_{s+v}(\cdot,w)) \dot{B}^{n,p}(s+v,w)\, \widetilde{\mu}^{ipj}_{s,w,X}(dv) \\
&\quad \times \left(B^{n,p}\left(\frac{k+1}{2^n}, w\right) - B^{n,p}(s,w) \right) ds \\
&\quad + \sum_{j=1}^{d} \int\limits_{k/2^n}^{(k+1)/2^n} \sum_{p=1}^{m} \widetilde{D}_j \sigma^{ip}(X^n_s(\cdot,w)) \sigma^{jp}(X^n_s(\cdot,w)) \dot{B}^{n,p}(s,w) \\
&\quad \times \left(B^{n,p}\left(\frac{k+1}{2^n}, w\right) - B^{n,p}(s,w) \right) ds \\
&= J_1(k) + \sum_{r=1}^{3} \sum_{j=1}^{d} J^j_{r2}(k)\,.
\end{aligned}
$$

As in [27], we write

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

and

$$
\begin{aligned}
H_2(t) &= \sum_{p=1}^{m} \Bigg(\sum_{k=1}^{m(t)-1} J_{11}(k) - \int\limits_{1/2^n}^{t_n^-} \sigma^{ip}(Y_s(\cdot,w))\, dw^p(s) \\
&\quad + \sum_{k=1}^{m(t)-1} J_{12}(k) + \sum_{k=1}^{m(t)-1} J_{13}(k) + \sum_{k=1}^{m(t)-1} J_{14}(k)
\end{aligned}
$$

$$+\sum_{j=1}^{d}\Bigg(\sum_{k=1}^{m(t)-1}\sum_{r=1}^{3}J_{r2}^{j}(k)-\frac{1}{2}\int\limits_{1/2^n}^{t_n^-}\widetilde{D}_j\sigma^{ip}(Y_s(\cdot,w))\sigma^{jp}(Y_s(\cdot,w))\,ds\Bigg)\Bigg)$$
$$=\sum_{p=1}^{m}(\widetilde{I}_1(t)+\ldots+\widetilde{I}_5(t))$$

(for simplicity we ignore the dependence on p in notation). Now we estimate each of $\widetilde{I}_1,\ldots,\widetilde{I}_5$, successively. For every $t_1\in[0,T]$, using the martingale inequality, (2.4.1), ($\widetilde{\mathrm{A}}$4), and (7.2) in [27], we obtain

$$E[\sup_{0\le t\le t_1}|\widetilde{I}_1(t)|^2]=E\Big[\sup_{0\le t\le t_1}\Big|\int\limits_{1/2^n}^{t_n^-}(\sigma^{ip}(X_{s_n^- -\delta}(\cdot,w))-\sigma^{ip}(Y_s(\cdot,w)))\,dw^p(s)\Big|^2\Big]$$
$$\le c_{10}\Big(\int\limits_0^{t_1}E[|Y(s,w)-X^n(s,w)|^2]\,ds$$
$$+\int\limits_0^{t_1}\Big(\int\limits_{-\infty}^{0}E[|Y_{s_1}(\cdot,w)-X^n_{s_1}(\cdot,w)|^2]\,dK(s_1)\Big)\,ds+\frac{n}{2^n}\Big).$$

Now from (2.4.1) and the Hölder inequality we get

$$E[\sup_{0\le t\le t_1}|\widetilde{I}_2(t)|^2]$$
$$\le E\Bigg[\sum_{k=1}^{m(T)-1}(\sigma^{ip}(X^n_{k/2^n}(\cdot,w))$$
$$-\sigma^{ip}(X^n_{k/2^n-\delta}(\cdot,w)))^2\sum_{k=1}^{m(T)-1}\Big(B^{n,p}\Big(\frac{k+1}{2^n},w\Big)-B^{n,p}\Big(\frac{k}{2^n},w\Big)\Big)^2\Bigg]$$
$$\le c_{11}\Bigg(m(T)\sum_{k=1}^{m(T)-1}E\Big[L^1\Big(\int\limits_{-\infty}^{0}|X^n_{k/2^n}(\theta)-X^n_{k/2^n-\delta}(\theta)|\,dK(\theta)$$
$$+L^2\Big|X\Big(\frac{k}{2^n},w\Big)-X^n\Big(\frac{k}{2^n}-\delta,w\Big)\Big|\Big)^4\Big]$$
$$\times m(T)\sum_{k=1}^{m(T)-1}E\Big[\Big|B^{n,p}\Big(\frac{k+1}{2^n},w\Big)-B^{n,p}\Big(\frac{k}{2^n},w\Big)\Big|^4\Big]\Bigg)^{1/2}$$
$$\le c_{12}\Bigg(m(T)\sum_{k=1}^{m(T)-1}E\Big[\Big|X^n\Big(\frac{k}{2^n},w\Big)-X^n\Big(\frac{k}{2^n}-\delta,w\Big)\Big|^4\Big]$$
$$\times m(T)\sum_{k=1}^{m(T)-1}E\Big[B^{n,p}\Big(\frac{k+1}{2^n},w\Big)-B^{n,p}\Big(\frac{k}{2^n},w\Big)\Big]^4\Bigg)^{1/2}$$

$$\le c_{13}\left((m(T))^2\left(\frac{1}{n}\frac{1}{2^n}\right)^2(m(T))^2\left(\frac{1}{2^n}\right)^2\right)^{1/2}$$
$$\le c_{14}\frac{1}{n}\to 0 \qquad \text{as } n\to\infty\,.$$

Let
$$\eta_l(w)=\sum_{k=1}^{l}\sigma^{ip}(X^n_{k/2^n-\delta}(\cdot,w))B^{n,p}(0,\theta_{(k+1)/2^n}w)\,,$$
where $\theta_t w(s)=w(t+s)-w(t)$. It is obvious that η_l is an F_l-martingale for $F_l=\mathfrak{B}_{(l+2)/2^n}$.

Since $\widetilde{I}_3(t)$ as well as $\widetilde{I}_4(t)$ written as $\eta_l(w)$ are F_l-martingales, from the martingale inequality we have
$$E[\sup_{0\le t\le T}|\widetilde{I}_3(t)|^2]\le c_{15}\frac{1}{n}\to 0 \qquad \text{as } n\to\infty$$
and
$$E[\sup_{0\le t\le T}|\widetilde{I}_4(t)|^2]\le c_{16}\frac{1}{n}\to 0 \qquad \text{as } n\to\infty\,.$$
We write $\widetilde{I}_5(t)=\sum_{j=1}^d J_5^j(t)$, where
$$\begin{aligned}J_5^j(t)=&\sum_{k=1}^{m(t)-1}\sum_{r=1}^{3}J^j_{r2}(k)\\&-\frac{1}{2}\sum_{k=1}^{m(t)-1}\int_{k/2^n}^{(k+1)/2^n}\widetilde{D}_j\sigma^{ip}(Y_s(\cdot,w))\sigma^{jp}(Y_s(\cdot,w))\,ds\,.\end{aligned}$$

It follows that
$$\begin{aligned}J_5^j(t)=&\sum_{k=1}^{m(t)-1}\int_{k/2^n}^{(k+1)/2^n}(\widetilde{D}_j\sigma^{ip}(X^n_s(\cdot,w))\sigma^{jp}(X^n_s(\cdot,w))\\&-\widetilde{D}_j\sigma^{ip}(X^n_{k/2^n}(\cdot,w))\sigma^{jp}(X^n_{k/2^n}(\cdot,w)))\dot{B}^{n,p}(s,w)\left(\dot{B}^{n,p}\left(\frac{k+1}{2^n},w\right)\right.\\&\left.-B^{n,p}(s,w)\right)ds+\sum_{k=1}^{m(t)-1}\int_{k/2^n}^{(k+1)/2^n}\int_{-\infty}^{0}b^j(X^n_{s+v}(\cdot,w))\,\mu^{ipj}_{s,w}(dv)\\&\times\left(B^{n,p}\left(\frac{k+1}{2^n},w\right)-B^{n,p}(s,w)\right)ds\\&+\sum_{k=1}^{m(t)-1}\int_{k/2^n}^{(k+1)/2^n}\widetilde{D}_j\sigma^{ip}(X^n_{k/2^n}(\cdot,w))\sigma^{jp}(X^n_{k/2^n}(\cdot,w))\end{aligned}$$

$$
\begin{aligned}
&\times\bigg(\dot{B}^{n,p}(s,w)\bigg(B^{n,p}\bigg(\frac{k+1}{2^n},w\bigg)-B^{n,p}(s,w)\bigg)-c_p\bigg(\frac{1}{2^n},n\bigg)\bigg)\,ds\\
&+\frac{1}{2}\sum_{k=1}^{m(t)-1}\int\limits_{k/2^n}^{(k+1)/2^n}(\widetilde{D}_j\sigma^{ip}(X^n_{k/2^n}(\cdot,w))\sigma^{jp}(X^n_{k/2^n}(\cdot,w))\\
&-\widetilde{D}_j\sigma^{ip}(Y_s(\cdot,w))\sigma^{jp}(Y_s(\cdot,w)))\,ds\\
&+\sum_{k=1}^{m(t)-1}\frac{1}{2^n}\widetilde{D}_j\sigma^{ip}(X^n_{k/2^n}(\cdot,w))\sigma^{jp}(X^n_{k/2^n}(\cdot,w))\bigg(c_p\bigg(\frac{1}{2^n},n\bigg)-\frac{1}{2}\bigg)\\
&+\sum_{k=1}^{m(t)-1}J^j_{22}(k)\\
=\;&I_{51}(t)+\ldots+I_{55}(t)+I_{56}(t)\,,
\end{aligned}
$$

where

$$
c_j(t,n)=(1/t)E\Big[\int\limits_0^t\dot{B}^{n,j}_s(\cdot,w)(B^{n,j}(t,w)-B^{n,j}(s,w))\,ds\Big]
$$

and $\lim_{n\to\infty}c_j(1/2^n,n)=1/2$ (see [27], Lemma 7.1).

Since X^n is uniformly continuous on every finite interval, X^n_s is continuous as a function of the variable s with functional values and we may estimate (analogously to [27])

$$
E[\sup_{0\le t\le T}|I_{5j}(t)|^2]\to 0\qquad\text{as } n\to\infty,\ \text{for } j=1,2,3,5,6\,.
$$

Further, using ($\widetilde{\mathrm{A}}4$) and (7.68) of [27], for every $t_1\in[0,T]$, we obtain as in [79]

$$
\begin{aligned}
E[\sup_{0\le t\le t_1}|I_{54}(t)|^2]&\le c_{17}\Big(\int\limits_0^{t_1}E[|Y(s,w)-X^n(s^-_n,w)|^2]\,ds\\
&\quad+\int\limits_0^{t_1}\Big(\int\limits_{-\infty}^{0}E[|Y_{s_1}(\cdot,w)-X^n_{s^-_{1,n}}(\cdot,w)|^2]\,dK(s_1)\Big)\,ds\Big)\\
&\le c_{18}\Big(\int\limits_0^{t_1}E[|Y(s,w)-X^n(s,w)|^2]\,ds\\
&\quad+\int\limits_0^{t_1}\Big(\int\limits_{-\infty}^{0}E[|Y_{s_1}(\cdot,w)-X^n_{s_1}(\cdot,w)|^2]\,dK(s_1)\Big)\,ds+\frac{2n}{2^n}\Big)\,.
\end{aligned}
$$

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

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

For every $t_1 \in [0,T]$, using the Hölder inequality and $(\widetilde{\mathrm{A}}4)$ we obtain

$$E[\sup_{0\le t\le t_1} |H_4(t)|^2] \le c_{19}\Big(\int_0^{t_1} E[|Y(s,w)-X^n(s,w)|^2]\,ds + \int_0^{t_1}\Big(\int_{-\infty}^{0} E[|Y_{s_1}(\cdot,w)-X^n_{s_1}(\cdot,w)|^2]\,dK(s_1)\Big)\,ds\Big).$$

Now we shall use the general Gronwall lemma ([40], Lemma 4.13).

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

$$u(t) \le k_0 + k_1\int_0^t v(s)u(s)\,ds + k_2\int_0^t v(s)\int_{-\infty}^{0} u(s_1)\,dK(s_1)\,ds\,.$$

Then

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

We use this lemma for $u(t)=E[\sup_{0\le s\le t}|Y(t,w)-X^n(t,w)|^2]$ and $v(t)=1$. We obtain

$$\begin{aligned} E[\sup_{0\le t\le t_1} |X^n(t,w)-Y(t,w)|^2] &\le o(1)+c_{20}\Big(\int_0^{t_1} E[\sup_{0\le s\le t}|X^n(s,w)-Y(s,w)|^2]\,ds \\ &\quad + \int_0^{t_1}\Big(\int_{-\infty}^{0} E[\sup_{0\le s\le t}|X^n_{s_1}(\cdot,w)-Y_{s_1}(\cdot,w)|^2]\,dK(s_1)\Big)\,ds\Big)\,, \end{aligned}$$

with $o(1)$ uniformly convergent for $t_1\in[0,T]$. Consequently, we have

$$E[\sup_{0\le t\le t_1} |X^n(t,w)-Y(t,w)|^2] \le o(1)\exp(c_{21}T) \to 0 \qquad \text{as } n\to\infty\,.$$

This completes the proof.

Remark 2.4.1. Instead of the interval $J=(-\infty,0]$ we can consider $J=[-r,0]$, $r>0$. Then, instead of considering $X^i(t_i^n+s)-X^i(t_{i-1}^n+s)$ on the whole interval of definition (t_{i-1}^n, t_i^n are certain points of a partition of the time axis), we observe that

$$X^i(t_i^n+s)-X^i(t_{i-1}^n+s)$$
$$=\begin{cases} X_0^i(t_i^n+s)-X_0^i(t_{i-1}^n+s) \quad \text{for } t_i^n+s\le 0,\\ X_0^i(0)-X_0^i(t_{i-1}^n+s)+\int_0^{t_i^n+s} b^i(X_u(\cdot))\,du \\ \qquad +\sum_{p=1}^m\int_0^{t_i^n+s}\sigma^{ij}(X_u(\cdot))\,dw^p(u) \quad \text{for } t_{i-1}^n+s\le 0\le t_i^n+s,\\ \int_{t_{i-1}^n+s}^{t_i^n+s} b^i(X_u(\cdot))\,du+\sum_{p=1}^m\int_{t_{i-1}^n+s}^{t_i^n+s}\sigma^{ij}(u,X_u(\cdot))\,dw^p(u) \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{for } t_{i-1}^n+s>0,\end{cases}$$

and we estimate each part separately by expressions converging to zero.

Instead of the constant initial condition $X_0(\cdot)$ we can take a function satisfying the Hölder condition

$$\|X_0(t)-X_0(s)\|^2\le Z(\omega)|t-s|^\beta,\qquad \beta>0\,.$$

Then

$$\|X_0(t_i^n+s)-X_0(t_{i-1}^n+s)\|^2\le Z(\omega)|t_i^n-t_{i-1}^n|^\beta\,,$$

and we obtain convergence to zero. Under suitable assumptions on the random variable $Z(\omega)$ we can prove the convergence in the mean-square sense or with probability one.

