Acta Applicandae Mathematicae **43**: 317–359, 1996.

Wong–Zakai Approximations for Stochastic Differential Equations

KRYSTYNA TWARDOWSKA*
Institute of Mathematics, Warsaw University of Technology, Pl. Politechniki 1, 00–661 Warsaw, Poland. e-mail: tward@alpha.im.pw.edu.pl

(Received: 14 December 1994)

Abstract. The aim of this paper is to give a wide introduction to approximation concepts in the theory of stochastic differential equations. The paper is principally concerned with Wong–Zakai approximations. Our aim is to fill a gap in the literature caused by the complete lack of monographs on such approximation methods for stochastic differential equations; this will be the objective of the author's forthcoming book. First, we briefly review the currently-known approximation results for finite- and infinite-dimensional equations. Then the author's results are preceded by the introduction of two new forms of correction terms in infinite dimensions appearing in the Wong–Zakai approximations. Finally, these results are divided into four parts: for stochastic delay equations, for semilinear and nonlinear stochastic equations in abstract spaces, and for the Navier–Stokes equations. We emphasize in this paper results rather than proofs. Some applications are indicated.

Mathematics Subject Classifications (1991): 60H15, 60H10, 60H05, 34K50, 41A10, 65C20, 35Q30, 35R60.

Key words: stochastc evolution equation, stochastic delay equation, Wong–Zakai approximation, approximation method.

1. Introduction

We survey a part of the approximation theory of stochastic differential equations, namely Wong–Zakai approximation theorems. We also wish to familiarize the reader with other approximation problems for stochastic differential equations. We try to emphasize the global treatment of these problems within the framework of general approximation theory (see Section 4).

The problem of approximation of a Wiener process by its piecewise linear counterpart has arisen from Wong and Zakai's work [123], which proves to be of great significance to many subjects concerning the limit behaviour of stochastic differential equations. The main characteristic property of these theorems is the so-called correction term in the limit equation.

The paper is organized as follows. Later in this section we give a review of approximations of different types other than the Wong–Zakai approximations for stochastic differential equations which are generalizations of the deterministic

* The author's research was partially supported by KBN grant No. 2 P301 052 03.

and numerical approaches to approximations of differential equations. We also indicate the so-called UT approximations. For methodological reasons, we give a table of systematic treatment of the results in the Wong–Zakai approximations.

Further, we present only the Wong–Zakai approximations. There are several results described in Section 2 for the finite-dimensional case. Section 3 discusses these problems for linear stochastic differential equations in infinite dimensions.

The problem of constructing two new forms of the correction term is discussed in Section 4. They were introduced by the author in [112]–[116]. This was necessary to enable us to formulate and prove the Wong–Zakai type approximation theorems for stochastic delay equations (the first correction term) as well as for semilinear equations, nonlinear stochastic evolution equations and Navier–Stokes equations (the second correction term); in the latter three cases, the disturbances are Hilbert space-valued Wiener processes (Section 5). The idea that the author followed in these constructions is presented in detail and a discussion and comparison between these correction terms and with their finite-dimensional form from previous sections is also included.

We have to stress that three quite different methods are used to prove the Wong–Zakai type theorems of Section 5: one for stochastic delay equations, another for semilinear equations, and still another for both nonlinear and Navier–Stokes equations.

We also discuss (Section 6) some applications of approximation theorems to other mathematical aspects connected with the theory of stochastic differential equations as well as to engineering and physical sciences; the Wong–Zakai approximation methods are of wide practical significance (when the white noise is approximated by the coloured noise).

Didactic aims were the guiding principle during writing this paper, so as to present the approximation theory of stochastic differential equations not only to researchers of a wide spectrum in the field of differential equations but also for postgraduate students. Therefore, we grade difficulties, emphasize the methodology, and give a systematic table of results. We refer to the bibliography for proofs. The stages of generalization of results are strictly isolated.

What is new in the approximation theory of deterministic differential equations versus approximation of stochastic differential equations as well as what is new in the Wong–Zakai approximations of stochastic differential equations versus other aspects of the theory of these equations? Why are approximation theorems such a difficult part of the theory of stochastic differential equations? The author would like to answer these and other questions in a forthcoming book.

As already mentioned, a 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 such term. These are, e.g., numerical schemes.

Consider the Itô stochastic equation of the form

$$x(t) = x(0) + \int_0^t m(x(s), s)\,\mathrm{d}s + \int_0^t \sigma(x(s), s)\,\mathrm{d}w(s), \quad t \geqslant 0, \tag{1.1}$$

where the initial value $x(0)$ may be a random variable. The second integral in (1.1) is the Itô stochastic integral with respect to the Wiener process $w(t)$, $t \geqslant 0$. This equation may be written in the differential form

$$\mathrm{d}x(t) = m(x(t), t)\,\mathrm{d}t + \sigma(x(t), t)\,\mathrm{d}w(t), \quad x(0) = x_0. \tag{1.2}$$

Effective analytical solution of stochastic differential equations is only possible in some simple cases. So there exists an obvious interest to develop numerical methods for such equations. There are now a number of papers which deal with these methods to relate solutions to approximate exact solutions.

A thorough description of numerical problems in this area can be found in the books of Kloeden and Platen [48] and Sobczyk [97].

The approximation methods have followed five directions (see Sobczyk [97] and Talay [102]): mean-square approximation (Clark and Cameron [21], Milshtein [75] and Platen [88]), pathwise approximation (Talay [103]), approximation of expectations of the solutions (Milshtein [75], Talay [102]), numerical computation of the Lyapunov exponents (Talay [101]) and asymptotically efficient schemes for minimization of the normalized quadratic mean error (Clark [20] and Newton [79]).

Now we survey various time discrete numerical methods which simulate the sample paths and functionals of the Itô processes that are solutions to (1.1).

The first method is the one of successive approximations (see Kawabata [46] and Tudor [108]). It has been used before for proving the existence of the solutions and can be used for numerical calculations. The recursive formula is

$$\begin{aligned} x^{(n)}(t) &= x(0) + \int_0^t m(x^{(n-1)}(s), s)\,\mathrm{d}s + \\ &\quad + \int_0^t \sigma(x^{(n-1)}(s), s)\,\mathrm{d}w(s), \end{aligned} \tag{1.3^n}$$

where $n = 1, 2, \ldots$ We calculate the first (Riemann–Stieltjes) integral and the second (Itô) integral by known discretization algorithms. The methods of successive approximations are, however, not popular in practice since we have to keep at the disposal all the values of the previous approximations for all $s \in [t_0, t]$.

Therefore one-step difference methods are of great interest.

Let us stop for a moment and consider the types of convergence of approximate schemes (compare Kloeden and Platen [48]).

When we wish to approximate a solution to Equation (1.2) by a time-discretized recurrent formula

$$x(t_{i+1}) = f(x(t_i), w(s);\ t_i \leqslant s \leqslant t_{i+1}), \tag{1.4}$$

then the 'goodness' of the scheme depends on the type of the chosen approximation and the required type of convergence.

We consider a time discretization of the interval $[0,T]$,

$$\begin{aligned} &\tau_n = \{t_i\colon\ 0,1,\dots,n;\ 0 = t_0 < t_1 < \cdots < t_n = T\},\\ &\Delta t_i = t_{i+1} - t_i = h, \quad \Delta w_i = w(t_{i+1}) - w(t_i),\\ &x^{(n)}(t) = \bar{x}(t), \quad \bar{x}(t_i) = \bar{x}_i \end{aligned} \tag{1.5}$$

and the integral counterpart of Equation (1.2) on the interval $[t_i, t_{i+1}]$:

$$x(t_{i+1}) = x(t_i) + \int_{t_i}^{t_{i+1}} m(x(s),s)\,\mathrm{d}s + \int_{t_i}^{t_{i+1}} \sigma(x(s),s)\,\mathrm{d}w(s). \tag{1.6}$$

We are looking for a scheme such that for all $t \in [0,T]$,

$$E[x(t) - \bar{x}(t)]^2 \longrightarrow 0, \quad \text{as } h \to \infty.$$

To go further we have to settle that the natural way of classifying numerical methods for stochastic differential equations is to compare them with strong and weak Taylor approximations. The stochastic Taylor formula allows a function of an Itô process, i.e. $f(x(t))$, to be expanded about $f(x(t_0))$ in terms of stochastic integrals weighted by coefficients evaluated at $x(t_0)$.

The simplest strong Taylor approximation is the Euler approximation

$$\bar{x}_{i+1} = \bar{x}_i + m(\bar{x}_i, t_i)h + \sigma(\bar{x}_i, t_i)\Delta w_i \quad \text{for } s \in [t_i, t_{i+1}]. \tag{1.7}$$

It was shown by Maruyama [68] that the Euler scheme converges uniformly in the mean-square sense to the $x(t)$ determined by (1.2) as $h \to 0$. However, the order of approximation of this scheme is too low.