2.5. Examples

Example 2.5.1. Consider the equation

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

where $\Sigma:\mathfrak{C}_-\to\mathbb{R}$, $\Sigma(\varphi)=\varphi(-1)$, that is, $\varphi(-1)=X_t(-1)=X(t-1)$. Then

$$dX(t)=X(t-1)\,dw(t),\qquad X_0=\eta\,,$$

and equation (2.3.3) is of the form

$$dY(t)=Y(t-1)\,dw(t),\qquad Y_0=\eta\,,$$

because the measure μ is concentrated on the set $\{-1\}$ only and hence $\mu(\{0\})=0$. Therefore, there is no difference between the initial and limit equations.

Using the step-by-step method of solving delay equations we take, for example, the equation

$$\begin{aligned} dX(t)&=X(t-1)dw(t) &&\text{for } t\ge 0\,,\\ X(t)&=1 &&\text{for } t\in[-1,0]\end{aligned}$$

and

$$\begin{aligned} dX^n(t)&=X^n(t-1)\dot w_n(t)\,dt &&\text{for } t\ge 0\,,\\ X^n(t)&=1 &&\text{for } t\in[-1,0]\,.\end{aligned}$$

We obtain in the first step for $t\in[0,1]$,

$$dX(t)=dw(t),\qquad X(0)=1$$

and

$$dX^n(t)=\dot w_n(t)\,dt,\qquad X^n(0)=1\,.$$

After integrating we get

$$X(t) = X(0) + \int_0^t dw(s) = 1 + w(t)$$

and

$$X^n(t) = X^n(0) + \int_0^t \dot{w}_n(s)\, ds = 1 + w_n(t)\,.$$

In the second step we consider for $t \in [1, 2]$,

$$dX(t) = (1 + w(t-1))dw(t), \qquad X(1) = 1 + w(1)$$

and

$$dX^n(t) = (1 + w_n(t-1))w_n(t)\, dt, \qquad X^n(1) = 1 + w_n(1)\,.$$

We obtain

$$X(t) = 1 + w(t) + \int_1^t w(s-1)\, dw(s)$$

and

$$X^n(t) = 1 + w_n(1) + \int_1^t w_n(s-1)\dot{w}_n(s)\, ds\,.$$

It is easy to observe that $X^n(t) \to X(t) = Y(t)$ as $n \to \infty$ (in the mean-square sense).

Example 2.5.2. Now we consider the equation

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

where for some constants b_0, b_1, σ_0 and σ_1, we define $B, \Sigma : \mathfrak{C}_- \to \mathbb{R}$,

$$B(\varphi) = b_0\varphi(0) + b_1\varphi(-r), \qquad \Sigma(\varphi) = \sigma_0\varphi(0) + \sigma_1\varphi(-r)\,.$$

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

$$\begin{aligned} dX(t) &= (b_0X(t) + b_1X(t-r))\, dt + (\sigma_0X(t) + \sigma_1X(t-r))dw(t)\,,\\ X_0 &= \eta\,. \end{aligned}$$

Then equation (2.3.3) is of the form

$$\begin{aligned} dY(t) &= (b_0Y(t) + b_1Y(t-r))dt + (\sigma_0Y(t) + \sigma_1Y(t-r))dw(t)\\ &\quad + \tfrac{1}{2}\sigma_0(\sigma_0Y(t) + \sigma_1Y(t-r))dt\,,\\ Y_0 &= \eta \end{aligned}$$

because $\sigma_0X(t)$ is the only term for which the support of the measure contains zero. Therefore, $\mu(\{0\}) = \sigma_0$.

3. An extension of the Wong–Zakai theorem to stochastic evolution equations in Hilbert spaces

3.1. Introductory remarks. In this chapter we examine an approximation theorem of Wong–Zakai type for stochastic evolution equations in a Hilbert space with the noise being the generalized derivative of the Wiener process with values in another Hilbert space. As a consequence of the approximation of the Wiener process we get in the limit equation the Itô correction term for the infinite-dimensional case. The result obtained includes the case of stochastic delay equations. The uniqueness and existence of solutions are guaranteed by known theorems for mild solutions. The model considered here is more complicated than that examined by the author in [78]; also, the approximation theorem is proved under weaker assumptions than in [78]. Namely, a nonlinear term with a nonlinear operator $\mathcal{C}$ is added, assumption ($\widetilde{\mathrm{A}}5$) is weakened. Moreover, the proof is performed from the very beginning for the infinite-dimensional Wiener process while in [78] it was done for the one-dimensional Wiener process and then the necessary changes were indicated.

3.2. Definitions and notation. Let H and H_1 be real separable Hilbert spaces with the norms $\|\cdot\|_H$, $\|\cdot\|_{H_1}$ and the scalar products $\langle\cdot,\cdot\rangle_H$, $\langle\cdot,\cdot\rangle_{H_1}$. Let $(\Omega,\mathfrak{F},(\mathfrak{F}_t)_{t\in[0,T]},P)$ be a filtered probability space on which an increasing and right-continuous family $(\mathfrak{F}_t)_{t\in[0,T]}$ of complete sub-σ-algebras of $\mathfrak{F}$ is defined. $L(H,H_1)$ denotes the space of bounded linear operators from H to H_1. Let $L_2(H,H_1)$ be the space of Hilbert–Schmidt operators with the norm $\|\cdot\|_{\mathrm{HS}}$.

We take an H-valued Wiener process $w(t)$, $t\in[0,T]$, with nuclear covariance operator $Q\in L(H)=L(H,H)$.

It is known [14] that there are real-valued independent Wiener processes $\{w_i(t)\}_{i=0}^{\infty}$ on $[0,T]$ such that

$$w(t)=\sum_{i=0}^{\infty} w_i(t)e_i$$

almost everywhere in $(t,\omega)\in[0,T]\times\Omega$, where $\{e_i\}_{i=0}^{\infty}$ is an orthonormal basis of eigenvectors of Q corresponding to eigenvalues $\{\lambda_i\}_{i=0}^{\infty}$, $\sum_{i=0}^{\infty}\lambda_i<\infty$, with $E[\Delta w_i\Delta w_j]=(t-s)\lambda_i\delta_{ij}$ for $\Delta w_i=w_i(t)-w_i(s)$ and $s<t$ (δ_{ij} is the Kronecker delta).

Let (see [10])

$$\begin{aligned}&\Lambda_T(w,H,H_1)\\ &\quad=\Big\{\overline{\Psi}:\overline{\Psi}:[0,T]\times\Omega\to L(H,H_1)\text{ is a progressively measurable process,}\\ &\qquad E\Big[\int_0^T\|\overline{\Psi}Q^{1/2}\|_{\mathrm{HS}}^2\,ds\Big]=\|\overline{\Psi}\|_{\Lambda_T}^2=\sum_{i=0}^{\infty}E\Big[\int_0^T\|\overline{\Psi}(s,\omega)e_i\|_{H_1}^2\,ds\Big]<\infty\Big\}.\end{aligned}$$

It is known that for $\overline{\Psi} \in \Lambda_T$ the stochastic integral $\int \overline{\Psi}\, dw$ is well defined and it can be represented by

$$\int_0^t \overline{\Psi}(s,\omega)\, dw(s) = \sum_{i=0}^{\infty} \int_0^t \overline{\Psi}(s,\omega) e_i \, dw_i(s)\,. \tag{3.2.1}$$

The convergence in (3.2.1) is in $L^2(\Omega)$ for each $t > 0$.

3.3. Description of the model. We consider the stochastic differential equation

$$\begin{aligned} dz(t) &= \mathcal{A}z(t)dt + \mathcal{C}(z(t))dt + \mathcal{B}(z(t))dw(t)\,,\\ z(0) &= z_0\,, \end{aligned} \tag{3.3.1}$$

where

(A1) $(z(t))_{t\in[0,T]}$ is an H_1-valued stochastic process, $(w(t))_{t\geq 0}$ is an H-valued Wiener process with the covariance operator Q, $\mathcal{A} : H_1 \supset D(\mathcal{A}) \to H_1$ is the infinitesimal generator of a strongly continuous semigroup $(S(t))_{t\geq 0}$, $\mathcal{C} : H_1 \to H_1$ and $\mathcal{B} : H_1 \to L(H, H_1)$ are bounded nonlinear operators. Moreover, we assume that $(S(t))_{t\geq 0}$ is a semigroup of contraction type, i.e., there exists a constant $\beta \in \mathbb{R}_+$ such that $\|S(t)\|_{H_1} \leq \exp(\beta t)$ for all $t \in [0, T]$,

(A2) $z_0 \in D(\mathcal{A})$ is an H_1-valued square integrable $\mathfrak{F}_0$-measurable initial random variable.

Apart from (3.3.1) we consider the equation

$$\begin{aligned} d\widehat{z}(t) &= \mathcal{A}\widehat{z}(t)dt + \mathcal{C}(\widehat{z}(t))dt + \mathcal{B}(\widehat{z}(t))dw(t)\\ &\qquad + \tfrac{1}{2}\widetilde{\operatorname{tr}}(QD\mathcal{B}(\widehat{z}(t))\mathcal{B}(\widehat{z}(t)))dt,\\ \widehat{z}(0) &= z_0\,, \end{aligned} \tag{3.3.2}$$

where $\widetilde{\operatorname{tr}}(QD\mathcal{B}(\widehat{z}(t))\mathcal{B}(\widehat{z}(t)))$ is defined below.

We observe that the Fréchet derivative $D\mathcal{B}(h_1) \in L(H_1, L(H, H_1))$ for $h_1 \in H_1$ and we consider the composition $D\mathcal{B}(h_1)\mathcal{B}(h_1) \in L(H, L(H, H_1))$. We view the Fréchet derivative of $\mathcal{B}(h_1)$ as $D\mathcal{B}(h_1, h_2)$ since $h_2 \to D\mathcal{B}(h_1, h_2)$, $h_2 \in H_1$, is linear and belongs to $L(H_1, L(H, H_1))$. Let $\Psi \in L(H, L(H, H_1))$ and define (see [19]) $\mathcal{B}_{\widetilde{h}_1}(h, h') := \langle \Psi(h)(h'), \widetilde{h}_1\rangle_{H_1} \in \mathbb{R}$ for $h, h' \in H$. From the Riesz theorem for the form Ψ on H we conclude that for every $\widetilde{h}_1 \in H_1$ there exists an operator $\widetilde{\Psi}(\widetilde{h}_1) \in L(H)$ such that for every $h, h' \in H$

$$\mathcal{B}_{\widetilde{h}_1}(h, h') = \langle \widetilde{\Psi}(\widetilde{h}_1)(h), h'\rangle_H = \langle \Psi(h)(h'), \widetilde{h}_1\rangle_{H_1}\,. \tag{3.3.3}$$

Now, the covariance operator Q has finite trace and therefore the mapping

$$\widetilde{\xi} : H_1 \ni \widetilde{h}_1 \to \operatorname{tr}(Q\widetilde{\Psi}(\widetilde{h}_1)) \in \mathbb{R}$$

is a linear bounded functional on H_1. Therefore, using the Riesz theorem we find a unique $\widetilde{\widetilde{h}}_1 \in H_1$ such that $\widetilde{\xi}(\widetilde{h}_1) = \langle \widetilde{\widetilde{h}}_1, h_1\rangle_{H_1}$.

Define

$$\widetilde{\widetilde{h}}_1 = \widetilde{\operatorname{tr}}(Q\Psi)\,.$$

We observe that $\langle \widetilde{\widetilde{h}}_1, \widetilde{h}_1\rangle_{H_1}$ is the trace of the operator $Q\widetilde{\Psi}(\widetilde{h}_1) \in L(H)$ but $\widetilde{\operatorname{tr}}(Q\Psi)$ is merely a symbol for $\widetilde{\widetilde{h}}_1$.

Since

$$\begin{aligned}\operatorname{tr}[Q\widetilde{\Psi}(\widetilde{h}_1)] &= \sum_{j=0}^{\infty}\langle\widetilde{\Psi}(\widetilde{h}_1)e_j, e_j\rangle_H = \sum_{j=0}^{\infty}\langle\widetilde{\Psi}(\widetilde{h}_1)e_j, Q^*e_j\rangle_H \\ &= \sum_{j=0}^{\infty}\langle\widetilde{\Psi}(\widetilde{h}_1)e_j, Qe_j\rangle_H = \sum_{j=0}^{\infty}\langle\Psi(e_j)(Qe_j), \widetilde{h}_1\rangle_{H_1} \\ &= \sum_{j=0}^{\infty}\langle\Psi(e_j)(\lambda_j e_j), \widetilde{h}_1\rangle_{H_1}\,,\end{aligned}$$

taking in particular $\Psi = D\mathcal{B}(h_1)\mathcal{B}(h_1)$ we get

$$\widetilde{\widetilde{h}}_1 = \widetilde{\operatorname{tr}}[Q\Psi] = \sum_{j=0}^{\infty}\Psi(e_j)(Qe_j) = \sum_{j=0}^{\infty}[D\mathcal{B}(h_1)\mathcal{B}(h_1)(e_j)](\lambda_j e_j)\,. \tag{3.3.4}$$

Remark 3.3.1. In particular, if $H = \mathbb{R}^1$, then $w(\cdot)$ is the one-dimensional Wiener process, $\widetilde{\operatorname{tr}}(Q\Psi) = \Psi$. Equation (3.3.2) including the correction term is now of the form

$$d\widehat{z}(t) = \mathcal{A}\widehat{z}(t)dt + \mathcal{B}(\widehat{z}(t))d\,\overline{w}(t) + \tfrac{1}{2}D\mathcal{B}(\widehat{z}(t))\mathcal{B}(\widehat{z}(t))dt\,. \tag{3.3.2$'$}$$

We actually use $\overline{w}(t) = \sqrt{\lambda}w(t)$ in the one-dimensional case to simplify the notation in the correction term. Then $\sqrt{\lambda}$ does not appear in the correction term as it does in the infinite-dimensional case when $\sqrt{\lambda}$ is present owing to the definition of Q. Let us observe from (3.3.4) that for the infinite-dimensional $w(\cdot)$ the correction term is the series of correction terms derived for the one-dimensional Wiener processes.

Moreover, we assume

(A3) there is a constant $K > 0$ and a positive definite symmetric nuclear operator R which commutes with S such that $P(R^{-1}z_0 \in H_1) = 1$ and

(i) $$\|R^{-1}\mathcal{C}(h_1)\|_{H_1}^2 + \|R^{-1}\mathcal{B}(h_1)Q^{1/2}\|_{\mathrm{HS}}^2 + \|R^{-1}\operatorname{tr}(QD\mathcal{B}(h_1)\mathcal{B}(h_1))\|_{H_1}^2 \le K(1 + \|h_1\|_{H_1}^2)\,,$$

(ii) $$\|\mathcal{C}(h_1) - \mathcal{C}(\widetilde{h}_2)\|_{H_1}^2 + \operatorname{tr}((\mathcal{B}(h_1) - \mathcal{B}(\widetilde{h}_1))Q(\mathcal{B}(h_1) - \mathcal{B}(\widetilde{h}_1))^*) \le K\|h_1 - \widetilde{h}_1\|_{H_1}^2$$

for $h_1, \widetilde{h}_1 \in H_1$, where $*$ denotes the adjoint operator,

(A4) the operator $\mathcal{C}$ is of class C^1 and the operator $\mathcal{B} \in C_{\mathrm{b}}^1$, i.e., is of class C^1 with bounded derivative, and this derivative is assumed to be globally Lipschitzean, (A5) the operator $D\mathcal{B}(h_1)\mathcal{A} : H_1 \supset D(\mathcal{A}) \to L(H, H_1)$ can be uniquely extended to a bounded operator from H_1 to $L(H, H_1)$, so there exists a positive constant k such that for $h_1 \in H_1$

$$\|D\mathcal{B}(h_1)\mathcal{A}h_1\|_{L(H,H_1)} \le k\|h_1\|_{H_1} . \tag{3.3.5}$$

Remark 3.3.2. The mapping $(h_1, h) \to D\mathcal{B}(h_1)\mathcal{A}h$ has domain $H_1 \times D(\mathcal{A})$ and is linear in the second variable. If assumption (A5) is satisfied, this mapping can be extended to $H_1 \times H_1$. Now $D\mathcal{B}(h_1)\mathcal{A}h$ is equal to $D\mathcal{B}(h_1, \mathcal{A}h)$.

Remark 3.3.3. In the present paper the assumptions are rather strong. Assumptions of this kind are also used in [63], [74], [76].

Assumption (A5) here is weaker than in [78]. That is, we assume there instead of inequality (3.3.5) that

$$\|D\mathcal{B}(h_1)\mathcal{A}h_1\|_{L(H,H_1)} \le k . \tag{3.3.6}$$

We can weaken it because of a localization property. Namely, we observe that for every $\varepsilon > 0$ there exist $n_0 \in \mathbb{N}$ and $r > 0$ such that we have

$$P(u^n(t) = u_r^n(t)) \ge 1 - \varepsilon \qquad \text{for } t \in [0, T],\ n \ge n_0 .$$

Here $u_r^n(t)$ denotes the solution to equation (3.3.11) below with

$$\mathcal{B}_r = \begin{cases} \widetilde{\mathcal{B}} & \text{in } K_r, \\ 0 & \text{for } h_1 \notin K_{2r}, \end{cases}$$

instead of $\mathcal{B}$, where $K_r = \{h_1 \in H_1 : \|h_1\|_{H_1} \le r\}$ and $\widetilde{\mathcal{B}}$ satisfies (3.3.6).

We now define the nth approximation to the Wiener process $(w(t))_{t\ge0}$ as follows:

$$w^n(t) = \sum_{j=0}^{\infty} w_j^n(t) e_j ,$$

where $0 = t_0^n < \ldots < t_n^n$ and for $t_{i-1}^n < t \le t_i^n$,

$$w_j^n(t) = \frac{t - t_{i-1}^n}{t_i^n - t_{i-1}^n} w_j(t_i^n) + \frac{t_i^n - t}{t_i^n - t_{i-1}^n} w_j(t_{i-1}^n) . \tag{3.3.7}$$

For $\overline{\Psi} \in \Lambda_T(w, H, H_1)$ we have

$$\int_0^t \overline{\Psi}(s,\omega)\, dw^n(s) = \sum_{j=0}^{\infty} \int_0^t \overline{\Psi}(s,\omega)(e_j)\, dw_j^n(s) , \tag{3.3.8}$$

where the integrals on the right-hand side are the classical Stieltjes integrals.

We rewrite (3.3.1) in the mild integral form

$$\begin{aligned} &z(t) = S(t)z_0 + \int_0^t S(t-s)\mathcal{C}(z(s))\, ds + \int_0^t S(t-s)\mathcal{B}(z(s))\, dw(s) , \\ &z(0) = z_0 . \end{aligned} \tag{3.3.9}$$

Similarly, from (3.3.2) we get

$$
\begin{aligned}
\widehat{z}(t) &= S(t)z_0 + \int_0^t S(t-s)\mathcal{C}(\widehat{z}(s))\,ds + \int_0^t S(t-s)\mathcal{B}(\widehat{z}(s))\,dw(s) \\
&\quad + \tfrac{1}{2}\int_0^t S(t-s)\widetilde{\mathrm{tr}}(QD\mathcal{B}(\widehat{z}(s))\mathcal{B}(\widehat{z}(s)))\,ds\,, \\
\widehat{z}(0) &= z_0\,.
\end{aligned}
\tag{3.3.10}
$$

Consider now the sequence of integral equations

$$
\begin{aligned}
u^n(t) &= S(t)z_0 + \int_0^t S(t-s)\mathcal{C}(u^n(s))\,ds \\
&\quad + \int_0^t S(t-s)\mathcal{B}(u^n(s))\,dw^n(s)\,, \\
u^n(0) &= z_0,
\end{aligned}
\tag{3.3.11}
$$

$n = 1, 2, \ldots$ First observe that under our assumptions the integrals are well defined.

Definition 3.3.1. Suppose we are given an H_1-valued initial random variable z_0 and an H-valued Wiener process $(w(t))_{t\ge 0}$. Moreover, assume that an H_1-valued stochastic process $(z(t))_{t\in[0,T]}$ has the following properties:

(i) $(z(t))_{t\in[0,T]}$ is progressively measurable,

(ii) $\mathcal{B}(z(\cdot)) \in \Lambda_T(w, H, H_1)$,

(iii) for every $t \in [0,T]$ there exists Ω_t with $P(\Omega_t) = 1$ such that for every $\omega \in \Omega_t$ equation (3.3.9) is satisfied.

Then $(z(t))_{t\in[0,T]}$ is called a *mild solution* to (3.3.1) with the initial condition z_0.

The uniqueness of solution is understood in the sense of trajectories.

The existence and uniqueness of solution to (3.3.1) under hypotheses (A1)–(A3) is obtained from Theorem 2.1 of [74].

Definition 3.3.2. Let $n \in \mathbb{N}$. We say that a mapping $u^n : [0,T] \to H_1$ is a *mild integral solution* to equation (3.3.11) if u^n is continuous and equation (3.3.11) is satisfied for all $0 \le t \le T$.

For each $n \in \mathbb{N}$ the existence and uniqueness of solution to equation (3.3.11) follows from [35], [56]. Indeed, for each $n \in \mathbb{N}$ and almost every $\omega \in \Omega$ we write (3.3.11) on $(t^n_{i-1}, t^n_i]$, $i = 1, \ldots, n$, in the form

$$
du^n(t) = \mathcal{A}u^n(t)dt + \mathcal{B}_{a_i}(u^n(t))dt, \tag{3.3.12}
$$

where $\mathcal{B}_{a_i}(\cdot) = \mathcal{B}(\cdot)a_i$ for $a_i \in H$. Here $\mathcal{B}_{a_i} = \mathcal{B}_{\Delta w_i}$.

We know ([56], Ch. VI, Th. 1.5) that a sufficient condition for the mild solution to (3.3.11) to be continuously differentiable on $[0,T]$ is $\mathcal{B} \in C^1$ and $z_0 \in D(\mathcal{A})$. It is easy to observe that equation (3.3.12) satisfies these assumptions.

Notice that (A1)–(A4) ensure the existence and uniqueness of a mild solution to equation (3.3.10). Indeed, under condition (A4) we see that the term $\widetilde{\mathrm{tr}}(QD\mathcal{B}(\widehat{z}(t))\mathcal{B}(\widehat{z}(t)))$ satisfies condition (A3) because the series in (3.3.4) converges. It is obvious that conditions (A1), (A2) are also satisfied and that equation (3.3.10) has exactly one mild solution.

3.4. The main theorem. We shall prove the following

THEOREM 3.4.1. *Let $(w^n(t))_{t\ge 0}$ be the n-th approximation of the Wiener process $(w(t))_{t\ge 0}$ as given in* (3.3.7). *Let $(u^n(t))_{t\in[0,T]}$ be the solution to equation* (3.3.11) *and $\widehat{z}(t)$ to equation* (3.3.10). *Assume that hypotheses* (A1)–(A5) *are satisfied and $E[\|R^{-1}z_0\|^2_{H_1}] < \infty$. Then, for each T, $0<T<\infty$ and given $\varepsilon>0$*

$$\lim_{n\to\infty} P(\sup_{0\le t\le T} \|u^n(t,\omega) - \widehat{z}(t,\omega)\|_{H_1} \ge \varepsilon) = 0\,. \tag{3.4.1}$$

Proof. In order to prove the theorem we need a finite-difference approximation scheme for equation (3.3.1). We take a sequence of partitions $\{t_0^n,\ldots,t_n^n\}$ of $[0,T]$ such that $0=t_0^n<\ldots<t_n^n=T$.

Let $h^n = \sup\{t_i^n - t_{i-1}^n : i=1,\ldots,n\}$. Define the process $\widetilde{\xi}^n$ by $\widetilde{\xi}^n(0) = z_0$ and

$$\begin{aligned}\widetilde{\xi}^n(t) = S(t-t_{i-1}^n)\widetilde{\xi}^n(t_{i-1}^n) &+ \int_{t_{i-1}^n}^t S(t-t_{i-1}^n)\mathcal{C}(\widetilde{\xi}^n(t_{i-1}^n))\,ds \\ &+ \int_{t_{i-1}^n}^t S(t-t_{i-1}^n)\mathcal{B}(\widetilde{\xi}^n(t_{i-1}^n))\,dw(s)\end{aligned} \tag{3.4.2}$$

for $t_{i-1}^n < t \le t_i^n$.

Now we introduce two operator-valued functions on $\{(t,s): 0<s<t<T\}$ as follows:

$$\mathcal{C}_n(\widetilde{\xi}^n,t,s) = \begin{cases} S(t-t_{i-1}^n)\mathcal{C}(\widetilde{\xi}^n(t_{i-1}^n)) & \text{on } (t_{i-1}^n, t_i^n], \\ S(t-t_{m-1}^n)\mathcal{C}(\widetilde{\xi}^n(t_{m-1}^n)) & \text{on } (t_{m-1}^n, t], \end{cases} \tag{3.4.3}$$

$$\mathcal{B}_n(\widetilde{\xi}^n,t,s) = \begin{cases} S(t-t_{i-1}^n)\mathcal{B}(\widetilde{\xi}^n(t_{i-1}^n)) & \text{on } (t_{i-1}^n, t_i^n], \\ S(t-t_{m-1}^n)\mathcal{B}(\widetilde{\xi}^n(t_{m-1}^n)) & \text{on } (t_{m-1}^n, t], \end{cases} \tag{3.4.4}$$

where $t_{m-1}^n < t \le t_m^n$, $0<m\le n$ and $i=1,\ldots,m-1$. The values of $\mathcal{C}_n$ are in H_1 and those of $\mathcal{B}_n$ are operators from H to H_1. Therefore, for any $t\in[0,T]$ we have

$$\widetilde{\xi}^n(t) = S(t)z_0 + \int_0^t \mathcal{C}_n(\widetilde{\xi}^n,t,s)\,ds + \int_0^t \mathcal{B}_n(\widetilde{\xi}^n,t,s)\,dw(s)\,. \tag{3.4.5}$$

We may state the following

LEMMA 3.4.1. *Assume that* $(z(t))_{t\in[0,T]}$ *is the mild solution to equation* (3.3.1), *the hypotheses* (A1)–(A3) *are satisfied and* $E[\|R^{-1}z_0\|^2_{H_1}]<\infty$. *Then for every* $\varepsilon>0$

$$\lim_{n\to\infty} P(\sup_{0\le t\le T}\|\widetilde{\xi}^n(t,\omega)-z(t,\omega)\|_{H_1}\ge\varepsilon)=0\,. \tag{3.4.6}$$

To prove this lemma we apply a similar argument to the proof of Lemma 2.6 of [74].

To prove the theorem we further transform equations (3.3.10) and (3.3.11). In analogy to equation (3.4.5) corresponding to (3.4.2) we derive from (3.3.10) the equation

$$\begin{aligned}\widehat{\xi}^n(t)&=S(t)z_0+\int_0^t \mathcal{C}_n(\widehat{\xi}^n,t,s)\,ds\\&\quad+\int_0^t \mathcal{B}_n(\widehat{\xi}^n,t,s)\,dw(s)+\frac12\int_0^t G_n(\widehat{\xi}^n,t,s)\,ds\,,\end{aligned} \tag{3.4.7}$$

where $\mathcal{C}_n$ and $\mathcal{B}_n$ are defined similarly to (3.4.3) and (3.4.4). Moreover, G_n is an operator-valued function on $\{(t,s):0<s<t<T\}$ with values in H_1, defined by

$$G(\xi,t,s)=\begin{cases}S(t-t^n_{i-1})\widetilde{\mathrm{tr}}(QD\mathcal{B}(\widehat{\xi}^n(t^n_{i-1}))\mathcal{B}(\widehat{\xi}^n(t^n_{i-1}))) & \text{on } (t^n_{i-1},t^n_i],\\ S(t-t^n_{m-1})\widetilde{\mathrm{tr}}(QD\mathcal{B}(\widehat{\xi}^n(t^n_{m-1}))\mathcal{B}(\widehat{\xi}^n(t^n_{m-1}))) & \text{on } (t^n_{m-1},t],\end{cases} \tag{3.4.8}$$

where $t^n_{m-1}<t\le t^n_m$, $0<m\le n$ and $i=1,\ldots,m-1$.