Therefore we include the next (second) term from the stochastic Taylor formula to the scheme (1.7). We obtain the Milshtein scheme

$$\begin{aligned} \bar{x}_{i+1} &= \bar{x}_i + [m(\bar{x}_i, t_i) - \tfrac{1}{2}\sigma(\bar{x}_i, t_i)\sigma'(\bar{x}_i, t_i)]h +\\ &\quad + \sigma(x_i, t_i)\Delta w_i + \tfrac{1}{2}\sigma(x_i, t_i)\sigma'(x_i, t_i)(\Delta w_i)^2,\\ \bar{x}_0 &= x(0) \end{aligned} \tag{1.8}$$

for $s \in [t_i, t_{i+1}]$.

The Milshtein scheme can also be obtained as the Euler scheme for the Stratonovich version of Equation (1.2) (see Pardoux and Talay [84]). This is one of the practical applications of the correction term appearing in the Wong–Zakai approximation theorems when transition is made from the Itô to Stratonovich integral.

The rate of convergence is optimal within a large class of approximations because the last term in (1.8) contains additional information about the sample paths of the Wiener process (see Clark and Cameron [21]).

A practical disadvantage of the above Taylor approximations is that the derivatives of various orders of the drift and diffusion coefficients must be evaluated. There are time discrete approximations which avoid the use of derivatives, that is, analogs of Runge–Kutta schemes for ordinary differential equations.

A stochastic counterpart of the two-stage Runge–Kutta method (called the Heun method (see Blum [13])) has the form

$$\begin{aligned}\bar{x}_{i+1} &= \bar{x}_i + m(\bar{x}_i, t_i) + \tfrac{1}{2}[m(\bar{x}_i, t_i) + m(\widehat{x}_{i+1}, t_{i+1})]h + \\ &\quad + \tfrac{1}{2}[\sigma(\bar{x}_i, t_i) + \sigma(\widehat{x}_{i+1}, t_{i+1})]\, \mathrm{d}w_i, \end{aligned} \tag{1.9}$$

where $\widehat{x} = \bar{x}_i + m(\bar{x}_i, t_i)h + \sigma(\bar{x}_i, t_i)\, \mathrm{d}w_i$.

McShane [70] has shown that this scheme converges in the mean-square sense to the solution of the Stratonovich counterpart of Equation (1.2). Other results of the stochastic Runge–Kutta method can be found in the paper of Klauder and Peterson [47].

All the above procedures have given approximations of the solution in the mean-square sense. The numerical schemes which give the pathwise convergence, i.e. with probability one, are discussed by Pardoux and Talay in [84].

In numerical practice, additional approximation is also necessary. Namely, we have to approximate the Wiener process by its suitable simulations, e.g., by polygonal approximations. For problems associated with such approximations, see papers dealing with the Wong–Zakai approximation theorems.

Another type of possible approximation of the Wiener process was investigated, e.g., in the paper of Gorostiza [30], where the Wiener process approximation called the transport process was considered.

As for other approximation methods, the objective of Janssen's paper [43] was to extend the Cauchy–Maruyama approximation method to delay stochastic differential equations based on seminartingales with spatial parameters. This procedure is also applicable to nondelay equations.

A very important contribution to approximation methods was made by Kushner: in [55] for ordinary differential equations with wide-band random right-hand sides; and in [56] and in [57] together with Dupuis for problems arising in stochastic control. In [58], 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.

Pettersson [86] and Słomiński [95] considered the convergence of a recursive projection scheme for a stochastic differential equation reflecting on the boundary of a convex domain G. Dependening on the shape of the domain G, we obtain mean-square or pointwise convergence of different rates. This scheme is essentially the Euler method forced to remain in the constraining set G.

The paper of M. Tudor [111] gives several approximation schemes for two-parameter Itô equations, mainly by separation of the diffusion and drift terms.

The convergence of the algorithms and the rate of the convergence are considered under suitable conditions.

The papers of Jakubowski, Mémin and Pagès [42], Kurtz and Protter [52]–[54], Mémin and Słomiński [71] give the UT (uniform tightness) condition for a sequence $\{z^n\}_{n\in\mathbb{N}}$ of $\mathcal{F}^n$-adapted semimartingales and introduce an approximation of the noise in the stochastic differential equation. A stability result is proved. The theorems are 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 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 above results can be applied. The UT condition is sometimes difficult to verify in practice. An alternative condition was given by Kurtz and Protter in [53]. Applications to stochastic differential equations, that is, approximation theorems, are also considered in these papers.

To be more precise (see Kurtz and Protter [52]), for $n = 1, 2, \ldots$, let $\Xi_n = (\Omega, \mathcal{F}^n, (\mathcal{F}^n_t)_{t\geqslant 0}, P^n)$ be a filtered probability space, let H^n càdlàg and adapted, and let X^n be a càdlàg semimartingale. A fundamental question is: Under what conditions does the convergence in the distribution of (H^n, X^n) to (H, X) imply that X is a semimartingale and that $\int_0^t H^n_{s-}\,\mathrm{d}X^n_s$ converges in distribution to $\int_0^t H_{s-}\,\mathrm{d}X_s$? A slightly more general formulation would put conditions on the sequence X^n only, such that the convergence above holds for all such sequences H^n. A sequence with this property will be called *good*. Let $\mathbb{M}^{km}$ denote the real-valued $k \times m$ matrices, and let $\mathbb{D}_{\mathbb{E}}[0, \infty)$ denote the space of càdlàg, E-valued functions with Skorokhod topology.

DEFINITION 1.1. For $n = 1, 2, \ldots$, let X^n be an $\mathbb{R}^k$-valued $(\mathcal{F}^n_t)$-semimartingale, and let the sequence $(X^n)_{n\geqslant 1}$ converge in distribution in the Skorokhod topology to a process X. The sequence $(X^n)_{n\geqslant 1}$ is said to be good if for any sequence $(H^n)_{n\geqslant 1}$ of $\mathbb{M}^{km}$-valued, càdlàg processes, H^n $(\mathcal{F}^n_t)$-adapted, such that (H^n, X^n) converges in distribution in the Skorokhod topology on $\mathbb{D}_{\mathbb{M}^{km}\times\mathbb{R}^m}[0, \infty)$ to a process (H, X), there exists a filtration $(\mathcal{F}_t)$ such that H is $(\mathcal{F}_t)$-adapted, X is an $(\mathcal{F}_t)$-semimartingale, and

$$\int_0^t H^n_{s-}\,\mathrm{d}X^n_s \Longrightarrow \int_0^t H_{s-}\,\mathrm{d}X_s.$$

On Ξ_n, let $\mathcal{H}^n$ denote the set of elementary predictable processes bounded by 1, that is,

$$\mathcal{H}^n = \Big\{H^n\colon\ H^n \text{ has the representation } H^n_t = H^n_0 1_{\{0\}}(t) + \sum_{i=1}^{p-1} H^n_i 1_{[t_i, t_{i+1})}(t),$$
$$\text{with } H^n_i \in \mathcal{F}^n_{t_i},\ p \in \mathbb{N},\ \text{and } 0 = t_0 < t_1 < \cdots < t_p < \infty,\ |H^n_i| \leqslant 1\Big\}$$

DEFINITION 1.2. A sequence of semimartingales $(X^n)_{n\geqslant 1}$, X^n defined on Ξ_n, satisfies the condition UT if for each $t > 0$ the set $\{\int_0^t H_s^n \,\mathrm{d}X_s^n,\ H^n \in \mathcal{H}^n, n \in \mathbb{N}\}$ is stochastically bounded.

THEOREM 1.1 (Jakubowski, Mémin and Pagès [42]). *If (H^n, X^n) on Ξ_n converges in distribution to (H, X) in the Skorokhod topology and if $(X^n)_{n\geqslant 1}$ satisfies condition* UT, *then there exists a filtration $(\mathcal{F}_t)$ such that X is an $(\mathcal{F}_t)$-semimartingale and $\int H_{s-}^n \,\mathrm{d}X_s^n$ converges in distribution in the Skorokhod topology to $\int H_{s-} \,\mathrm{d}X_s$. That is, the sequence $(X^n)_{n\geqslant 1}$ is good.*

The above presentation does not claim to be a complete and up-to-date work on the subject of approximation theorems of stochastic differential equations. Our aim is to present, in the introductory manner only, all the main steps of generalization of the original Wong–Zakai theorem. We may arrange these results in the following table and discuss them later in the paper. We denote the types of disturbances here in the following way:

TABLE I.

	Finite 1-dim. state & noise	Finite multi-dim. state & noise	Infinite state & finite noise	Infinite state & noise
Linear drift and diffusion	—	—	[2], [14] W	—
Linear drift and nonlinear diffusion	—	—	[33]–[35] CS	—
Nonlinear drift and diffusion	[122], [123] W	[26] CS [28] DS [32] CS [39], [40] W [49] CS [51] DS [64], [70] W [78] CQ [87] CM [91] DS [99] W	[116]	[112]–[115] W

W — Wiener process, CS — continuous semimartingale,
DS — discontinuous semimartingale, CM — continuous martingale,
CQ — continuous quasimartingale.