By Lemma 3.4.1, for every $\varepsilon>0$ we obtain

$$\lim_{n\to\infty} P(\sup_{0\le t\le T}\|\widehat{\xi}^n(t,\omega)-\widehat{z}(t,\omega)\|_{H_1}\ge\varepsilon)=0\,. \tag{3.4.9}$$

Now we put (3.3.11) in another form which may be easily compared with the solutions to (3.4.7). On $(t^n_{i-1},t^n_i]$ we have

$$\begin{aligned}u^n(t^n_i)&=S(h^n)u^n(t^n_{i-1})+\int_{t^n_{i-1}}^{t^n_i} S(t^n_i-t^n_{i-1})\mathcal{C}(u^n(t^n_{i-1}))\,ds\\&\quad+\int_{t^n_{i-1}}^{t^n_i} S(t^n_i-s)(\mathcal{C}(u^n(s))-\mathcal{C}(u^n(t^n_{i-1})))\,ds\\&\quad+\int_{t^n_{i-1}}^{t^n_i} (S(t^n_i-s)-S(t^n_i-t^n_{i-1}))\mathcal{C}(u^n(t^n_{i-1}))\,ds\end{aligned} \tag{3.4.10}$$

$$+ \int_{t_{i-1}^n}^{t_i^n} (S(t_i^n - s) - S(t_i^n - t_{i-1}^n))\mathcal{B}(u^n(t_{i-1}^n))\, dw^n(s)$$
$$+ \int_{t_{i-1}^n}^{t_i^n} S(t_i^n - s)(\mathcal{B}(u^n(s)) - \mathcal{B}(u^n(t_{i-1}^n)))\, dw^n(s)$$
$$+ \int_{t_{i-1}^n}^{t_i^n} S(t_i^n - t_{i-1}^n)\mathcal{B}(u^n(t_{i-1}^n))\, dw^n(s) = \sum_{l=1}^{7} I_l\,.$$

We transform I_6 as follows:

$$I_6 = \int_{t_{i-1}^n}^{t_i^n} S(t_i^n - s)\Big(\int_{t_{i-1}^n}^{s} \frac{d}{d\tau}\mathcal{B}(u^n(\tau))\, d\tau\Big) dw^n(s) \tag{3.4.11}$$
$$= \int_{t_{i-1}^n}^{t_i^n} S(t_i^n - s)\Big(\int_{t_{i-1}^n}^{s} D\mathcal{B}(u^n(\tau))\frac{du^n(\tau)}{d\tau}\, d\tau\Big) dw^n(s)\,.$$

It follows that

$$I_6 = \int_{t_{i-1}^n}^{t_i^n} S(t_i^n - s)\Big(\int_{t_{i-1}^n}^{s} D\mathcal{B}(u^n(\tau))\mathcal{A}u^n(\tau)\, d\tau\Big)\, dw^n(s)$$
$$+ \int_{t_{i-1}^n}^{t_i^n} S(t_i^n - s)\Big(\int_{t_{i-1}^n}^{s} D\mathcal{B}(u^n(\tau))\mathcal{C}(u^n(\tau))\, d\tau\Big)\, dw^n(s)$$
$$+ \int_{t_{i-1}^n}^{t_i^n} S(t_i^n - s)\Big(\int_{t_{i-1}^n}^{s} (D\mathcal{B}(u^n(\tau))\mathcal{B}(u^n(\tau))$$
$$- D\mathcal{B}(u^n(t_{i-1}^n))\mathcal{B}(u^n(t_{i-1}^n)))\, dw^n(\tau)\Big)\, dw^n(s)$$
$$+ \int_{t_{i-1}^n}^{t_i^n} (S(t_i^n - s) - S(t_i^n - t_{i-1}^n))\Big(\int_{t_{i-1}^n}^{s} D\mathcal{B}(u^n(t_{i-1}^n))$$
$$\circ\, \mathcal{B}(u^n(t_{i-1}^n))\, dw^n(\tau)\Big)\, dw^n(s)$$
$$+ \int_{t_{i-1}^n}^{t_i^n} S(t_i^n - t_{i-1}^n)\Big(\int_{t_{i-1}^n}^{s} D\mathcal{B}(u^n(t_{i-1}^n))\mathcal{B}(u^n(t_{i-1}^n))\, dw^n(\tau)\Big) dw^n(s)$$
$$= \widetilde{H}_1 + \widetilde{H}_2 + (\widetilde{H}_3 - \widetilde{H}_4) + (\widetilde{H}_4 - \widetilde{H}_5) + \widetilde{H}_4\,.$$

We recall that

$$\Delta_i^n w_j = w_j^n(t_i^n) - w_j^n(t_{i-1}^n) = w_j(t_i^n) - w_j(t_{i-1}^n)\,. \tag{3.4.12}$$

Now we use (3.3.8), (3.4.12) and assumptions on $\mathcal{B}$ to deduce that

$$\begin{aligned}(3.4.13)\quad \widetilde{H}_4 &= \sum_{j=0}^{\infty}\sum_{k=0}^{\infty}(S(t_i^n - t_{i-1}^n)D\mathcal{B}(u^n(t_{i-1}^n))\mathcal{B}(u^n(t_{i-1}^n))(e_j))(e_k)\\ &\quad\times \int_{t_{i-1}^n}^{t_i^n}\Big(\int_{t_{i-1}^n}^{s} dw_j^n(\tau)\Big)\,dw_k^n(s)\\ &= \frac{1}{2}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}(S(t_i^n - t_{i-1}^n)D\mathcal{B}(u^n(t_{i-1}^n))\mathcal{B}(u^n(t_{i-1}^n))(e_j))(e_k)\\ &\quad\times \dot{w}_j^n\dot{w}_k^n(t_i^n - t_{i-1}^n)^2\\ &= \frac{1}{2}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}(S(t_i^n - t_{i-1}^n)D\mathcal{B}(u^n(t_{i-1}^n))\mathcal{B}(u^n(t_{i-1}^n))(e_j))(e_k)\\ &\quad\times \Delta_i^n w_j \Delta_i^n w_k\,,\end{aligned}$$

because $\dot{w}_j^n(t_i^n - t_{i-1}^n)\dot{w}_k^n(t_i^n - t_{i-1}^n) = \Delta_i^n w_j \Delta_i^n w_k$, where $\dot{w}_j^n$, $\dot{w}_k^n$ are constants; they are derivatives of w_j^n and w_k^n, respectively, on $(t_{i-1}^n, t_i^n]$. Further,

$$\begin{aligned}(3.4.14)\quad E[\widetilde{H}_4] &= \frac{1}{2}\sum_{j=0}^{\infty}(S(t_i^n - t_{i-1}^n)D\mathcal{B}(u^n(t_{i-1}^n))\mathcal{B}(u^n(t_{i-1}^n))(e_j))(\lambda_j e_j)\\ &\quad\times (t_i^n - t_{i-1}^n)\\ &= \frac{1}{2}\sum_{j=0}^{\infty}\int_{t_{i-1}^n}^{t_i^n} S(t_i^n - t_{i-1}^n)D\mathcal{B}(u^n(t_{i-1}^n))\mathcal{B}(u^n(t_{i-1}^n))(e_j)(\lambda_j e_j)\,ds\\ &= \frac{1}{2}\int_{t_{i-1}^n}^{t_i^n} S(t_i^n - t_{i-1}^n)\widetilde{\operatorname{tr}}(QD\mathcal{B}(u^n(t_{i-1}^n))\mathcal{B}(u^n(t_{i-1}^n)))\,ds\,,\end{aligned}$$

because

$$(3.4.15)\qquad E[\Delta_i^n w_j \Delta_i^n w_k] = (t_i^n - t_{i-1}^n)\lambda_j\delta_{jk}\,.$$

Notice for further reference that by (3.4.12),

$$\begin{aligned}(3.4.16)\quad &\int_{t_{i-1}^n}^{t_i^n} S(t_i^n - t_{i-1}^n)\mathcal{B}(u^n(t_{i-1}^n))\,dw^n(s)\\ &\qquad = \int_{t_{i-1}^n}^{t_i^n} S(t_i^n - t_{i-1}^n)\mathcal{B}(u^n(t_{i-1}^n))\,dw(s)\,.\end{aligned}$$

Applying the (by now standard) discretization to equation (3.3.11), for $0 < m \le n$, with the help of (3.4.11), (3.4.13), (3.4.16), we obtain the following equation:

$$
\begin{aligned}
(3.4.17)\qquad u^n(t_m^n) &= S(t_m^n)z_0 + \int_0^{t_m^n} \mathcal{C}_n(u^n, t_m^n, s)\,ds \\
&+ \sum_{i=1}^{m} \int_{t_{i-1}^n}^{t_i^n} S(t_m^n - s)(\mathcal{C}(u^n(s)) - \mathcal{C}(u^n(t_{i-1}^n)))\,ds \\
&+ \sum_{i=1}^{m} \int_{t_{i-1}^n}^{t_i^n} (S(t_m^n - s) - S(t_m^n - t_{i-1}^n))\mathcal{C}(u^n(t_{i-1}^n))\,ds \\
&+ \sum_{i=1}^{m} \int_{t_{i-1}^n}^{t_i^n} (S(t_m^n - s) - S(t_m^n - t_{i-1}^n))\mathcal{B}(u^n(t_{i-1}^n))\,dw^n(s) \\
&+ \sum_{i=1}^{m} \int_{t_{i-1}^n}^{t_i^n} S(t_m^n - s)\Big(\int_{t_{i-1}^n}^{s} D\mathcal{B}(u^n(\tau))\mathcal{A}u^n(\tau)\,d\tau\Big)dw^n(s) \\
&+ \sum_{i=1}^{m} \int_{t_{i-1}^n}^{t_i^n} S(t_m^n - s)\Big(\int_{t_{i-1}^n}^{s} D\mathcal{B}(u^n(\tau))\mathcal{C}(u^n(\tau))\,d\tau\Big)dw^n(s) \\
&+ \sum_{i=1}^{m} \int_{t_{i-1}^n}^{t_i^n} S(t_m^n - s)\Big(\int_{t_{i-1}^n}^{s} (D\mathcal{B}(u^n(\tau))\mathcal{B}(u^n(\tau)) \\
&- D\mathcal{B}(u^n(t_{i-1}^n))\mathcal{B}(u^n(t_{i-1}^n)))dw^n(\tau)\Big)\,dw^n(s) \\
&+ \sum_{i=1}^{m} \int_{t_{i-1}^n}^{t_i^n} (S(t_m^n - s) - S(t_m^n - t_{i-1}^n))\Big(\int_{t_{i-1}^n}^{s} D\mathcal{B}(u^n(t_{i-1}^n)) \\
&\circ \mathcal{B}(u^n(t_{i-1}^n))\,dw^n(\tau)\Big)\,dw^n(s) \\
&+ \frac{1}{2}\sum_{i=1}^{m}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty} S(t_m^n - t_{i-1}^n)D\mathcal{B}(u^n(t_{i-1}^n))\mathcal{B}(u^n(t_{i-1}^n))(e_j)(e_k) \\
&\times \varDelta_i^n w_j \varDelta_i^n w_k + \int_0^{t_m^n} \mathcal{B}_n(u^n, t_m^n, s)\,dw(s) = \sum_{l=1}^{11} \widehat{I}_l\,.
\end{aligned}
$$

We shall prove that $\widehat{I}_3$–$\widehat{I}_9$ in (3.4.17) converge to zero. We shall also show that $\widehat{I}_{10}$ gives the correction term occurring in (3.3.2). Moreover, $\widehat{I}_1$, $\widehat{I}_2$ and $\widehat{I}_{11}$ yield our initial equation (3.3.1). Below, $L_1, \ldots, L_{18}$, $\widetilde{L}$, $\widetilde{\widetilde{L}}$ denote positive constants.

First we estimate

$$
\begin{aligned}
E[\sup_{t_{i-1}^n < s \le t_i^n} \|u^n(s) - u^n(t_{i-1}^n)\|_{H_1}^2] &= E[\sup_{t_{i-1}^n < s \le t_i^n} \|(S(s - t_{i-1}^n) - I)u^n(t_{i-1}^n)\|_{H_1}^2] \\
&+ \widetilde{L}(t_i^n - t_{i-1}^n)^2 + \widetilde{\widetilde{L}}E[\sup_{t_{i-1}^n < s \le t_i^n} \|w^n(s) - w^n(t_{i-1}^n)\|_{H_1}^2]\,.
\end{aligned}
$$

Notice that by Lemma 4.1 of [76] the family of processes $\{u^n\}$ is relatively compact in $C([0,T],H_1)$. We denote by $\mathsf{C}(\widetilde{K})$ the complement of the set $\widetilde{K}$. If we choose a compact set $\widetilde{K}\subset C([0,T],H_1)$ such that $P(u^n\in\mathsf{C}(\widetilde{K}))\le\varepsilon$ and we define the compact set $\widehat{K}=\{f(t): t\le T, f\in\widetilde{K}\}$, then we have ([74], p. 193)

$$\begin{aligned}\widehat{I}&=E[\sup_{t^n_{i-1}<s\le t^n_i}\|(S(s-t^n_{i-1})-I)u^n(t^n_{i-1})\|^2_{H_1}]\\&\le P(u^n\in\mathsf{C}(\widetilde{K}))+\sup_{y\in\hat{K}}\sup_{t\le h_n}\|(S(t)-I)y\|^2_{H_1}\end{aligned}\tag{3.4.18}$$

so $\widehat{I}\to 0$ as $n\to\infty$. Since each w^n is linear on any interval $(t^n_{i-1},t^n_i]$, for $t\in(t^n_{i-1},t^n_i]$ we have

$$\begin{aligned}w^n(t)&=w^n(t^n_{i-1})+\alpha^n_i(t)(w^n(t^n_i)-w^n(t^n_{i-1}))\\&=w(t^n_{i-1})+\alpha^n_i(t)(w(t^n_i)-w(t^n_{i-1}))\,,\end{aligned}$$

where $\alpha^n_i(t)=(t-t^n_{i-1})/(t^n_i-t^n_{i-1})$ is monotonic and $0\le\alpha^n_i(t)\le 1$. Therefore, using (3.4.18) we have

$$\lim_{n\to\infty}E[\sup_{t^n_{i-1}<s\le t^n_i}\|u^n(s)-u^n(t^n_{i-1})\|^2_{H_1}]=0\,.\tag{3.4.19}$$

Now, we estimate $\widehat{I}_3$ using (A1), (3.4.18), a Chebyshev type inequality and the Schwarz inequality:

$$\begin{aligned}\|\widehat{I}_3\|_{H_1}&\le L_1\sum_{i=1}^m\int\limits_{t^n_{i-1}}^{t^n_i}\|S(t^n_m-s)(u^n(s)-u^n(t^n_{i-1}))\|_{H_1}\,ds\\&\le L_2\sup_{\substack{t^n_{i-1}<s\le t^n_i\\ i=1,\ldots,m}}\|u^n(s)-u^n(t^n_{i-1})\|_{H_1}\Big|\sum_{i=1}^m\int\limits_{t^n_{i-1}}^{t^n_i}ds\Big|\\&=L_2T\sup_{\substack{t^n_{i-1}<s\le t^n_i\\ i=1,\ldots,m}}\|u^n(s)-u^n(t^n_{i-1})\|_{H_1}\,.\end{aligned}$$

Therefore, for each $\varepsilon>0$,

$$\begin{aligned}E[\sup_{t^n_m\le T}\|\widehat{I}_3\|_{H_1}]&\le L_3E[\sup_{t^n_m\le T}\sup_{\substack{t^n_{i-1}<s\le t^n_i\\ i=1,\ldots,m}}\|u^n(s)-u^n(t^n_{i-1})\|_{H_1}]\\&\le L_3(E[\sup_{t^n_m\le T}\sup_{\substack{t^n_{i-1}<s\le t^n_i\\ i=1,\ldots,m}}\|u^n(s)-u^n(t^n_{i-1})\|_{H_1}]^2)^{1/2}\end{aligned}$$

and so

$$P(\sup_{t^n_m\le T}\|\widehat{I}_3\|_{H_1}\ge\varepsilon)\le\frac{1}{\varepsilon}E[\sup_{t^n_m\le T}\|\widehat{I}_3\|_{H_1}]\to 0\qquad\text{as } n\to\infty\,.$$