2. Wong–Zakai Approximations in the Finite-Dimensional Case

The theorem on the convergence of ordinary integrals to stochastic integrals was first proved by Wong and Zakai [123, 124] for a one-dimensional state space and one-dimensional Wiener process. The solution $x(t)$, $a \leqslant t \leqslant b$, to the stochastic differential equation

$$\mathrm{d}x(t) = m(x(t), t)\,\mathrm{d}t + \sigma(x(t), t)\,\mathrm{d}w(t), \quad x(a) = x_a, \tag{2.1}$$

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 $x(t)$ (compare Wong and Zakai [123], Arnold [3], Liptser and Shiryayev [60]). Let $x_n(t)$ be the solution of the ordinary differential equation

$$\mathrm{d}x_n(t) = m(x_n(t), t)\,\mathrm{d}t + \sigma(x_n(t), t)\,\mathrm{d}w_n(t), \quad x_n(a) = x_a \tag{2.2$_n$}$$

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 (2.1), but it satisfies

$$\begin{aligned}\mathrm{d}y(t) &= m(y(t), t)\,\mathrm{d}t + \tfrac{1}{2}\sigma(y(t), t)\frac{\partial \sigma(y(t), t)}{\partial y}\,\mathrm{d}t + \\ &\quad + \sigma(y(t), t)\,\mathrm{d}w(t), \quad y(a) = x_a.\end{aligned} \tag{2.3}$$

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 $\mathrm{d}w(t) \approx \sqrt{\mathrm{d}t}$ (see Arnold [3], Wong [122], Wong and Zakai [123]).

More precisely, we have

THEOREM 2.1 (Wong and Zakai [123]). *Suppose that* $(\Omega, \mathcal{F}, P)$ *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* $-\infty < x < \infty$, $a \leqslant t \leqslant b$.
(ii) $m(x,t)$, $\sigma(x,t)$, $\sigma(x,t)(\partial\sigma(x,t)/\partial x)$ *satisfy the Lipschitz condition with a constant* $k > 0$.
(iii) $\sigma(x,t) \geqslant \beta > 0$ *(or* $-\sigma(x,t) \geqslant \beta > 0$*) and* $|\partial\sigma(x,t)/\partial t| \leqslant k\sigma^2(x,t)$.
(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* $[a,b]$, $w_n(t,\omega) \leqslant k(\omega)$.
(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* (2.2$_n$) *and* (2.3), *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 Ikeda *et al.* [39], Ikeda and Watanabe [40], McShane [70], Strook and Varadhan [99]). If we consider the functions m: $[a, b] \times \mathbb{R}^d \to \mathbb{R}^d$ and σ: $[a, b] \times \mathbb{R}^d \to \mathbb{R}^{d\times m}$, then the correction term has the form

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

Recall one of the main results in this area:

THEOREM 2.2 (Ikeda and Watanabe [40]). *Let $X \in \mathbb{R}^d$, $m \in C_b^1(\mathbb{R}^d, \mathbb{R}^d)$, $\sigma \in C_b^2(\mathbb{R}^d, \mathbb{R}^{d\times m})$ (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_0^T, P) and the following equations are satisfied:*

$$\begin{aligned} x_n^i(t,w) &= X^i(w) + \int_0^t m^i(x_n(s,w))\,\mathrm{d}s + \\ &\quad + \sum_{p=1}^{m}\int_0^t \sigma^{ip}(x_n(s,w))\dot{B}_n^p(s,w)\,\mathrm{d}s, \\ y^i(t,w) &= X^i(w) + \int_0^t m^i(y(s,w))\,\mathrm{d}s + \sum_{p=1}^{m}\int_0^t \sigma^{ip}(y(s,w))\,\mathrm{d}w(s) + \\ &\quad + \frac{1}{2}\sum_{p=1}^{m}\sum_{j=1}^{d}\int_0^t \frac{\partial\sigma^{ip}(y(t),t)}{\partial y^j}\sigma^{jp}(y(t),t)\,\mathrm{d}t \end{aligned}$$

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

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

Further generalizations deal with problems with more general noises than the Wiener process. The result due to Protter [91] for continuous semimartingale differentials can be stated in simplified form as

THEOREM 2.3. *Consider the equations*

$$\begin{aligned} &\mathrm{d}x_n(t) = f(t, x_n(t), Z_n(t))\,\mathrm{d}Z_n(t), \\ &\mathrm{d}x(t) = f(t, x(t), Z(t)) \circ \mathrm{d}Z(t), \end{aligned} \tag{2.4}$$

$$\begin{aligned} \mathrm{d}y(t) &= f(t, y(t), Z(t))\,\mathrm{d}Z(t) + \\ &\quad + \tfrac{1}{2}\{f(\partial f/\partial y) + (\partial f/\partial Z)\}(t, y(t), Z(t))\,\mathrm{d}[Z^c, Z^c](t), \end{aligned} \tag{2.5}$$

where Z_n are piecewise C^1 approximations of a continuous semimartingale Z, the circle '$\circ$' denotes the Stratonovich integral and Z^c is the continuous part of Z. Under suitable assumptions, if Z_n tends to Z, then x_n tends to x satisfying (2.4) (*and hence satisfying* (2.5) *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 [67] and Kushner [55]. Results of this type were also examined by Bally [5], Ferreyra [28], Gyöngy [32], Mackevičius [63] and Picard [87]. In the second direction, pioneering work was done in [124] by Wong and Zakai, and more recently by Doss [26], Koneczny [49] and also by Nakao and Yamato [78].

In the paper [51] of Kurtz, Pardoux and Protter, stochastic differential equations driven by semimartingales with jumps are examined. The existence and uniqueness of solutions is established and the result on the Wong–Zakai type weak convergence is proved when the approximating differentials are smooth, in particular continuous, even though the limits are discontinuous.

Mackevičius and Žibaitis in [65] considered both polygonal and mollifier approximations of the Wiener process. They give the limit theorems for such cases. More exactly, given a Brownian motion B, Gaussian approximations B^δ, $\delta > 0$, of the form

$$B_t^\delta = \int_0^t \int_{\mathbb{R}} K^\delta(u,s)\,\mathrm{d}B_s\,\mathrm{d}u, \quad t \geqslant 0,$$

including polygonal and mollifier approximations, are considered. A limit theorem is proved for the integrals $\int_0^T X_t\,\mathrm{d}B_t^\delta$ as $\delta \to 0$. In particular, in the case of symmetric kernels K^δ, the limit is the Fisk–Stratonovich integral $\int_0^T X_t \circ \mathrm{d}B_t$.

3. Wong–Zakai Approximations in Infinite Dimensions for the Linear Case

In the infinite-dimensional case, some generalizations are known where the Wiener process is one-dimensional and the state space is infinite-dimensional (Aquistapace and Terreni [2], Brzeźniak, Capiński and Flandoli [14], Da Prato [23], Doss [26], Gyöngy [33–35]).

In [2], the following result is stated:

THEOREM 3.1 (Aquistapace and Terreni [2]). *Let $(\Omega, \mathcal{F}, P)$ be a probability space. Consider the stochastic problem*

$$\mathrm{d}u(t) = (A(t)u(t) + \tfrac{1}{2}B^2u(t))\,\mathrm{d}t + Bu(t)\,\mathrm{d}w(t) + f(t)\,\mathrm{d}t,$$

$$u(0) = u_0, \tag{3.1}$$

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 generates 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 (3.1) and it is the limit of the solutions of the approximating deterministic problems

$$\begin{aligned} \mathrm{d}u_n(t) &= A(t)u_n(t)\,\mathrm{d}t + Bu_n(t)\dot{w}_n(t)\,\mathrm{d}t + f_n(t)\,\mathrm{d}t, \\ u_n(0) &= u_0, \end{aligned} \tag{3.2$_n$}$$

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

We observe that the correction term is here of the form $\frac{1}{2}B^2u(t)$ and we see that it is the same form of correction term as in the linear case in infinite dimensions. Some slight modifications of the above theorem are given by Brzeźniak, Capiński and Flandoli in [14] and by Da Prato in [23].

In the papers of Gyöngy [33–35], 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 equations considered admit many meaningful physical applications.

Gyöngy in his papers [33–35] generalizes approximation results to the following stochastic evolution equation

$$\begin{aligned} u(t) &= u_0 + \int_0^t (A(s)u(s) + F(s))\,\mathrm{d}H(s) + \\ &\quad + \int_0^t (B_i(s)u(s) + G(s)\,\mathrm{d}M^i(s), \end{aligned} \tag{3.3}$$

where A, B_i are unbounded linear operators (second and first-order differential operators), $M = (M^1, \ldots, M^d)$ is a continuous semimartingale. His assumptions make it possible to cover many stochastic partial differential equations appearing in applications.

4. Correction Terms in the Wong–Zakai Approximations

4.1. CONSTRUCTION OF CORRECTION TERMS IN INFINITE DIMENSIONS

We begin with an example and some results from the author's papers [115] and [116].

EXAMPLE 4.1.1. Using the step-by-step method of solving delay equations, we consider the following stochastic delay differential equation:

$$\begin{aligned} \mathrm{d}X(t) &= X(t-1)\,\mathrm{d}w(t) \quad \text{for } t \geqslant 0, \\ X(t) &= 1 \quad \text{for } t \in [-1,0], \end{aligned}$$

where $w(t)$ is the one-dimensional Wiener process. Further,

$$\begin{aligned} \mathrm{d}X^n(t) &= X^n(t-1)\dot{w}_n(t)\,\mathrm{d}t \quad \text{for } t \geqslant 0, \\ X^n(t) &= 1 \quad \text{for } t \in [-1,0]. \end{aligned}$$

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

$$\mathrm{d}X(t) = \mathrm{d}w(t), \quad X(0) = 1$$

and

$$\mathrm{d}X^n(t) = \dot{w}_n(t)\,\mathrm{d}t, \quad X^n(0) = 1.$$

Integrating, we get

$$X(t) = X(0) + \int_0^t \mathrm{d}w(s) = 1 + w(t)$$

and

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

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

$$\mathrm{d}X(t) = (1 + w(t-1))\,\mathrm{d}w(t), \quad X(1) = 1 + w(1)$$

and

$$\mathrm{d}X^n(t) = (1 + w_n(t-1))\,\mathrm{d}w_n(t), \quad X^n(1) = 1 + w_n(1).$$

We obtain

$$X(1) = 1 + w(1) + \int_0^t w(s-1)\,\mathrm{d}w(s)$$

and

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

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

We now consider (see also Section 5.1 below) the following more general stochastic differential equation with delayed argument on the space $\mathcal{C}_- = C(J, \mathbb{R}^d)$, $J = [-r, 0]$, $+\infty \geqslant r > 0$:

$$X^i(t,w) = X_0^i(w) + \int_0^t b^i(X_s(\cdot,w))\,\mathrm{d}s + + \sum_{p=1}^m \int_0^t \sigma^{ip}(X_s(\cdot,w))\,\mathrm{d}w^p(s) \tag{4.1.1}$$

for $i = 1, \dots, d$, where $X_t(\cdot)$ is a segment of the trajectory of X on $[-r, 0]$.

By replacing the Wiener process by its piecewise linear approximations B^n, we obtain the following approximations of (4.1.1):

$$X^{n,i}(t,w) = X_0^{n,i}(w) + \int_0^t b^i(X_s{}^n(\cdot,w))\,\mathrm{d}s + + \sum_{p=1}^m \int_0^t \sigma^{ip}(X_s^n(\cdot,w))\dot{B}^{n,p}(s,w)\,\mathrm{d}s. \tag{4.1.2^n}$$

We also introduce another stochastic differential equation:

$$Y^i(t,w) = Y_0^i(w) + \int_0^t b^i(Y_s(\cdot,w))\,\mathrm{d}s + \sum_{p=1}^m \int_0^t \sigma^{ip}(Y_s(\cdot,w))\,\mathrm{d}w^p(s) + + \frac{1}{2}\sum_{p=1}^m \sum_{j=1}^d \int_0^t \widetilde{D}_j\sigma^{ip}(Y_s(\cdot,w))\sigma^{jp}(Y_s(\cdot,w))\,\mathrm{d}s \tag{4.1.3}$$

for every $i = 1, \dots, d$. Further, $D\sigma^{ip}$ is the Fréchet derivative from $\mathcal{C}_-$ to $L(\mathcal{C}_-, R)$ (the necessary assumptions are given in Section 5.1) while

$$\widetilde{D}_j\sigma^{ip}(Y_s(\cdot,w)) = \mu_{s,w,Y}^{ipj}(\{0\}) \tag{4.1.4}$$

is the jth coordinate of a measure $\mu = \mu_{s,w,Y}^{ip}$ on $\mathcal{C}_-$ such that

$$\mu(\Phi) = \sum_{j=1}^d \int_{-r}^0 \Phi(v_j)\,\mu^j(\mathrm{d}v).$$

We have

$$\mu(A) = \mu(A \cap (-r,0)) + \mu(A \cap \{0\}) = \widetilde{\mu}(A) + \mu(\{0\})\delta_0(A), \tag{4.1.5}$$

where δ_0 is the Dirac measure, $A \in \mathcal{B}((-r,0))$. It is obvious that

$$D\sigma^{ip}(g)(\Phi) = \sum_{j=1}^d \int_{-r}^0 \Phi_j(v)\mu_{s,w,g}^{ipj}(\mathrm{d}v) \tag{4.1.6}$$

is a directional derivative.

Let us return to our first example. How do we relate the form of the correction term in (4.1.3) to the fact that we have no correction term in Example 4.1.1 for the delay constant in time?

We can conclude in the following way. If we compute the integrals $\int_0^t \sigma(X_s(\cdot,w))\,\mathrm{d}w(s)$ in Example 4.1.1 then there is no correction term if $\sigma(X_t(\theta)) = \sigma(X_t(-r,-\varepsilon))$, that is, if $\sigma(X_t(-r,-\varepsilon))$ does not depend on the value of σ at zero while the Wiener process depends on this value in zero. So if we take $\sigma(X_t(\theta)) = X(t-1)$, then σ depends on $t \in (-r,-\varepsilon)$ but we integrate on the interval $[0,\varepsilon]$. So σ and w are independent and there is no correction term. This suggests that we introduce formula (4.1.5) for the measure μ and to conclude that the correction term for stochastic delay equations will only depend on $\mu(\{0\})$ and will have the form

$$\frac{1}{2}\sum_{p=1}^{m}\sum_{j=1}^{d}\int_0^t \widetilde{D}_j\sigma^{ip}(Y_s(\cdot,w))\sigma^{jp}(Y_s(\cdot,w))\,\mathrm{d}s. \tag{4.1.7}$$

The correctness of our definition of the correction term can be seen from the following examples. Consider once more the equation

$$\mathrm{d}X(t) = \sigma(X_t)\,\mathrm{d}w(t), \quad X_0(\theta,w) = \eta(w)$$

for $\theta \in J$, where $\sigma\colon C_- \to R$, $\sigma(\varphi) = \varphi(-1)$, that is, $\varphi(-1) = X_t(-1) = X(t-1)$. Then

$$\mathrm{d}X(t) = X(t-1)\,\mathrm{d}w(t), \quad X_0 = \eta,$$

and Equation (4.1.3) is of the form

$$\mathrm{d}Y(t) = Y(t-1)\,\mathrm{d}w(t), \quad Y_0 = \eta,$$

because the measure μ is concentrated on the set $\{-1\}$ only and, hence, $\mu(\{0\}) = 0$. The latter can be seen after computing the directional derivative (4.1.6), that is, $D_\varphi\sigma(u) = \varphi(-1)$. Therefore, there is no difference between the initial and limit equations.

Consider the second new correction term for the case of stochastic evolution equations with nonlinear diffusion term with the state in a Hilbert space and with a Hilbert space-valued Wiener process. Again, we start with an example.

EXAMPLE 4.1.2. First we introduce the space M^2. For fixed $r \in \mathbb{R}_+$ we put $I = [-r,0]$, $0 < r < \infty$, and $M^2 = \mathbb{R}^n \times L^2(I,\mathbb{R}^n)$. The elements of M^2 are defined by

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

In M^2 the natural inner product is introduced

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

Consider the following stochastic delay equation of semilinear type:

$$\begin{aligned} \mathrm{d}x(t) &= \left(\sum_{i=0}^{m} A_i x(t - r_i) + \int_{-r}^{0} A(\theta) x(t+\theta)\,\mathrm{d}\theta\right)\mathrm{d}t + \Sigma(x_t)\,\mathrm{d}w(t), \\ x_0(\theta) &= \psi(\theta), \end{aligned} \tag{4.1.8}$$

where $t \in [0, T]$ and $r_i \in \mathbb{R}_+$ are fixed, $0 = r_0 < \cdots < r_m = r$, $(x(t))_{t\in[0,T]}$ is an $\mathbb{R}^n$-valued stochastic process and $(w(t))_{t\geqslant 0}$ 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, and Σ: $L^2(I, \mathbb{R}^n) \to L(H, \mathbb{R}^n)$ is an operator.

We denote by $\tilde{A}$ the infinitesimal generator of a contraction semigroup $(T(t))_{t\geqslant 0}$ on $L^2(I, \mathbb{R}^n)$. Let

$$D(\tilde{A}) = \{\varphi \in W^{1,2}(I, \mathbb{R}^n)\colon\ \varphi(0) = 0\}, \quad \tilde{A}\varphi = \frac{\mathrm{d}\varphi}{\mathrm{d}\theta}$$

and

$$[T(t)\varphi(\cdot)](\theta) = \begin{cases} \varphi(t+\theta) & \text{for } t \leqslant \theta, \\ 0 & \text{otherwise,} \end{cases}$$

where $t > 0$, $\theta \in I$.