Similarly, we estimate

$$\begin{aligned}\|\widehat{I}_4\|_{H_1} &= \Big\| \sum_{i=1}^{m} \int\limits_{t_{i-1}^n}^{t_i^n} (S(t_m^n - s) - S(t_m^n - t_{i-1}^n))\mathcal{C}(u^n(t_{i-1}^n))\, ds \Big\|_{H_1} \\ &\le \sum_{i=1}^{m} \int\limits_{t_{i-1}^n}^{t_i^n} \|(S(s - t_{i-1}^n) - I)S(t_m^n - s)\mathcal{C}(u^n(t_{i-1}^n))\|_{H_1}\, ds \\ &\le L_4 \sup_{\substack{t_{i-1}^n < s \le t_i^n \\ i=1,\dots,m}} \|(S(s - t_{i-1}^n) - I)u^n(t_{i-1}^n)\|_{H_1} \Big| \sum_{i=1}^{m} \int\limits_{t_{i-1}^n}^{t_i^n} \exp(\beta(t_m^n - s))\, ds \Big| .\end{aligned}$$

Therefore, for each $\varepsilon > 0$,

$$\begin{aligned}E[\sup_{t_m^n \le T} \|\widehat{I}_4\|_{H_1}] &\le L_4 E[\sup_{t_m^n \le T} \sup_{\substack{t_{i-1}^n < s \le t_i^n \\ i=1,\dots,m}} \|(S(s - t_{i-1}^n) - I)u^n(t_{i-1}^n)\|_{H_1} \\ &\le L_5 (E[\sup_{t_m^n \le T} \sup_{\substack{t_{i-1}^n < s \le t_i^n \\ i=1,\dots,m}} \|(S(s - t_{i-1}^n) - I)u^n(t_{i-1}^n)\|_{H_1}]^2)^{1/2}\end{aligned}$$

and therefore

$$P(\sup_{t_m^n \le T} \|\widehat{I}_4\|_{H_1} \ge \varepsilon) \le \frac{1}{\varepsilon} E[\sup_{t_m^n \le T} \|\widehat{I}_4\|_{H_1}] \to 0 \qquad \text{as } n \to \infty .$$

Further, we apply the fact that

$$E\Big[\frac{\dot{w}_j^n(s)\sqrt{h_n}}{\sqrt{\lambda_j}}\Big]^2 = 1 .$$

To estimate $\widehat{I}_5$ we use (A1), (3.4.18), a Chebyshev type inequality and the Schwarz inequality to get

$$\begin{aligned}\|\widehat{I}_5\|_{H_1} &= \Big\| \frac{1}{\sqrt{h_n}} \sum_{i=1}^{m} \sum_{j=0}^{\infty} \sqrt{\lambda_j} \\ &\quad \times \int\limits_{t_{i-1}^n}^{t_i^n} (S(s - t_{i-1}^n) - I)S(t_m^n - s)\mathcal{B}(u^n(t_{i-1}^n))(e_j) \frac{\sqrt{h_n}}{\sqrt{\lambda_j}} \dot{w}_j^n(s)\, ds \Big\|_{H_1} \\ &\le L_6 \frac{1}{\sqrt{h_n}} \sup_{\substack{t_{i-1}^n < s \le t_i^n \\ i=1,\dots,m}} \|(S(s - t_{i-1}^n) - I)u^n(t_{i-1}^n)\|_{H_1} \\ &\quad \times \sum_{j=0}^{\infty} \sqrt{\lambda_j} \Big| \sum_{i=1}^{m} \int\limits_{t_{i-1}^n}^{t_i^n} \exp(\beta(t_m^n - s)) \frac{\sqrt{h_n}}{\sqrt{\lambda_j}} \dot{w}_j^n(s)\, ds \Big| .\end{aligned}$$

Therefore, for each $\varepsilon > 0$,

$$
\begin{aligned}
&E[\sup_{t_m^n\le T}\|\widehat{I}_5\|_{H_1}]\\
&\quad\le L_7\frac{1}{\sqrt{h_n}}E\Big[\sup_{t_m^n\le T}\sup_{\substack{t_{i-1}^n<s\le t_i^n\\ i=1,\dots,m}}\|(S(s-t_{i-1}^n)-I)\mathcal{B}(u^n(t_{i-1}^n))\|_{H_1}\\
&\qquad\times\sup_{t_m^n\le T}\Big|\sum_{i=1}^m\int\limits_{t_{i-1}^n}^{t_i^n}\frac{\sqrt{h_n}}{\sqrt{\lambda_j}}\dot{w}_j^n(s)\,ds\Big|\Big]\\
&\quad\le L_8\frac{1}{\sqrt{h_n}}(E[\sup_{t_m^n\le T}\sup_{\substack{t_{i-1}^n<s\le t_i^n\\ i=1,\dots,m}}\|(S(s-t_{i-1}^n)-I)u^n(t_{i-1}^n)\|_{H_1}]^2)^{1/2}\to 0
\end{aligned}
$$

as $n\to\infty$. Thus, for each $\varepsilon > 0$,

$$
P(\sup_{t_m^n\le T}\|\widehat{I}_5\|_{H_1}\ge\varepsilon)\le\frac{1}{\varepsilon}E[\sup_{t_m^n\le T}\|\widehat{I}_5\|_{H_1}]\to 0\qquad\text{as } n\to\infty\,.
$$

Also using (A1), (A5), (3.4.15), a Chebyshev type inequality and the Schwarz inequality we get

$$
\begin{aligned}
\|\widehat{I}_6\|_{H_1} &= \Big\|\frac{1}{\sqrt{h_n}}\sum_{i=1}^m\sum_{j=0}^\infty\sqrt{\lambda_j}\int\limits_{t_{i-1}^n}^{t_i^n}S(t_m^n-s)\Big(\int\limits_{t_{i-1}^n}^{s}D\mathcal{B}(u^n(\tau))\\
&\qquad\circ\mathcal{A}u^n(\tau)\,d\tau\Big)(e_j)\frac{\sqrt{h_n}}{\sqrt{\lambda_j}}\dot{w}_j^n(s)\,ds\Big\|_{H_1}\\
&\le L_9\frac{1}{\sqrt{h_n}}\Big|\sum_{i=1}^m\int\limits_{t_{i-1}^n}^{t_i^n}\exp(\beta(t_m^n-s))\Big(\int\limits_{t_{i-1}^n}^{s}d\tau\Big)\frac{\sqrt{h_n}}{\sqrt{\lambda_j}}\dot{w}_j^n(s)\,ds\Big|\,.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
&E[\sup_{t_m^n\le T}\|\widehat{I}_6\|_{H_1}]\\
&\quad\le L_{10}\frac{1}{\sqrt{h_n}}\bigg(E\Big[\sup_{t_m^n\le T}\Big|\sum_{i=1}^m\int\limits_{t_{i-1}^n}^{t_i^n}\exp(\beta(t_m^n-s))\\
&\qquad\qquad\times\Big(\int\limits_{t_{i-1}^n}^{s}d\tau\Big)\frac{\sqrt{h_n}}{\sqrt{\lambda_j}}\dot{w}_j^n(s)\,ds\Big|\Big]^2\bigg)^{1/2}\\
&\quad\le L_{11}\frac{h_n}{\sqrt{h_n}}\to 0
\end{aligned}
$$

as $n \to \infty$, because for $v = (s - t_{i-1}^n)/h_n$ we get

$$(*) \qquad \begin{aligned}\widetilde{\beta} &= \sum_{i=1}^{m} \int_{t_{i-1}^n}^{t_i^n} \exp(\beta(t_m^n - s))\Big(\int_{t_{i-1}^n}^{s} d\tau\Big)\, ds \\ &= h_n \sum_{i=1}^{m} \int_{t_{i-1}^n}^{t_i^n} \exp(\beta(t_m^n - s)) \frac{s - t_{i-1}^n}{h_n}\, ds \\ &= h_n \sum_{i=1}^{m} \int_{0}^{1} \exp(\beta(t_m^n - t_{i-1}^n)) \exp(\beta(-h_n)v) v h_n \, dv \\ &\le h_n^2 \sum_{i=1}^{m} \exp(\beta(t_m^n - t_{i-1}^n)) \\ &\le h_n^2 m \exp(\beta T) \le L_{12} h_n.\end{aligned}$$

Thus, for each $\varepsilon > 0$,

$$P(\sup_{t_m^n \le T} \|\widehat{I}_6\|_{H_1} \ge \varepsilon) \le \frac{1}{\varepsilon} E[\sup_{t_m^n \le T} \|\widehat{I}_6\|_{H_1}] \to 0 \qquad \text{as } n \to \infty\,.$$

In a similar way we estimate $\widehat{I}_7$.

Further, using (A1), (A4), $(*)$, (3.4.15), the Schwarz inequality and a Chebyshev type inequality, we obtain

$$\begin{aligned}\|\widehat{I}_8\|_{H_1} = \Bigg\| & \frac{1}{\sqrt{h_n}} \frac{1}{\sqrt{h_n}} \sum_{j=0}^{\infty} \sqrt{\lambda_j} \sum_{k=0}^{\infty} \sqrt{\lambda_k} \\ & \times \sum_{i=1}^{m} \int_{t_{i-1}^n}^{t_i^n} S(t_m^n - s) \Bigg(\int_{t_{i-1}^n}^{s} (D\mathcal{B}(u^n(\tau))\mathcal{B}(u^n(\tau)) \\ & - D\mathcal{B}(u^n(t_{i-1}^n))\mathcal{B}(u^n(t_{i-1}^n)))(e_j) \\ & \times \frac{\sqrt{h_n}}{\sqrt{\lambda_j}} \dot{w}_j^n(\tau)\, d\tau \Bigg)(e_k) \frac{\sqrt{h_n}}{\sqrt{\lambda_k}} \dot{w}_k^n(s)\, ds \Bigg\|_{H_1} \\ \le L_{13} & \frac{1}{h_n} \sup_{\substack{t_{i-1}^n < \tau \le t_i^n \\ i=1,\ldots,m}} \|u^n(\tau) - u^n(t_{i-1}^n)\|_{H_1} \\ & \times \Bigg| \sum_{i=1}^{m} \int_{t_{i-1}^n}^{t_i^n} \exp(\beta(t_m^n - s)) \Bigg(\int_{t_{i-1}^n}^{s} \frac{\sqrt{h_n}}{\sqrt{\lambda_j}} \dot{w}_j^n(\tau)\, d\tau \Bigg) \frac{\sqrt{h_n}}{\sqrt{\lambda_k}} \dot{w}_k^n(s)\, ds \Bigg|\,.\end{aligned}$$

Thus

$$
\begin{aligned}
&E[\sup_{t_m^n\le T}\|\widehat{I}_8\|_{H_1}]\\
&\quad\le L_{14}\frac{1}{h_n}E\Big[\sup_{t_m^n\le T}\sup_{\substack{t_{i-1}^n<\tau\le t_i^n\\ i=1,\ldots,m}}\|u^n(\tau)-u^n(t_{i-1}^n)\|_{H_1}\\
&\qquad\times\sup_{t_m^n\le T}\Big|\sum_{i=1}^m\int\limits_{t_{i-1}^n}^{t_i^n}\exp(\beta(t_m^n-s))\Big(\int\limits_{t_{i-1}^n}^{s}\frac{\sqrt{h_n}}{\sqrt{\lambda_j}}\dot{w}_j^n(\tau)\,d\tau\Big)\frac{\sqrt{h_n}}{\sqrt{\lambda_k}}\dot{w}_k^n(s)\,ds\Big|\Big]\\
&\quad\le L_{15}\frac{h_n}{h_n}(E[\sup_{t_m^n\le T}\sup_{t_{i-1}^n<\tau\le t_i^n}\|u^n(\tau)-u^n(t_{i-1}^n)\|_{H_1}]^2)^{1/2}\\
&\quad\to 0\quad\text{ as } n\to\infty,
\end{aligned}
$$

and hence for each $\varepsilon>0$,

$$
P(\sup_{t_m^n\le T}\|\widehat{I}_8\|_{H_1}\ge\varepsilon)\le\frac{1}{\varepsilon}E[\sup_{t_m^n\le T}\|\widehat{I}_8\|_{H_1}]\to 0\quad\text{ as } n\to\infty\,.
$$

Finally, using (A1), (A4), $(*)$, (3.4.15), the Schwarz inequality and a Chebyshev type inequality, in the same way as for $\widehat{I}_5$ we have

$$
\begin{aligned}
&\|\widehat{I}_9\|_{H_1}\\
&=\Big\|\frac{1}{\sqrt{h_n}}\frac{1}{\sqrt{h_n}}\sum_{j=0}^{\infty}\sqrt{\lambda_j}\sum_{k=0}^{\infty}\sqrt{\lambda_k}\\
&\quad\times\sum_{i=1}^m\int\limits_{t_{i-1}^n}^{t_i^n}(S(s-t_{i-1}^n)-I)S(t_m^n-s)\Big(\int\limits_{t_{i-1}^n}^{s}D\mathcal{B}(u^n(t_{i-1}^n))\mathcal{B}(u^n(t_{i-1}^n))(e_j)\\
&\quad\times\frac{\sqrt{h_n}}{\sqrt{\lambda_j}}\dot{w}_j^n(\tau)\,d\tau\Big)(e_k)\frac{\sqrt{h_n}}{\sqrt{\lambda_k}}\dot{w}_k^n(s)\,ds\Big\|_{H_1}\\
&\le L_{16}\sup_{\substack{t_{i-1}^n<s\le t_i^n\\ i=1,\ldots,m}}\|(S(s-t_{i-1}^n)-I)u^n(t_{i-1}^n)\|_H\\
&\quad\times\Big|\sum_{i=1}^m\int\limits_{t_{i-1}^n}^{t_i^n}\exp(\beta(t_m^n-s))\Big(\int\limits_{t_{i-1}^n}^{s}\frac{\sqrt{h_n}}{\sqrt{\lambda_j}}\dot{w}_j^n(\tau)\,d\tau\Big)\frac{\sqrt{h_n}}{\sqrt{\lambda_k}}\dot{w}_k^n(s)\,ds\Big|\,.
\end{aligned}
$$