In case $\Sigma \equiv 0$, we define a family $(\widetilde{S}(t))_{t\geqslant 0}$ of operators acting on M^2 by

$$\widetilde{S}(t)\begin{pmatrix} a \\ \varphi \end{pmatrix} = \begin{pmatrix} x(t) \\ x_t(\cdot) \end{pmatrix} \quad \text{for } \begin{pmatrix} a \\ \varphi \end{pmatrix} \in M^2.$$

The family $(\widetilde{S}(t))_{t\geqslant 0}$ is a C_0-semigroup of bounded linear operators. Following the idea used in [6, p. 500], we can introduce an equivalent norm in M^2 such that $(\widetilde{S}(t))_{t\geqslant 0}$ becomes a semigroup of contraction type.

Now we rewrite (4.1.8) in the following form for $z(t) = \binom{x(t)}{x_t(\cdot)}$:

$$\begin{aligned} \mathrm{d}z(t) = {} & \begin{pmatrix} \sum_{i=0}^{m} A_i x(t-r_i) + \int_{-r}^{0} A(\theta)x(t+\theta)\,\mathrm{d}\theta \\ \tilde{A}x_t(\cdot) \end{pmatrix} \mathrm{d}t + \\ & + \widehat{\Sigma}(x_t(\cdot))\,\mathrm{d}w(t), \\ z(0) = {} & z_0, \end{aligned}$$

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

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

We now define the operators $\mathcal{A}$ and $\mathcal{B}$ in the following manner. Let

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

$$\mathcal{A} \begin{pmatrix} \varphi(0) \\ \varphi(\cdot) \end{pmatrix} = \begin{pmatrix} \sum_{i=0}^{m} A_i \varphi(r_i) + \int_{-r}^{0} A(\theta)\varphi(\theta)\,\mathrm{d}\theta \\ \tilde{A}\varphi \end{pmatrix}$$

and

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

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 (4.1.8) has the form

$$\mathrm{d}z(t) = \mathcal{A}z(t)\,\mathrm{d}t + \mathcal{B}(z(t))\,\mathrm{d}w(t), \quad z(0) = z_0 \tag{4.1.9}$$

and assumptions (A1)–(A5) from Section 5.2 are imposed.

Now we present the construction of the correction term for such an evolution equation. It was H. Doss [26] who gave only the first ideas of how to constract this correction term but he did not prove the Wong–Zakai-type approximation theorem with this term. The latter was done in the author's papers [112–115].

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 [26])

$$\mathcal{B}_{\tilde{h}_1}(h, h') := \langle \Psi(h)(h'), \tilde{h}_1 \rangle_{H_1} \in \mathbb{R} \quad \text{for } h, h' \in H.$$

From the Riesz theorem for the form Ψ on H we conclude that for every $\tilde{h}_1 \in H_1$ there exists an operator $\tilde{\Psi}(\tilde{h}_1) \in L(H)$ such that for every $h, h' \in H$,

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

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

$$\tilde{\xi}\colon\ H_1 \ni \tilde{h}_1 \longrightarrow \mathrm{tr}(Q\tilde{\Psi}(\tilde{h}_1)) \in \mathbb{R}$$

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

$$\widetilde{\mathrm{tr}}(Q\Psi) = \tilde{\tilde{h}}_1. \tag{4.1.10}$$

Returning to the last example, we see that $\langle \tilde{\tilde{h}}_1, \tilde{h}_1 \rangle_{H_1}$ is the trace of the operator $Q\widetilde{\Psi}(\tilde{h}_1) \in L(H)$ but $\widetilde{\mathrm{tr}}(Q\Psi)$ is merely a symbol for $\tilde{\tilde{h}}_1$.

We observe that the term which is needed for the construction of the correction term (4.1.10) in Example 4.1.2 is

$$D\mathcal{B}(h_1)\mathcal{B}(h_1) = (D\Sigma(\varphi_0((\cdot))(\psi(\cdot))) \quad \text{for} \begin{pmatrix} a \\ \varphi(\cdot) \end{pmatrix} = h_1 \in M^2. \quad (4.1.11)$$

Now we would like to compare these two types of correction terms derived for stochastic delay equations in (4.1.7) and (4.1.10) for the different spaces occurring in these two models. In Example 4.1.2, we use the convention

$$\mathcal{B}(h_1) = \begin{pmatrix} \Sigma(\varphi(\cdot)) \\ 0 \end{pmatrix} = \begin{pmatrix} \psi(0) \\ \psi(\cdot) \end{pmatrix} \quad \text{for } h_1 = \begin{pmatrix} \varphi(0) \\ \varphi(\cdot) \end{pmatrix}, \quad (4.1.12)$$

and, hence, for the one-dimensional Wiener process we have

$$\begin{aligned} D\mathcal{B}(h_1)\begin{pmatrix} \psi(0) \\ \psi(\cdot) \end{pmatrix} &= \begin{pmatrix} D\Sigma(\varphi(\cdot))\psi(\cdot) \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} \alpha\psi(0) + \displaystyle\int_{-r}^{0} \psi(s)u(s)\,\mathrm{d}s \\ 0 \end{pmatrix}, \end{aligned} \quad (4.1.13)$$

and the correction term will only be $\alpha\psi(0)$, since the second coordinate $\int_{-r}^{0} \psi(s)u(s)\,\mathrm{d}s$ in (4.1.13) is zero because the second coordinate of the vector in (4.1.11) is zero. We observe that $\alpha = \mu(\{0\})$ and $u(s)$ is the density of the measure $\widetilde{\mu}$ (μ and $\widetilde{\mu}$ are defined above). So the term $\alpha\psi(0)$ is the main part of the correction term in (4.1.7).

4.2. FINITE-DIMENSIONAL CASE

Finally, it is also proved in papers [114], [115] that this infinite-dimensional form of the correction term gives the already known finite multi-dimensional form of the correction term.

We consider the case where $H = \mathbb{R}^d$, $H_1 = \mathbb{R}^n$. Let $x, z \in \mathbb{R}^n$. Then $\mathcal{B}\colon \mathbb{R}^n \to L(\mathbb{R}^d, \mathbb{R}^n)$, $D\mathcal{B}(z) \in L(\mathbb{R}^n, L(\mathbb{R}^d, \mathbb{R}^n))$ and $D\mathcal{B}(z)(x) \in L(\mathbb{R}^d, \mathbb{R}^n)$ are given by a matrix

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

We put

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

Let $D\mathcal{B}(z)\mathcal{B}(z)X = \Psi(X)$. Then

$$\Psi(X)Y = \sum_{j=1}^{d}\left(\sum_{l=1}^{n}\frac{\partial B_{ij}(z)}{\partial z_l}\left(\sum_{k=1}^{d} B_{lk}(z)\xi_k\right)\right)\eta_j$$

and, for $p \in \mathbb{R}^n$, we write the inner product

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

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

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

Consider the trace with $(Qm)_{jk}$, $j, k = 1, \dots, m$, being the restriction of the covariance operator Q to $\mathbb{R}^m$

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

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

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

The first vector,

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

is exactly the correction term $\tilde{\tilde{h}}_1$ obtained in (4.1.10).

It would also be interesting to compare other types of correction terms in finite dimensions for more general disturbances than the Wiener process, e.g., for semimatingales.

4.3. RELATION TO GENERAL APPROXIMATION THEORY

One of the most useful ways of approximation of functions in deterministic analysis is the polynomial approximation based on the Weierstrass theorem (see Akhiezer [1]) and Fourier series. This type of theorem can also be formulated for random functions and stochastic processes (see Onicescu [81]). We would expect that the sum of the Fourier series of a function f at x is $(f(x^+) - f(x^-))/2$ when we write the solution to a stochastic differential equation in terms of the initial value and a noise process. How do we generalize the Wong–Zakai approximations for stochastic differential equations (to motivate the occurrence of the correction term) by approximate counterparts using general approximation theory? This will be our interest when we try to unify the Wong–Zakai approximations theory.

5. Wong–Zakai Approximations in Infinite Dimensions for the Nonlinear Case

5.1. STOCHASTIC DELAY EQUATIONS

In this section we give the generalization of the Wong–Zakai theorem to nonlinear stochastic functional differential equations with values in the space $\mathbb{R}^d$ ($d \geqslant 1$) (see the author's papers [113, 114]). 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, that is, we obtain the so-called Itô correction term of the form (4.1.7).

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

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

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

Let $\mathcal{B}(\mathcal{X})$ denote the topological σ-algebra of the space $\mathcal{X}$. It is obviously identical with the σ-algebra generated by the family of all Borel cylinder sets in $\mathcal{X}$. So we construct the Wiener space $(\mathcal{C}_2^0, \mathcal{B}(\mathcal{C}_2^0), P^{\mathrm{w}})$, where P^{w} is the Wiener measure ([40, Chapter I]). The coordinate process $B(t,w) = w(t)$, $w \in \mathcal{C}_2^0$, is the m-dimensional Wiener process.

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

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

$$B^{n,p}(t,w) = w^p\left(\frac{k}{2^n}\right) + 2^n\left(t - \frac{k}{2^n}\right)\left(w^p\left(\frac{k+1}{2}\right) - w^p\left(\frac{k}{2^n}\right)\right) \quad (5.1.1)$$

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

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

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

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

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

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

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

Now we consider $\widetilde{\Omega} = C_2^0$. Let X be a continuous stochastic process $X(t,w)$: $(-\infty,T] \times \widetilde{\Omega} \to R^d$, that is, X: $\widetilde{\Omega} \to \mathcal{X} = C_1$.

We take some fixed initial constant stochastic processes

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

We also consider operators b: $\mathcal{C}_- \to \mathbb{R}^d$, σ: $\mathcal{C}_- \to L(\mathbb{R}^m,\mathbb{R}^d)$ (where $(L(\mathbb{R}^m,\mathbb{R}^d)$ 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

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

to give a meaning to the stochastic integral in (4.1.1).

We assume that

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

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

$$\widetilde{b}\colon \mathcal{C}_- \ni g \longrightarrow (J \ni \tau \to b(g(\cdot+\tau)) \in \mathbb{R}^d),$$
$$\widetilde{\sigma}\colon \mathcal{C}_- \ni g \longrightarrow (J \ni \tau \to \sigma(g(\cdot+\tau)) \in L(\mathbb{R}^m,\mathbb{R}^d)).$$

Remark 5.1.1. This construction explains why we take $(-\infty,0]$ for the domain of the initial function. If we took the interval $[-r,0]$, $0 < r < \infty$, instead, it would be impossible to define $\widetilde{b}$ and $\widetilde{\sigma}$ correctly (that is, for $X_t = Y_t$ it could happen that $\widetilde{\sigma}(X_t) \neq \widetilde{\sigma}(Y_t)$).

Let us introduce the following conditions:

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

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

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

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

($\tilde{\rm A}5$) For every $\varphi \in \mathcal{C}_-$ the following growth condition is satisfied:

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

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

($\tilde{\rm A}6$) We have

$$P\left(\int_0^T |b(X_s)| \, \mathrm{d}s < \infty \right) = 1, \qquad P\left(\int_0^T |\sigma(X_s)|_L^2 \, \mathrm{d}s < \infty \right) = 1.$$

($\tilde{\rm A}7$) We have $b^i, \sigma^{ip} \in C_b^1(\mathcal{C}_-)$ for every $i = 1, \ldots, d$, $p = 1, \ldots, m$, where C_b^1 denotes the space of bounded mappings with continuous first derivative.

Notice that our conditions ensure the existence and uniqueness of the strong solutions to Equations (4.1.1)–(4.1.3).

We have proved the following

THEOREM 5.1.1 (Twardowska [113, 114]). *Let conditions* ($\tilde{\rm A}2$)–($\tilde{\rm A}5$) *and* ($\tilde{\rm A}7$) *be satisfied. Let* $B^n(t, w)$ *be an approximation of type* (5.1.1) *of the Wiener process. We assume that* X^n *and* Y *are solutions to* ($4.1.2^n$) *and* (4.1.3), *respectively, with a constant initial stochastic process. Then conditions* ($\tilde{\rm A}1$) *and* ($\tilde{\rm A}6$) *are satisfied and for every* $T > 0$,

$$\lim_{n \to \infty} \sup_{0 \leqslant t \leqslant T} E\left[|X^n(t, w) - Y(t, w)|^2 \right] = 0. \tag{5.1.2}$$

Remark 5.1.2. Instead of the interval $J = (-\infty, 0]$, we can consider $I = [-r, 0]$, $0 < r < \infty$. Then, instead of considering $X^i(t_i^n + s) - X^i(t_{i-1}^n + s)$ on the whole interval of definition, we observe that

$$X^i(t_i^n + s) - X^i(t_{i-1}^n + s)$$

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

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

5.2. SEMILINEAR EVOLUTION EQUATIONS

In this section, we examine an approximation theorem of the Wong–Zakai-type for stochastic evolution equations in a Hilbert space such that the noise is 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 correction term (4.1.10) for the infinite-dimensional case (see Section 4.1). The uniqueness and existence of solutions are guaranteed by known theorems for mild solutions.

Let H and H_1 be real separable Hilbert spaces with norms $\|\cdot\|_H$, $\|\cdot\|_{H_1}$ and scalar products $(\cdot,\cdot)_H$, $(\cdot,\cdot)_{H_1}$ and with orthonormal bases $\{e_n\}_{n=1}^{\infty}$, $\{l_n\}_{n=1}^{\infty}$, respectively. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, P)$ be a filtered probability space on which an increasing and right-continuous family $(\mathcal{F}_t)_{t\in[0,T]}$ of complete sub-σ-algebras of $\mathcal{F}$ is defined. $L(H, H_1)$ denotes the space of bounded linear operators from H to H_1. Let $\mathcal{L}^2(H, H_1)$ 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 (see Curtain and Pritchard [22], Da Prato and Zabczyk [24]) that there are real-valued independent Wiener processes $\{\overset{j}{w}(t)\}_{j=1}^{\infty}$ on $[0,T]$ such that

$$
w(t) = \sum_{j=1}^{\infty} \overset{j}{w}(t) e_j \tag{5.2.1}
$$

almost everywhere in $(t,\omega) \in [0,T] \times \Omega$, where $\{e_j\}_{j=1}^{\infty}$ is an orthonormal basis of eigenvectors of Q corresponding to eigenvalues $\{\lambda_j\}_{j=1}^{\infty}$, $\sum_{j=1}^{\infty} \lambda_j < \infty$, with

$$
E[\Delta \overset{i}{w}\, \Delta \overset{j}{w}] = (t-s)\lambda_i \delta_{ij} \quad \text{for } \Delta \overset{j}{w} = \overset{j}{w}(t) - \overset{j}{w}(s) \text{ and } s < t
$$

(δ_{ij} is the Kronecker delta).

We consider the stochastic differential equation

$$\begin{aligned} \mathrm{d}z(t) &= \mathcal{A}z(t)\,\mathrm{d}t + \mathcal{C}(z(t))\,\mathrm{d}t + \mathcal{B}(z(t))\,\mathrm{d}w(t), \\ z(0) &= z_0, \end{aligned} \tag{5.2.2}$$

where

(A1) $(z(t))_{t\in[0,T]}$ is an H_1-valued stochastic process, $(w(t))_{t\geqslant 0}$ 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 $(S(t))_{t\geqslant 0}$, $\mathcal{C}$: $H_1 \to H_1$ and $\mathcal{B}$: $H_1 \to L(H, H_1)$ are bounded nonlinear operators. Moreover, we assume that $(S(t))_{t\geqslant 0}$ is of contraction type, i.e., there exists a constant $\beta \in R_+$ such that $\|S(t)\|_{H_1} \leqslant \exp(\beta t)$ for all $t \in [0, T]$,

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

Apart from (5.2.2) we consider the equation

$$\begin{aligned} \mathrm{d}\hat{z}(t) &= \mathcal{A}\hat{z}(t)\,\mathrm{d}t + \mathcal{C}(\hat{z}(t))\,\mathrm{d}t + \mathcal{B}(\hat{z}(t))\,\mathrm{d}w(t) + \\ &\quad + \tfrac{1}{2}\widetilde{\mathrm{tr}}(QD\mathcal{B}(\hat{z}(t))\mathcal{B}(\hat{z}(t)))\,\mathrm{d}t, \\ \hat{z}(0) &= z_0, \end{aligned} \tag{5.2.3}$$

where $\widetilde{\mathrm{tr}}(QD\mathcal{B}(\hat{z}(t))\mathcal{B}(\hat{z}(t)))$ is defined in Section 4.1.

Moreover, we assume that

(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

$$\begin{aligned} &\|R^{-1}\mathcal{C}(h_1)\|_{H_1}^2 + \|R^{-1}B(h_1)Q^{1/2}\|_{\mathrm{HS}}^2 + \\ &\quad + \|R^{-1}\mathrm{tr}(QD\mathcal{B}(h_1)\mathcal{B}(h_1))\|_{H_1}^2 \leqslant K(1 + \|h_1\|_{H_1}^2), \end{aligned} \tag{i}$$

$$\begin{aligned} &\|\mathcal{C}(h_1) - \mathcal{C}(\tilde{h}_1)\|_{H_1}^2 + \mathrm{tr}((\mathcal{B}(h_1) - \mathcal{B}(\tilde{h}_1))Q(\mathcal{B}(h_1) - \mathcal{B}(\tilde{h}_1))^*) \\ &\quad \leqslant K\|h_1 - \tilde{h}_1\|_{H_1}^2 \end{aligned} \tag{ii}$$

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

(A4) the operator $\mathcal{C}$ is of class C^1 and the operator $B \in C_b^1$, i.e. is of class C^1 with bounded derivative; this derivative is assumed to be globally Lipschitzian,

(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 $h_1 \in H_1$,

$$\|D\mathcal{B}(h_1)\mathcal{A}h_1\|_{L(H,H_1)} \leqslant k. \tag{5.2.4}$$

We now define the nth approximation of the Wiener process $(w(t))_{t\geqslant 0}$ as follows:

$$w_{(n)}(t) = \sum_{j=1}^{\infty} w_{j,n}(t)e_j, \tag{5.2.5}$$

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

$$w_{j,n}(t) = \frac{t - t_{i-1}^n}{t_i^n - t_{i-1}^n} \overset{j}{w} (t_i^n) + \frac{t_i^n - t}{t_i^n - t_{i-1}^n} \overset{j}{w} (t_{i-1}^n). \tag{5.2.6}$$

Consider now the sequence of integral equations

$$\begin{aligned} u_{(n)}(t) &= S(t)z_0 + \int_0^t S(t-s)\mathcal{C}(u_{(n)}(s))\,\mathrm{d}s + \\ &\quad + \int_0^t S(t-s)\mathcal{B}(u_{(n)}(s))\,\mathrm{d}w^n(s), \\ u_{(n)}(0) &= z_0, \quad n = 1, 2, \ldots . \end{aligned} \tag{5.2.7$_n$}$$

We consider solutions in the mild sense (see [22, 24]). The uniqueness of solutions is understood in the sense of trajectories. Our assumptions ensure the existence and uniqueness of solutions.