Therefore

$$
\begin{aligned}
&E[\sup_{t_m^n\le T}\|\widehat{I}_9\|_{H_1}]\\
&\quad\le L_{17}\frac{1}{h_n}E\Big[\sup_{t_m^n\le T}\sup_{\substack{t_{i-1}^n<s\le t_i^n\\ i=1,\dots,m}}\|(S(s-t_{i-1}^n)-I)u^n(t_{i-1}^n)\|_{H_1}\\
&\qquad\times\sup_{t_m^n\le T}\Big|\sum_{i=1}^m\int_{t_{i-1}^n}^{t_i^n}\exp(\beta(t_m^n-s))\Big(\int_{t_{i-1}^n}^{s}\frac{\sqrt{h_n}}{\sqrt{\lambda_j}}\dot{w}_j^n(\tau)\,d\tau\Big)\frac{\sqrt{h_n}}{\sqrt{\lambda_k}}\dot{w}_k^n(s)\,ds\Big|\Big]\\
&\quad\le L_{18}\frac{h_n}{h_n}(E[\sup_{t_m^n\le T}\sup_{t_{i-1}^n<\tau\le t_i^n}\|(S(s-t_{i-1}^n)-I)u^n(t_{i-1}^n)\|_{H_1}]^2)^{1/2}\to 0\,,
\end{aligned}
$$

and for each $\varepsilon>0$,

$$
P(\sup_{t_m^n\le T}\|\widehat{I}_9\|_{H_1}\ge\varepsilon)\le\frac{1}{\varepsilon}E[\sup_{t_m^n\le T}\|\widehat{I}_9\|_{H_1}]\to 0\qquad\text{as } n\to\infty\,.
$$

Now we consider a slight modification of (3.4.17):

$$
\begin{aligned}
(3.4.20)\qquad \widetilde{u}^n(t)&=S(t)z_0+\int_0^t\mathcal{B}_n(\widetilde{u}^n,t,s)\,dw(s)+\int_0^t\mathcal{C}_n(\widetilde{u}^n,t,s)\,ds\\
&\quad+\frac12\sum_{j=0}^\infty\sum_{k=0}^\infty\sum_{i=1}^m S(t-t_{i-1}^n)D\mathcal{B}(u^n(t_{i-1}^n))\mathcal{B}(u^n(t_{i-1}^n))(e_j)(e_k)\Delta_i^n w_j\Delta_i^n w_k\,.
\end{aligned}
$$

By (3.4.13), (3.4.14) and the definition of $G_n(\widehat{\xi}^n;t,s)$, for every $\varepsilon>0$ we get

$$
(3.4.21)\qquad \lim_{n\to\infty}P(\sup_{0\le t\le T}\|\widetilde{u}^n(t,\omega)-\widehat{\xi}^n(t,\omega)\|_{H_1}\ge\varepsilon)=0\,.
$$

By the above estimates and by the continuity of u^n (comparing $\widetilde{u}^n$ in (3.4.20) with u^n in (3.4.17)) we conclude that for every $\varepsilon>0$,

$$
(3.4.22)\qquad \lim_{n\to\infty}P(\sup_{0\le t\le T}\|\widetilde{u}^n(t,\omega)-u^n(t,\omega)\|_{H_1}\ge\varepsilon)=0\,.
$$

Therefore, by (3.4.9), (3.4.21) and (3.4.22) the proof is complete.

3.5. Examples

3.5.1. *Equations satisfying the assumptions of Theorem* 3.4.1. Let us observe that inequality (3.3.6) is satisfied if we take $H=\mathbb{R}$, $\mathcal{B}:H_1\to H_1$ such that $\mathcal{B}(h_1)=b\cdot f(\langle c,h_1\rangle_{H_1})$ (Lurie type), $b\in D(\mathcal{A}^*\mathcal{A})$, $c\in D(\mathcal{A}^*)$ is an eigenvector of $\mathcal{A}^*$, $\mathcal{A}$ has a compact resolvent and a real eigenvalue $\lambda\in\sigma(\mathcal{A}^*)$, $f(x)=\tanh(x)$, $x=\langle c,u\rangle_{H_1}$, for example. Then $D\mathcal{B}(h_1)\mathcal{A}(h_1)=b\cdot f'(\langle c,h_1\rangle_{H_1})\langle c,\mathcal{A}u\rangle_{H_1}=b(1/(\cosh(\langle c,u\rangle_{H_1}))^2)\langle\mathcal{A}^*c,u\rangle_{H_1}$ and $\|D\mathcal{B}(h_1)\mathcal{A}h_1\|_{L(H,H_1)}\le k$. We may define $R=(\mathcal{A}^{-1})(\mathcal{A}^*)^{-1}$ to deduce that assumption (A3) is satisfied.

Now, let us take $H=H_1=L^2([0,1])$ with an orthonormal basis $\{e_i\}_{i=0}^\infty$ and let $\mathcal{A}:H_1\supset D(\mathcal{A})\to H_1$ be the infinitesimal generator of a semigroup of contraction type, that is, $\mathcal{A}\varphi = d\varphi(\theta)/d\theta$, $D(\mathcal{A}) = \{\varphi \in W^{1,2}([0,1]) : \varphi(0) = 0\}$. We consider a Hammerstein operator $\mathcal{C} : H_1 \to H_1$ such that

$$\mathcal{C}(h)(s) = \int_0^1 K(s,t) f(t, h(t))\, dt\,,$$

where $K(s,t) = K_1(s)K_2(t)$ with $K_1(s) \in C^4([0,1])$, $K_2(t) \in C([0,1])$, and $f \in C^1([0,1] \times \mathbb{R})$ with $|f(t,x)| \le a(t) + b|x|$. Similarly we define $\mathcal{B} : H_1 \to L(H, H_1)$ as a Hammerstein operator on H_1, putting

$$\mathcal{B}(h)(e_i)(s) = \int_0^1 K_i(s,t) f(t, h(t))\, dt\,,$$

where $K_i(s,t)$ have similar properties as $K(s,t)$ and they are uniformly bounded. We take an operator $R : H_1 \to H_1$ such that $R = (I - \Delta)^{-1}$, where $\Delta = d^2/dx^2$. Clearly, R is positive definite, symmetric and nuclear. Let us compute

$$D\mathcal{B}_{x_0}(h)(e_i)(s) = \int_0^1 K_i(s,t) f'_x(t, x_0(t)) h(t)\, dt\,.$$

Then assumptions (A1)–(A5) are satisfied. For example, let us verify (A3)(i) for $\widetilde{K}$, $K > 0$:

$$\begin{aligned}
\|R^{-1}\mathcal{C}(h_1)\|^2_{H_1} &= \sum_{n=0}^{\infty} |\langle R^{-1}\mathcal{C}(h_1), e_n\rangle_{H_1}|^2 \\
&= \sum_{n=0}^{\infty} \left| \left\langle \mathcal{C}(h_1), \left(1 - \frac{d^2}{dx^2}\right) e_n \right\rangle_{H_1} \right|^2 \\
&= \widetilde{K} \sum_{n=0}^{\infty} |\langle \mathcal{C}(h_1), (1+n^2)e_n\rangle_{H_1}|^2 \\
&\le \widetilde{K} \sum_{n=0}^{\infty} (|\langle \mathcal{C}(h_1), e_n\rangle_{H_1}|^2 + n^4 |\langle \mathcal{C}(h_1), e_n\rangle_{H_1}|^2) \\
&\le \widetilde{K}\|\mathcal{C}(h_1)\|^2_{H_1} + \widetilde{K}(1 + \|h_1\|^2_{H_1}) \le K(1 + \|h_1\|^2_{H_1})
\end{aligned}$$

because

$$\begin{aligned}
|\langle \mathcal{C}(h_1), e_n\rangle_{H_1}|^2 &= \Big| \int_0^1 \int_0^1 K(s,t) f(t, h_1(t))\, dt\, e_n(s)\, ds \Big|^2 \\
&\le \Big| \int_0^1 K_1(s) e_n(s)\, ds \int_0^1 K_2(t)(a(t) + b|h_1(t)|)\, dt \Big|^2 \\
&\le \frac{\widetilde{K}}{n^4} \Big(1 + b\Big(\int_0^1 |K_2(t)|^2\, dt \Big) \|h_1\|^2_{H_1} \Big) \\
&\le \frac{\widetilde{K}}{n^4} (1 + \|h_1\|^2_{H_1})\,.
\end{aligned}$$

Now we verify (A5) under a suitable definition of the function $f(t,x)$:

$$
\begin{aligned}
(D\mathcal{B}(h_1)\mathcal{A}h_1(e_i))(s) &= \int_0^1 K_i(s,t)f'_x(t,h_1(t))h'_1(t)\,dt \\
&= K_i(s,1)f'_x(1,h_1(1))h_1(1) - K_i(s,0)f'_x(0,h_1(0))h_1(0) \\
&\quad - \int_0^1 K'_{it}(s,t)f'_x(t,h_1(t))h_1(t)\,dt \\
&\quad - \int_0^1 K_i(s,t)f''_{tx}(t,h_1(t))h_1(t)\,dt \\
&\quad - \int_0^1 K_i(s,t)f''_{xx}(t,h_1(t))h'_1(t)h_1(t)\,dt
\end{aligned}
$$

and $\|D\mathcal{B}(h_1)\mathcal{A}h_1\|_{L(H,H_1)} \le k\|h_1\|_{H_1}$ if, for example, $f'_x(t,h_1(t)) = f''_{xx}(t,h_1(t)) \times h_1(t)$.

Finally, let us take $H = L^2([0,1])$ with an orthonormal basis $\{e_i\}_{i=0}^{\infty}$. Let H_1 be a real separable Hilbert space. For example, our assumptions are also satisfied by a linear operator $\mathcal{B} : H_1 \to L(H,H_1)$ such that $\mathcal{B}(h_1)(e_i) = b\langle c_i, h_1\rangle_{H_1}(e_i)$ ($\langle c_i, h_1\rangle_{H_1}$ are uniformly bounded) for a given $b \in H_1$ and fixed $c_i \in H_1$. Set $\mathcal{A} = \Delta^{\alpha}$, $\alpha > \frac{1}{2}$, $R^{-1} = (-\Delta)^{\alpha}$, $b \in D(-\mathcal{A})^{\alpha}$, $c_i \in D(\mathcal{A}^*)$. We observe that

$$\mathcal{B}_{h_1}(h_1)\mathcal{A}h_1(e_i) = b\langle c_i, \mathcal{A}h_1\rangle_{H_1}(e_i) = b\langle \mathcal{A}^*c_i, h_1\rangle_{H_1}(e_i)\,,$$

and assumption (3.3.5) is satisfied.

3.5.2. *Stochastic delay equations.* We first introduce the space M^2. For fixed $r \in \mathbb{R}_+$ we put $I = [-r,0]$ and $M^2 = \mathbb{R}^n \times L^2(I,\mathbb{R}^n)$. The elements of M^2 are denoted as follows:

$$\begin{pmatrix} a \\ \varphi \end{pmatrix} \in M^2 \qquad \text{for } a \in \mathbb{R}^n,\ \varphi \in L^2(I,\mathbb{R}^n)\,.$$

In M^2 the natural inner product is used:

$$\left\langle \begin{pmatrix} a_1 \\ \varphi_1 \end{pmatrix}, \begin{pmatrix} a_2 \\ \varphi_2 \end{pmatrix} \right\rangle_{M^2} = \langle a_1, a_2\rangle + \langle \varphi_1, \varphi_2\rangle_{L^2(I,\mathbb{R}^n)}\,.$$

Equation (3.3.1) can be a model for stochastic delay equations of semilinear type:

$$
\begin{aligned}
dx(t) &= \Big(\sum_{i=0}^{m} A_i x(t-r_i) + \int_{-r}^{0} A(\theta)x(t+\theta)\,d\theta\Big)dt + \Sigma(x_t)dw(t)\,, \\
x_0(\theta) &= \psi(\theta)\,,
\end{aligned}
\tag{3.5.1}
$$

where $t \in [0,T]$ and $r_i \in \mathbb{R}_+$ are fixed, $0 = r_0 < \ldots < r_m = r$.

The following hypotheses are assumed:

(B1) $(x(t))_{t\in[0,T]}$ is an $\mathbb{R}^n$-valued stochastic process and $(w(t))_{t\geq 0}$ is an H-valued Wiener process, $x_t(\theta) = x(t+\theta)$ for $\theta \in I$, $t \in [0,T]$, and A_i, $A(\theta)$ are $n \times n$ matrices, the elements of $A(\theta)$ being square-integrable on I, $\Sigma : L^2(I, \mathbb{R}^n) \to L(H, \mathbb{R}^n)$ is an operator,

(B2) $\psi : I \times \Omega \to \mathbb{R}^n$ is $\mathfrak{B}_{(I)} \times \mathfrak{F}_0$-measurable and such that
$$\begin{pmatrix} \psi(0) \\ \psi(\cdot) \end{pmatrix} \in M^2 \quad \text{and} \quad E[\|\psi\|^2_{M^2}] < \infty\,,$$
where $\mathfrak{B}_{(I)}$ is the σ-algebra of Borel sets on I,

(B3) the operator Σ satisfies the growth and Lipschitz conditions analogous to (A3) in §3.3,

(B4) $\Sigma \in C^1_{\mathrm{b}}$ and its derivative is globally Lipschitz (see §3.3),

(B5) the operator Σ satisfies the condition analogous to (A5) in §3.3.

It is known from [5], [77], [81] that there is exactly one solution $(x(t))_{t\in[0,T]}$ to equation (3.5.1) in the sense of Definition 3.3.1.

Remark 3.5.1. The linear term in equation (3.5.1) is sufficiently general ([5]) to include most linear autonomous functional differential equations arising in applications.

To write (3.5.1) in the form (3.3.1) we denote by $\widetilde{A}$ the infinitesimal generator of a contraction semigroup $(T(t))_{t\geq 0}$ on $L^2(I, \mathbb{R}^n)$. Let
$$D(\widetilde{A}) = \{\varphi \in W^{1,2}(I, \mathbb{R}^n) : \varphi(0) = 0\}, \qquad \widetilde{A}\varphi = \frac{d\varphi}{d\theta}$$
and
$$[T(t)\varphi(\cdot)](\theta) = \begin{cases} \varphi(t+\theta) & \text{for } t \leq -\theta, \\ 0 & \text{otherwise,} \end{cases}$$
where $t > 0$, $\theta \in I$.

In case $\Sigma \equiv 0$ we define a family $(\widetilde{S}(t))_{t\geq 0}$ of operators acting on M^2 by
$$\widetilde{S}(t)\begin{pmatrix} a \\ \varphi \end{pmatrix} = \begin{pmatrix} x(t) \\ x_t(\cdot) \end{pmatrix} \qquad \text{for } \begin{pmatrix} a \\ \varphi \end{pmatrix} \in M^2\,.$$
The family $(\widetilde{S}(t))_{t\geq 0}$ is a C_0-semigroup of bounded linear operators. Following the idea used in [5], p. 500, we can introduce an equivalent norm in M^2 such that $(\widetilde{S}(t))_{t\geq 0}$ becomes a semigroup of contraction type.

Now we rewrite (3.5.1) in the following form for $z(t) = \binom{x(t)}{x_t(\cdot)}$:

(3.5.2)
$$dz(t) = \begin{pmatrix} \sum\limits_{i=0}^{m} A_i x(t-r_i) + \int\limits_{-r}^{0} A(\theta)x(t+\theta)\,d\theta \\ \widetilde{A}x_t(\cdot) \end{pmatrix} dt + \widehat{\Sigma}(x_t(\cdot))\,dw(t)\,,$$
$$z(0) = z_0\,,$$

where for arbitrary $\xi \in L^2(I,\mathbb{R}^n)$ we define $\widehat{\Sigma}(\xi) : H \to M^2$ by

$$\widehat{\Sigma}(\xi)h = \begin{pmatrix} \Sigma(\xi)h \\ 0 \end{pmatrix} \qquad \text{for every } h \in H\,.$$

We now define the operators $\mathcal{A}$ and $\mathcal{B}$ of (3.3.1) for our example. Let

$$D(\mathcal{A}) = \left\{ \begin{pmatrix} \varphi(0) \\ \varphi(\cdot) \end{pmatrix} : \varphi \in W^{1,2}(I,\mathbb{R}^n) \right\},$$

$$\mathcal{A}\begin{pmatrix} \varphi(0) \\ \varphi(\cdot) \end{pmatrix} = \begin{pmatrix} \sum\limits_{i=0}^{m} A_i\varphi(r_i) + \int\limits_{-r}^{0} A(\theta)\varphi(\theta)\,d\theta \\ \widetilde{A}\varphi \end{pmatrix}$$

and

$$\mathcal{B}\begin{pmatrix} a \\ \varphi(\cdot) \end{pmatrix} = \begin{pmatrix} b \\ \psi(\cdot) \end{pmatrix} \equiv \widehat{\Sigma}(\varphi(\cdot))\,,$$

therefore for $h \in H$

$$\mathcal{B}\begin{pmatrix} a \\ \varphi(\cdot) \end{pmatrix}(h) = \begin{pmatrix} \Sigma(\varphi(\cdot))(h) \\ 0 \end{pmatrix}.$$

Here $\mathcal{A} : M^2 \supset D(\mathcal{A}) \to M^2$, $\mathcal{B} : M^2 \to L(H, M^2)$. We take $H_1 = M^2$ and now (3.5.1) has the form analogous to (3.3.1).

We observe that the term which is needed for the construction of the correction term in (3.3.2) is in this case

$$\begin{aligned} D\mathcal{B}(h_1)\mathcal{B}(h_1) &= D\mathcal{B}\begin{pmatrix} a \\ \varphi(\cdot) \end{pmatrix}\begin{pmatrix} b \\ \psi(\cdot) \end{pmatrix} \\ &= \begin{pmatrix} D\Sigma(\varphi(\cdot))(\psi(\cdot)) \\ 0 \end{pmatrix} \qquad \text{for } \begin{pmatrix} a \\ \varphi(\cdot) \end{pmatrix} = h_1 \in M^2\,. \end{aligned}$$

3.5.3. *Stochastic wave equations.* Consider (see [14], [17], [56]) the initial value problem for the wave equation in $\mathbb{R}^n$:

$$\frac{\partial^2 z}{\partial t^2} = \Delta z + \eta(t,\xi) \qquad \text{for } \xi \in \mathbb{R}^n,\ t > 0\,,$$

$$z(0,\xi) = z_0(\xi), \quad \frac{\partial z}{\partial t}(0,\xi) = z_1(\xi) \qquad \text{for } \xi \in \mathbb{R}^n\,,$$

where $\eta(t,\xi)$ represents a noise disturbance.

In the Hilbert space $H = H^1(\mathbb{R}^n)\times L^2(\mathbb{R}^n)$ we define the operator $\mathcal{A}$ as follows ($H^i(\mathbb{R}^n)$ are the usual Sobolev spaces for $i = 1, 2$):

$$D(\mathcal{A}) = H^2(\mathbb{R}^n) \times H^1(\mathbb{R}^n)$$

and for $z = (z_1, z_2) \in D(\mathcal{A})$ let (Δ denotes the Laplacian)

$$\mathcal{A}z = \begin{pmatrix} 0 & I \\ \Delta & 0 \end{pmatrix}\begin{pmatrix} z_1 \\ z_2 \end{pmatrix}.$$

The operator $\mathcal{A}$ is the infinitesimal generator of a C_0 group $T(t)$ of operators on H satisfying (compare [56], p. 219)

$$\|T(t)\|_{H_1} \le \exp(2|t|)\,.$$

We obtain the model

$$\begin{pmatrix} z \\ \frac{\partial z_t}{\partial t} \end{pmatrix} = T(t)\begin{pmatrix} z_0 \\ \frac{\partial z_0}{\partial t} \end{pmatrix} + \int_0^t T(t-s)\begin{pmatrix} 0 \\ I \end{pmatrix} dw(s)\,.$$

4. Comparison of the results

4.1. Finite-dimensional case. We consider the case (see Chapter 3) where $H = \mathbb{R}^d$, $H_1 = \mathbb{R}^n$ (see [27], [67]). Let $x, z \in \mathbb{R}^n$. Then $\mathcal{B} : \mathbb{R}^n \to L(\mathbb{R}^d, \mathbb{R}^n)$, $D\mathcal{B}(z) \in L(\mathbb{R}^n, L(\mathbb{R}^d, \mathbb{R}^n))$ and $D\mathcal{B}(z)(x) \in L(\mathbb{R}^d, \mathbb{R}^n)$ is given by a matrix

$$\widehat{A}(z)(x) = \begin{bmatrix} D\mathcal{B}_{11}(z)(x) & \ldots & D\mathcal{B}_{1d}(z)(x) \\ \ldots & \ldots & \ldots \\ D\mathcal{B}_{n1}(z)(x) & \ldots & D\mathcal{B}_{nd}(z)(x) \end{bmatrix}.$$

We put

$$\mathbb{R}^d \ni X = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_d \end{pmatrix}, \qquad \mathbb{R}^d \ni Y = \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_d \end{pmatrix}.$$

Let $D\mathcal{B}(z)\mathcal{B}(z)X = \Psi(X)$. Then for $i = 1, \ldots, n$,

$$\Psi(X)Y = \sum_{j=1}^{d}\Big(\sum_{l=1}^{n} \frac{\partial \mathcal{B}_{ij}(z)}{\partial z_l}\Big(\sum_{k=1}^{d} \mathcal{B}_{lk}(z)\xi_k\Big)\Big)\eta_j$$

and for $p \in \mathbb{R}^n$

$$\langle \Psi(X)Y, p\rangle_{\mathbb{R}^n} = \sum_{i=1}^{n}\sum_{j=1}^{d}\sum_{l=1}^{n}\sum_{k=1}^{d} \frac{\partial \mathcal{B}_{ij}(z)}{\partial z_l}\mathcal{B}_{lk}(z)\xi_k\eta_j p_i\,.$$

We omit X and Y in the last sum and obtain the matrix

$$\Big(\sum_{i=1}^{n}\sum_{l=1}^{n} \frac{\partial \mathcal{B}_{ij}(z)}{\partial z_l}\mathcal{B}_{lk}(z)p_i\Big)_{jk} = (\Psi_{jk})_{j,k=1,\ldots,d} = \widetilde{\Psi}(p)\,.$$

Consider its trace

$$\operatorname{tr}\widetilde{\Psi}(p) = \sum_{j=1}^{d}\Psi_{jj} = \sum_{j=1}^{d}\sum_{i=1}^{n}\sum_{l=1}^{n} \frac{\partial \mathcal{B}_{ij}(z)}{\partial z_l}\mathcal{B}_{lj}(z)p_i\,.$$

We rewrite it in the form of the inner product of two vectors in $\mathbb{R}^n$ (cf. (3.3.3)):

$$\operatorname{tr}\widetilde{\Psi}(p) = \sum_{i=1}^{n}\Big(\sum_{j=1}^{d}\sum_{l=1}^{n} \frac{\partial \mathcal{B}_{ij}(z)}{\partial z_l}\mathcal{B}_{lj}(z)\Big)p_i\,.$$

The first vector

$$\left(\sum_{j=1}^{d}\sum_{l=1}^{n}\frac{\partial \mathcal{B}_{ij}(z)}{\partial z_l}\mathcal{B}_{lj}(z)\right)_i, \qquad i=1,\dots,n\,,$$

is exactly the correction term $\approx\!\!\!\!\!h_1$ obtained in [27], [67] for the finite-dimensional case.

4.2. Stochastic delay equations. Now we would like to compare two types of correction terms derived for stochastic delay equations in §2.3 and in §3.5 for the different spaces occurring in these two models. In §3.5 we have

$$\mathcal{B}(h_1)=\begin{pmatrix}\Sigma(\varphi(\cdot))\\ 0\end{pmatrix}=\begin{pmatrix}\psi(0)\\ \psi(\cdot)\end{pmatrix}\quad \text{for } h_1=\begin{pmatrix}\varphi(0)\\ \varphi(\cdot)\end{pmatrix}, \tag{4.2.1}$$

and hence for the one-dimensional Wiener process

$$D\mathcal{B}(h_1)\begin{pmatrix}\psi(0)\\ \psi(\cdot)\end{pmatrix}=\begin{pmatrix}D\Sigma(\varphi(\cdot))\psi(\cdot)\\ 0\end{pmatrix}=\begin{pmatrix}\alpha\psi(0)+\int_{-r}^{0}\psi(s)u(s)\,ds\\ 0\end{pmatrix}, \tag{4.4.2}$$

$$D\mathcal{B}(h_1)\mathcal{B}(h_1)=\begin{pmatrix}\alpha\Sigma(\varphi)\\ 0\end{pmatrix}. \tag{4.2.3}$$

The last equality holds because the second coordinate in (4.2.1) is zero. We observe that if $\alpha=\mu(\{0\})$ and $u(s)$ is the density of the measure $\widetilde{\mu}$ (μ and $\widetilde{\mu}$ are defined in §2.3), then the term

$$\alpha\psi(0)+\int_{-r}^{0}\psi(s)u(s)\,ds$$

is the correction term in Chapter 2.

5. On the relation between the Itô and Stratonovich integrals in Hilbert spaces

The correction term introduced in Chapter 3 is the same as the one obtained in the course of transition from the Itô integral to the Stratonovich integral in [80].

As is well known, the Itô stochastic integral is convenient in some problems because it is a martingale. However, the Stratonovich integral is particularly suitable for applications to problems described by stochastic differential equations because it does not give any correction term in the approximation theorem of Wong–Zakai type. Also, the rules of classical calculus apply to that integral.

For this reason, the Stratonovich integral has been much discussed in the literature ([2], [18], [66]).

In the sequel, we use the notation of Chapter 3.

DEFINITION 5.1 (see [80]). We define the Stratonovich integral for an operator $\phi : [0,T] \times H_1 \to L(H, H_1)$ by

$$(5.1)\qquad (\mathrm{S})\int_a^b \phi(t, z(t))\, dw(t) = \lim_{n\to\infty} S_n$$

$$= \lim_{n\to\infty} \sum_{j=1}^n \phi\bigg(\frac{1}{2}(t_j^n + t_{j-1}^n), \frac{1}{2}(z(t_j^n) + z(t_{j-1}^n))\bigg)(w(t_j^n) - w(t_{j-1}^n))\,.$$

The limit is in the P-a.s. sense and $a = t_0^n < t_1^n < \ldots < t_n^n = b$ is a partition of the interval $[a, b]$. Set $h^n = t_j^n - t_{j-1}^n$ for $j = 1, \ldots, n$. We assume that the sequence of partitions is normal and the limit does not depend on the choice of the partition. The operator ϕ is continuous with respect to the first variable and it has the same properties as the operator $\mathcal{B}$ in §3.3 with respect to the second one.

We recall the definition of the Itô integral:

$$(5.2)\qquad (\mathrm{I})\int_a^b \phi(t, z(t))\, dw(t) = \lim_{j\to\infty} I_n$$

$$= \lim_{n\to\infty} \sum_{j=1}^n \phi(t_{j-1}^n, z(t_{j-1}^n))(w(t_j^n) - w(t_{j-1}^n))\,,$$

with the same assumptions as in Definition 5.1.

Moreover, this integral is a continuous, square-integrable H_1-valued martingale (see [17], [46]).

The following theorem is proved in [80]:

THEOREM 5.1. *Consider an operator* $\phi : [0,T] \times H_1 \to L(H, H_1)$ *satisfying the assumptions of Definition* 5.1 *and Lipschitz with respect to the first variable. Let* $(z(t))_{t\in[0,T]}$ *be the mild solution to the stochastic differential equation* (3.3.1). *Then the Stratonovich integral* (5.1) *exists and*

$$(5.3)\qquad (\mathrm{S})\int_0^t \phi(s, z(s))\, dw(s) = (\mathrm{I})\int_0^t \phi(s, z(s))\, dw(s)$$

$$+ \frac{1}{2}\int_0^t \widetilde{\mathrm{tr}}(QD_z\phi(s, z(s))\mathcal{B}(z(s)))\, ds\,.$$

Remark 5.1. If we put $\phi(t, z(t)) = \mathcal{B}(z(t))$ then the correction term in (5.3) has the form

$$\frac{1}{2}\int_0^t \widetilde{\mathrm{tr}}(QD\mathcal{B}(z(s))\mathcal{B}(z(s)))\, ds\,.$$

It is the same correction term that occurs in the approximation theorem of Wong–Zakai type.

R e m a r k 5.2. Our original problem in this chapter was to establish the transformation rule for the above types of integrals. In [80] it is pointed out that if we evaluate the integral as the limit value of approximating sums then the type of the result depends on the choice of intermediate points.

We can now define (see [2]) the stochastic differential equation in terms of the integral

$$(\mathrm{S})\int_0^t (0,\mathcal{B}(z(s)))\,d\begin{pmatrix} z(s)\\ w(s)\end{pmatrix} = (\mathrm{S})\int_0^t \mathcal{B}(z(s))\,dw(s)\,.$$

Then the Stratonovich equation can be symbolically written as

(S) $$dz(t) = \mathcal{A}z(t)dt + \mathcal{C}(z(t))dt + \mathcal{B}(z(t))dw(t), \qquad z(0) = z_0\,.$$

In our case, where $\Phi = (0,\mathcal{B})$ and the driven process is $\begin{pmatrix} z(s)\\ w(s)\end{pmatrix}$, the transition formula (5.3) yields the Itô equation

(I) $$\begin{aligned} dz(t) &= \mathcal{A}z(t)dt + \mathcal{C}(z(t))dt \\ &\qquad + \tfrac{1}{2}\widetilde{\mathrm{tr}}(QD\mathcal{B}(z(t))\mathcal{B}(z(t)))dt + \mathcal{B}(z(t))dw(t), \\ z(0) &= z_0 \end{aligned}$$

corresponding to the above Stratonovich equation in the sense of coincidence of solutions.

6. Conclusions

A new form of the correction term was computed for functional stochastic differential equations defined on some spaces of continuous functions (see also [79]). A new form of the correction term was also introduced for semi-linear stochastic evolution equations in Hilbert spaces. In the latter case the Wong–Zakai type theorem was proved for the first time in [78] with this correction term although the term itself had already appeared in the literature [19].