We have proved the following

THEOREM 5.2.1 (Twardowska [114, 115]). *Let* $(w_{(n)}(t))_{t\geqslant 0}$ *be the* nth *approximation of the Wiener process* $(w(t))_{t\geqslant 0}$ *as given in* (5.2.5). *Let* $(u_{(n)}(t))_{t\in[0,T]}$ *be the solution to Equation* (5.2.7$_n$) *and* $\widehat{z}(t)$ *to Equation* (5.2.3). *Assume that hypotheses* (A1)–(A5) *are satisfied and* $E[\|R^{-1}z_0\|_{H_1}^2] < \infty$. *Then, for each* T, $0 < T < \infty$, *and given* $\varepsilon > 0$

$$\lim_{n\to\infty} P\Big(\sup_{0\leqslant t\leqslant T} \|u_{(n)}(t,\omega) - \widehat{z}(t,\omega)\|_{H_1} \geqslant \varepsilon \Big) = 0. \tag{5.2.8}$$

5.3. NONLINEAR STOCHASTIC PARTIAL DIFFERENTIAL EQUATIONS

We consider a generalization of the Wong–Zakai approximation theorem for stochastic nonlinear partial differential equations with unbounded monotone and coercive operators defined in Gelfand triples (Lions [59], Pardoux [82]). A version of the theorem for sums of such operators is also included.

Similar equations have already been studied in the linear case, e.g., by Bensoussan [10], Gyöngy [35], Rozovskiĭ [93] and in the nonlinear case by Bensoussan and Temam [11], Krylov and Rozovskiĭ [50], as well as by Pardoux ([82, 83]) from whom we take the model of this paper. In the above papers, mainly the existence and uniqueness theorems were given.

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, P)$, H and H_1 be the spaces defined in Section 5.2. We also consider a real separable reflexive Banach space V which is continuously and densely embedded in the Hilbert space H_1. Then, identifying H_1 with its dual space H_1^* (by the scalar product in H_1), we have, denoting by V^* the dual space to V,

$$V \subset H_1 = H_1^* \subset V^*.$$

The embeddings are continuous with dense range. The above spaces are endowed with the norms $\|\cdot\|_V$, $\|\cdot\|_{H_1}$, and $\|\cdot\|_{V^*}$, respectively. The pairing between V and V^* is denoted by $\langle\cdot,\cdot\rangle$.

Let $L(X,Y)$ denote the vector space of continuous linear operators from X to Y, with the operator norm $\|\cdot\|_{L(X,Y)}$, where X and Y are arbitrary Banach spaces (we put $L(X) = L(X,X)$); $L^p(\Omega;X)$, $\infty \geqslant p \geqslant 1$, denotes the usual Banach space of equivalence classes of random variables with values in X which are p-integrable (essentially bounded for $p=\infty$) with respect to the measure P, with norm $\|\cdot\|_{L^p(\Omega;X)}$. Moreover, $\mathcal{L}^1(X,Y)$ is the Banach space of nuclear operators from X to Y with the trace norm $\|\cdot\|_{\mathcal{L}^1(X,Y)}$, and $\mathcal{L}^2(X,Y)$ is the Hilbert space of Hilbert–Schmidt operators with the norm $\|\cdot\|_{\mathrm{HS}}$, where X and Y are arbitrary separable Hilbert spaces. They are subspaces of $L(X,Y)$.

We put for the Wiener process $w(t)$ (defined in Section 5.2)

$$w^m(t) = \sum_{j=1}^{m} \overset{j}{w}(t) e_j = \sum_{j=1}^{m} (w(t), e_j)_H e_j. \tag{5.3.1}$$

Now we define the nth polygonal approximations of the process $(w(t))_{t\in[0,T]}$ by $w_{(n)}(t)$ in (5.2.5) and of $(w^m(t))_{t\in[0,T]}$ by

$$w_{(n)}^m(t) = \sum_{j=1}^{m} w_{j,n}(t) e_j, \tag{5.3.2}$$

where $w_{j,n}(t)$ is given by (5.2.6).

We consider the stochastic nonlinear differential equation

$$\begin{aligned} &\mathrm{d}u(t) + A(t,u(t))\,\mathrm{d}t + B(t,u(t))\,\mathrm{d}w(t) = f(t)\,\mathrm{d}t, \\ &u(0) = u_0, \end{aligned} \tag{5.3.3}$$

where $(u(t))_{t\in[0,T]}$ is an H_1-valued stochastic process, and

($\tilde{\tilde{\mathrm{A}}}$1) u_0 is an H_1-valued square integrable $\mathcal{F}_0$-measurable random variable, that is, $u_0 \in L^2(\Omega, \mathcal{F}_0, P; H_1)$.

For every $n \in \mathbb{N}$ we consider the approximation equation

$$\begin{aligned} &\mathrm{d}u_{(n)}(t) + A(t,u_{(n)}(t))\,\mathrm{d}t + B(t,u_{(n)}(t))\,\mathrm{d}w_{(n)}(t) = f(t)\,\mathrm{d}t, \\ &u_{(n)}(0) = u_0, \end{aligned} \tag{5.3.4$_n$}$$

where $w_{(n)}(t)$ are given by (5.2.5).

Moreover, we consider the equation

$$\begin{aligned} &\mathrm{d}\widehat{u}(t) + A(t,\widehat{u}(t))\,\mathrm{d}t + B(t,\widehat{u}(t))\,\mathrm{d}w(t)+\\ &\quad+\tfrac{1}{2}\widetilde{\mathrm{tr}}(QDB(t,\widehat{u}(t))B(t,\widehat{u}(t)))\,\mathrm{d}t = f(t)\,\mathrm{d}t, \\ &\widehat{u}(0) = u_0, \end{aligned} \tag{5.3.5}$$

where $\widetilde{\mathrm{tr}}(QDB(t,\widehat{u}(t))B(t,\widehat{u}(t)))$ is as described in Section 4.1.

We assume (compare Pardoux [82]) that the family of operators $A(t,\cdot)\colon V \to V^*$ defined for almost every (a.e.) $t \in (0,T)$ and for some $p > 1$ has the following properties:

$(\tilde{\tilde{\mathrm{A}}}2)$ growth restriction: there exists a constant β such that

$$\|A(t,u)\|_{V^*} \leqslant \beta\|u\|_V^{p-1} \quad \text{for every } u \in V \text{ and for a.e. } t,$$

$(\tilde{\tilde{\mathrm{A}}}3)$ hemicontinuity: the mapping

$$\mathbb{R} \ni \theta \longrightarrow \langle A(t,u+\theta v), w\rangle \in \mathbb{R}$$

is continuous for all $u, v, w \in V$ and for a.e. t,

$(\tilde{\tilde{\mathrm{A}}}4)$ measurability: the mapping

$$(0,T) \ni t \longrightarrow A(t,u) \in V^*$$

is Lebesgue measurable for every $u \in V$.

We further assume that the family of operators $B(t,\cdot)\colon V \to \mathcal{L}^2(H,H_1)$ defined for a.e. $t \in (0,T)$ satisfies the following assumptions:

$(\tilde{\tilde{\mathrm{A}}}5)$ boundedness: there exists a constant $\tilde{L}$ such that

$$\|B(t,u)\|^2_{L(H,H_1)} \leqslant \tilde{L}$$

for all $u \in V$,

$(\tilde{\tilde{\mathrm{A}}}6)$ the operator $B(t,\cdot) \in C^1_{\mathrm{b}}$, i.e., is of class C^1 with bounded derivative and this derivative is assumed to be globally Lipschitzian,

$(\tilde{\tilde{\mathrm{A}}}7)$ measurability: for every $u \in V$ the mapping

$$(0,T) \ni t \longrightarrow B(t,u) \in \mathcal{L}^2(H,H_1)$$

is Lebesgue measurable.