As already mentioned, these correction terms also appear in formulas giving the relations between the Itô and Stratonovich integrals ([18], [80]), as was to be expected from the one-dimensional case.

Another important application of the Wong–Zakai theorems is that they constitute an important part of the proofs of theorems on the support of measures connected with solutions of the appropriate stochastic differential equations. Namely, our future aim will be to describe the topological support of a measure connected with the solution. We expect that it equals the closure of a set of solutions to differential equations obtained when the noise process is approximated by a sequence of smooth functions. The correction terms considerd here also give some indications for computing the Itô formulas for stochastic differential equations considered in this paper.

Finally, one of the most important applications of new correction terms may be in filtering theory when their application can simplify the filtering equations in models with coloured noises. When the input and measurement noise in a model are not white noises but they are independent of each other, the construction of a filter is possible but it has a more complicated character. A standard way is to model these noises as the output of another model driven by the white noise. After the Kalman filter is constructed, it may turn out to be only suboptimal. The question arises how to design the filter equations by applying to coloured noise the approximation procedure and the correction term of the Wong–Zakai type approximation theorem.

More exactly, it is well known that the signal process $x(t)$ and the observation process $y(t)$ are governed by stochastic differential equations. The best estimate of $x(t)$, when $y(s)$ is known for any $s \le t$, is given by the conditional expectation $E(x(t) \mid \mathfrak{F}_t^y)$, where $\mathfrak{F}_t^y$ is the σ-algebra generated by $y(s)$ for $0 \le s \le t$. The filtering problem consists in comparing, under general assumptions, the conditional distribution of $x(t)$ given $\{y(s) : s \le t\}$ with a density $p(t,x)$. This can be obtained by normalizing the solution of the Duncan–Martenssen–Zakai equation ([15], [24]). As was already mentioned, applications of the stability results can be obtained by adding the correction term to the filter equations when the coloured noises appear in the model ([22], [24]). It is likely that one can obtain similar results for the models described by the equations considered in this paper.

References

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

[2] L. Arnold, *Stochastic Differential Equations: Theory and Applications*, Wiley, New York 1974.

[3] S. S. Artem"ev, *Numerical solution of stochastic differential equations*, in: Proc. Third Conf. Differential Equations and Applications, Ruse 1985, VTU "A. K"nchev", 1985, 15–18 (in Russian).

[4] V. Bally, *Approximation for the solutions of stochastic differential equations, III. Jointly weak convergence*, Stochastics Stochastics Rep. 30 (1990), 171–191.

[5] H. T. Banks and F. Kappel, *Spline approximation for functional differential equations*, J. Differential Equations 34 (1979), 496–522.

[6] D. R. Bell and S. E. A. Mohammed, *On the solution of stochastic ordinary differential equations via small delays*, Stochastics Stochastics Rep. 28 (1989), 293–299.

[7] Z. Brze/xniak, M. Capi/nski and F. Flandoli, *A convergence result for stochastic partial differential equations*, Stochastics 24 (1988), 423–445.

[8] H.-F. Chen and A. J. Gao, *Robustness analysis for stochastic approximation algorithms*, Stochastics Stochastics Rep. 26 (1989), 3–20.

[9] A. Chojnowska-Michalik, *Representation theorem for general stochastic delay equations*, Bull. Acad. Polon. Sci. 26 (7) (1978), 635–642.

[10] —, *Stochastic differential equations in Hilbert spaces and their applications*, thesis, 1976.

[11] J. M. C. Clark, *An efficient approximation for a class of stochastic differential equations*, in: Advances in Filtering and Optimal Stochastic Control, W. Fleming and L. G. Gorostiza (eds.), Proceedings of the IFIP Working Conference, Cocoyoc, Mexico, 1982, Lecture Notes in Control and Inform. Sci. 42, Springer, Berlin 1982, 69–78.

[12] J. M. C. Clark and R. J. Cameron, *The maximum rate of convergence of discrete approximations for stochastic differential equations*, in: Stochastic Differential Systems–Filtering and Control, B. Grigelionis (ed.), Proceedings of the IFIP Working Conference, Vilnius, Lithuania, USSR, 1978, Lecture Notes in Control and Inform. Sci. 25, Springer, Berlin 1980, 162–171.

[13] R. F. Curtain and P. F. Falb, *Itô's lemma in infinite dimensions*, J. Math. Anal. Appl. 31 (1970), 431–448.

[14] R. F. Curtain and A. J. Pritchard, *Infinite Dimensional Linear Systems Theory*, Springer, Berlin 1978.

[15] G. Da Prato, *Stochastic differential equations with noncontinuous coefficients in Hilbert space*, Rend. Sem. Mat. Univ. Politec. Torino, Numero Speciale 1982, 73–85.

[16] G. Da Prato, S. Kwapień and J. Zabczyk, *Regularity of solutions of linear stochastic equations in Hilbert spaces*, Stochastics 23 (1987), 1–23.

[17] G. Da Prato and J. Zabczyk, *Stochastic Equations in Infinite Dimensions*, Cambridge Univ. Press, 1991.

[18] A. L. Dawidowicz and K. Twardowska, *On the Stratonovich and Itô integrals with integrands of delay argument*, submitted to Stochastica (1991).

[19] H. Doss, *Liens entre équations différentielles stochastiques et ordinaires*, Ann. Inst. H. Poincaré 13 (2) (1977), 99–125.

[20] G. Ferreyra, *A Wong–Zakai type theorem for certain discontinuous semimartingales*, J. Theoret. Probab. 2 (3) (1989), 313–323.

[21] A. Greiner and W. Strittmatter, *Numerical integration of stochastic differential equations*, J. Statist. Phys. 51 (1–2) (1988), 95–108.

[22] I. Gyöngy, *On the approximation of stochastic differential equations*, Stochastics 23 (1988), 331–352.

[23] —, *On the approximation of stochastic partial differential equations*, *Part I*, ibid. 25 (1988), 59–85, *Part II*, ibid. 26 (1989), 129–164.

[24] —, *The stability of stochastic partial differential equations and applications. Theorems on supports*, in: Lecture Notes in Math. 1390, Springer, Berlin 1989, 91–118.

[25] J. Hale, *Theory of Functional Differential Equations*, Springer, Berlin 1977.

[26] N. Ikeda, S. Nakao and Y. Yamato, *A class of approximations of Brownian motion*, Publ. RIMS Kyoto Univ. 13 (1977), 285–300.

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

[28] K. Itô and M. Nisio, *On stationary solutions of a stochastic differential equation*, J. Math. Kyoto Univ. 4 (1) (1964), 1–75.

[29] R. Janssen, *Difference methods for stochastic differential equations with discontinuous coefficient*, Stochastics 13 (1984), 199–212.

[30] —, *Discretization of the Wiener process in difference methods for stochastic differential equations*, Stochastic Process. Appl. 18 (1984), 361–369.

[31] T. Kato, *Perturbation Theory for Linear Operators*, Springer, Berlin 1966.

[32] S. Kawabata, *On the successive approximation of solutions of stochastic differential equations*, Stochastics Stochastics Rep. 30 (1990), 69–84.

[33] P. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, Springer, Berlin 1992.

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

[35] P. Kotelenez, *A submartingale type inequality with applications to stochastic evolution equations*, Stochastics 8 (1982), 139–151.

[36] H. H. Kuo, *Gaussian Measures in Banach Spaces*, Springer, Berlin 1975.

[37] H. Kushner, *Jump-diffusion approximations for ordinary differential equations with wide-band random right hand sides*, SIAM J. Control Optim. 17 (1979), 729–744.

[38] —, *Probability Methods for Approximations in Stochastic Control and for Elliptic Equations*, Academic Press, New York 1977.

[39] H. J. Kushner and G. Yin, *Stochastic approximation algorithms for parallel and distributed processing*, Stochastics 22 (1987), 219–250.

[40] R. Liptser and A. Shiryayev, *Statistics of Random Processes*, Springer, Berlin 1977.

[41] W. Mackevičius, *SP-stability of symmetric stochastic differential equations*, Litovsk. Mat. Sb. 25 (4) (1985), 72–84.

[42] X. Mao, *Approximate solutions for a class of delay stochastic differential equations*, Stochastics Stochastics Rep. 35 (1991), 111–123.

[43] S. Marcus, *Modeling and approximation of stochastic differential equations driven by semimartingales*, Stochastics 4 (1981), 223–245.

[44] E. J. McShane, *Stochastic differential equations and models of random processes*, Proc. 6th Berkeley Sympos. Math. Statist. Probab., Vol. 3, University of California Press, Berkeley 1972, 263–294.

[45] J. Memin and L. Słonimski, *Condition UT et stabilité en loi des solutions d'équations différentielles stochastiques*, to appear in Séminaire de Probabilités 1991.

[46] M. Métivier, *Semimartingales*, Walter de Gruyter, Berlin 1982.

[47] M. Métivier and J. Pellaumail, *Stochastic Integration*, Academic Press, London 1980.

[48] G. N. Milshteĭn, *Approximate integration of stochastic differential equations*, Theor. Veroyatnost. i Primenen. 19 (1974), 583–588 (in Russian).

[49] —, *Weak approximation of solutions of systems of stochastic differential equations*, ibid. 30 (1985), 706–721 (in Russian).

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

[51] —, *Retarded Functional Differential Equations*, Pitman, London 1978.

[52] S. Nakao and Y. Yamato, *Approximation theorem on stochastic differential equations*, in: Proc. Internat. Sympos. SDE Kyoto 1976, Tokyo 1978, 283–296.

[53] N. J. Newton, *An asymptotically efficient difference formula for solving stochastic differential equations*, Stochastics 19 (1986), 175–206.

[54] —, *An efficient approximation for stochastic differential equations on the partition of symmetrical first passage times*, ibid. 29 (1990), 227–258.

[55] E. Pardoux and D. Talay, *Discretization and simulation of stochastic differential equations*, Acta Appl. Math. 3 (1985), 23–47.

[56] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer, Berlin 1983.

[57] J. Picard, *Approximation of stochastic differential equations and application of the stochastic calculus of variations to the rate of convergence*, in: Lecture Notes in Math. 1316, Springer, 1988, 267–287.

[58] E. Platen, *An approximation method for a class of Itô process*, Lietuvos Matematikos Rinkinys 21 (1) (1981), 121–133.

[59] E. Platen, *Approximation of the first exit times of diffusions and approximate solutions of parabolic equations*, Math. Nachr. 111 (1983), 127–146.

[60] —, *A Taylor formula for semimartingales solving a stochastic equation*, in: Lecture Notes in Control and Inform. Sci. 36, Springer, 1981, 157–174.

[61] P. Protter, *Approximations of solutions of stochastic differential equations driven by semimartingales*, Ann. Probab. 13 (3) (1985), 716–743.

[62] A. Shimizu, *Approximate solutions of stochastic differential equations*, Bull. Nagoya Inst. Tech. 36 (1984), 105–108.

[63] A. V. Skorokhod, *K-martingales and Stochastic Equations*, in: Proc. School-Seminar of the Theory of Random Processes (Druskininkai 1974), Part II, Inst. Fiz. Mat. AN Litovsk. SSR, Vilnius 1975, 195–234 (in Russian).

[64] K. Sobczyk, *Stochastic Differential Equations with Applications to Physics and Engineering*, Kluwer, Dordrecht 1991.

[65] J. L. Solé and F. Utzet, *Stratonovich integrals and trace*, Stochastics 29 (1980), 203–220.

[66] R. L. Stratonovich, *A new representation for stochastic integrals and equations*, SIAM J. Control Optim. 4 (2) (1966), 362–371.

[67] D. W. Stroock and S. R. S. Varadhan, *On the support of diffusion processes with applications to the strong maximum principle*, in: Proc. 6th Berkeley Sympos. Math. Statist. Probab., Vol. 3, University of California Press, Berkeley 1972, 333–359.

[68] H. Sussmann, *On the gap between deterministic and stochastic ordinary differential equations*, Ann. Probab. 6 (1) (1978), 19–41.

[69] D. Talay, *Approximation of upper Lyapunov exponents of bilinear stochastic differential systems*, submitted to INRIA, Report 965 (1989).

[70] —, *Efficient numerical schemes for the approximation of expectations of functionals of S.D.E.*, in: Filtering and Control of Random Process, H. Korezlioglu, G. Mazziotto and J. Szpirglas (eds.), Proceedings of the ENST-CNET Colloquium, Paris 1983, Lecture Notes in Control and Inform. Sci. 61, Springer, Berlin 1984, 294–313.

[71] —, *Résolution trajectorielle et analyse numérique des équations différentielles stochastiques*, Stochastics 9 (1983), 275–306.

[72] —, *Second-order discretization schemes of stochastic differential systems for the computation of the invariant law*, Stochastics Stochastics Rep. 29 (1990), 13–36.

[73] H. F. Trotter, *Approximation of semi-groups of operators*, Pacific J. Math. 8 (1958), 887–919.

[74] C. Tudor, *On stochastic evolution equations driven by continuous semimartingales*, Stochastics 24 (1988), 179–195.

[75] —, *On the successive approximations of solutions of delay stochastic evolution equations*, An. Univ. Bucureşti Mat. 34 (1985), 70–86.

[76] —, *On weak solutions of Volterra equations*, Boll. Un. Mat. Ital. 7 1-B (1987), 1033–1054.

[77] —, *Some properties of mild solutions of delay stochastic evolution equations*, Stochastics 17 (1986), 1–18.

[78] K. Twardowska, *An extension of Wong–Zakai theorem for stochastic evolution equations in Hilbert spaces*, Stochastic Anal. Appl. 10 (4) (1992), 471–500.

[79] —, *On the approximation theorem of the Wong–Zakai type for the functional stochastic differential equations*, Probab. Math. Statist. 12 (2) (1991), 319–334.

[80] —, *On the relation between the Itô and Stratonovich integrals in Hilbert spaces*, submitted to Statist. Probab. Lett. (1991).

[81] R. B. Vinter, *A Representation of Solutions to Stochastic Delay Equations*, Imperial College Report, 1975.

[82] —, *On the evolution of the state of linear differential delay equations in* M^2: *properties of the generator*, J. Inst. Math. Appl. 21 (1) (1978), 13–23.

[83] N. N. Vakhaniya, V. I. Tarieladze and S. A. Chobanyan, *Probability Distributions in Banach Space*, Nauka, Moscow 1985 (in Russian).

[84] E. Wagner, *Unbiased Monte-Carlo estimators for functionals of weak solutions of stochastic differential equations*, Stochastics Stochastics Rep. 28 (1989), 1–20.

[85] E. Wong, *Stochastic Processes in Information and Dynamical Systems*, McGraw-Hill, 1971.

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

[87] —, —, *Riemann–Stieltjes approximations of stochastic integrals*, Z. Warsch. Verw. Gebiete 12 (1969), 87–97.

[88] J. Zabczyk, *On decomposition of generators*, SIAM J. Control Optim. 16 (4) (1978), 523–534.