Moreover, we assume

$(\tilde{\tilde{\mathrm{A}}}8)$ coercivity: there exist constants $\alpha > 0$, λ and ν such that

$$2\langle A(t,u),u\rangle + \lambda\|u\|^2_{H_1} + \nu \geqslant \alpha\|u\|^2_V + \|B(t,u)\|^2_{\mathrm{HS}}$$

for every $u \in V$ and a.e. t,

($\tilde{\tilde{A}}9$) monotonicity:

$$2\langle A(t,u) - A(t,v), u - v\rangle + \lambda\|u - v\|_{H_1}^2 \geqslant \|B(t,u) - B(t,v)\|_{\mathrm{HS}}^2$$

for all $u, v \in V$ and a.e. t,

($\tilde{\tilde{A}}10$) boundedness of $DB(t,\cdot)$ on V in the sense of the norm in H_1: there exists a constant $\tilde{\tilde{L}}$ such that

$$\|DB(t,u)k\|_{\mathrm{HS}} \leqslant \tilde{\tilde{L}}\|k\|_{H_1}$$

for all $u \in V$, $k \in H_1$.

Finally, we assume

($\tilde{\tilde{A}}11$) $f \in L^{p'}((0,T) \times \Omega; V^*)$, where $\frac{1}{p} + \frac{1}{p'} = 1$, and f is nonanticipating.

DEFINITION 5.3.1. Suppose we are given an H_1-valued initial random variable u_0, and an H-valued Wiener process $(w(t))_{t\in[0,T]}$.

Suppose further that an H_1-valued stochastic process $(u(t))_{t\in[0,T]}$ has the following properties:

(i) $(u(t))_{t\in[0,T]}$ is well measurable (see [82]),

(ii) $u(t) \in L^p((0,T) \times \Omega; V) \cap L^2(\Omega; C(0,T; H_1))$,

(iii) there exists a set $\Omega' \subset \Omega$ such that $P(\Omega') = 1$ and for all $(t,\omega) \in [0,T] \times \Omega'$ and $y \in Y \subset H$ (Y is an everywhere dense, in the strong topology, subset of H) Equation (5.3.3) is satisfied in the following sense:

$$\begin{aligned}(y, u(t,\omega))_H &= (y, u_0(\omega))_H - \int_0^t \langle A(s, u(s,\omega)), y\rangle\,\mathrm{d}s - \\ &\quad - \left(y, \int_0^t B(s, u(s,\omega))\,\mathrm{d}w(s)\right)_H + \langle f(t,\omega), y\rangle.\end{aligned}$$

Equivalently, this equality may be understood in V^* (see Krylov and Rozovskiĭ [50]).

Then $(u(t))_{t\in[0,T]}$ is called a *solution to* (5.3.3) *with initial condition* u_0. The uniqueness of solution is understood in the sense of trajectories.

By our assumptions, the existence and uniqueness of solutions is guaranteed.

Denote by $V_m = H_{1,m} = V_m^*$ the vector space spanned by the vectors $l_1, \ldots, l_m$ and let $P_m \in L(H_1, H_{1,m})$ be the orthogonal projection. We introduce in $H_{1,m}$ the norm

$$\|u\| = \left(\sum_{j=1}^m |u_j|^2\right)^{1/2}$$

for $u = (u_1, \ldots, u_m)$, and the usual scalar product $(\cdot,\cdot)$.

We extend P_m to an operator $V^* \to V_m^*$ by

$$\widetilde{P}_m u = \sum_{j=1}^{m} \langle u, l_j \rangle l_j \quad \text{for } u \in V^*.$$

We denote by H_m the vector space spanned by the vectors $e_1, \ldots, e_m$. Let $\Pi_m \in L(H, H_m)$ be the orthogonal projection.

Now, we define the families of operators $A^m(t, \cdot)$: $V_m \to V_m^*$ by

$$A^m(t, u) := \widetilde{P}_m A(t, u) \quad \text{for } u \in V_m, \tag{5.3.6}$$

and $B^m(t, \cdot)$: $H_{1,m} \to \mathcal{L}^2(H_m, H_{1,m})$ by

$$B^m(t, u) := P_m B(t, u) \quad \text{for } u \in H_{1,m}. \tag{5.3.7}$$

Let $w^m(t)$ be the Wiener process with values in H_m defined by

$$w^m(t) = \Pi_m w(t).$$

Clearly, it can be represented by formula (5.3.1). Moreover, we put

$$f^m = P_m f \in L^{p'}(\Omega \times (0, T); V_m^*),$$

which can be represented by formula (2.2), and

$$u_0^m = P_m u_0 \in L^2(\Omega; H_{1,m}).$$

Now, we consider the following stochastic differential equation of Itô type in the space $\mathbb{R}^m$ for the ith coordinate of a process $v^m(t) = (v_1^m(t), \ldots, v_m^m(t)) \in H_{1,m}$:

$$\begin{aligned} &\mathrm{d}v_i^m(t) + A_i^m(t, v^m(t))\,\mathrm{d}t + B_i^m(t, v^m(t))\,\mathrm{d}w^m(t) + \\ &\quad + \frac{1}{2}\sum_{j=1}^{m}\sum_{l=1}^{m} \lambda_j \frac{\partial B_{ij}^m(t, v^m(t))}{\partial v_l^m} B_{lj}^m(t, v^m(t))\,\mathrm{d}t = f_i^m(t)\,\mathrm{d}t, \qquad (5.3.8^m) \\ &v^m(0) = u_0^m, \end{aligned}$$

where $(B_{ij}^m(t, v^m(t)))_{i,j=1,\ldots,m}$ is the matrix representation of elements of $B^m(t, v^m(t))$.

For every $n \in \mathbb{N}$, we also consider the approximation equation

$$\begin{aligned} &\mathrm{d}v_{(n)}^m(t) + A^m(t, v_{(n)}^m(t))\,\mathrm{d}t + B^m(t, v_{(n)}^m(t))\,\mathrm{d}w_{(n)}^m(t) \\ &\quad = f^m(t)\,\mathrm{d}t, \qquad (5.3.9_n^m) \\ &v_{(n)}^m(0) = u_0^m, \end{aligned}$$

where $w_{(n)}^m(t)$ are given by (5.3.2). We observe that $\mathrm{d}w_{(n)}^m(t) = \dot{w}_{(n)}^m \mathrm{d}t$ on every interval $(t_{i-1}^n, t_i^n]$, so Equations $(5.3.9_n^m)$ are of a deterministic nature for almost every $\omega \in \Omega$.

We start with

LEMMA 5.3.1 [112]. *The correction term*

$$\left(\frac{1}{2}\sum_{j=1}^{m}\sum_{l=1}^{m}\lambda_j \frac{\partial B_{ij}^m(t,v^m(t))}{\partial v_l^m} B_{lj}^m(t,v^m(t))\right)_{i=1,\dots,m}$$

in Equation ($5.3.8^m$) *is the result of applying the projection operators* P_m *and* Π_m *to the operator* $B(t,\cdot)$ *and to the Wiener process* $(w(t))_{t\in[0,T]}$ *in the construction of the term* $\frac{1}{2}\widetilde{\operatorname{tr}}(QDB(t,\widehat{u}(t))B(t,\widehat{u}(t)))$, *that is,*

$$\sum_{j=1}^{m}\sum_{l=1}^{m}\lambda_j \frac{\partial B_{\cdot j}^m(t,v^m(t))}{\partial v_l^m} B_{lj}^m(t,v^m(t)) \longrightarrow \widetilde{\operatorname{tr}}(QDB(t,\widehat{u}(t))B(t,\widehat{u}(t))) \quad \text{weakly in } H_1. \tag{5.3.10}$$

Further, we quote

LEMMA 5.3.2 (Pardoux [82, Theorem 2.1, p. 93]). *Under assumptions* ($\tilde{\tilde{\mathrm{A}}}1$)–($\tilde{\tilde{\mathrm{A}}}4$), ($\tilde{\tilde{\mathrm{A}}}6$)–($\tilde{\tilde{\mathrm{A}}}11$) *Equation* (5.3.8) *has exactly one solution* $v^m(t) \in H_{1,m}$ *and*

$$v^m \in L^p((0,T)\times\Omega;V_m)\cap L^2(\Omega;C(0,T;H_{1,m})). \tag{5.3.11}$$

LEMMA 5.3.3 [112]. *Let* $v^m(t)$, $\widehat{u}(t)$ *be the solutions to Equations* (5.3.8) *and* (5.3.5), *respectively, under assumptions* ($\tilde{\tilde{\mathrm{A}}}1$)–($\tilde{\tilde{\mathrm{A}}}4$), ($\tilde{\tilde{\mathrm{A}}}6$)–($\tilde{\tilde{\mathrm{A}}}11$). *Then for each* $t\in[0,T]$, $0<T<\infty$, *we have*

$$\lim_{m\to\infty} E\left[\|v^m(t)-\widehat{u}(t)\|_{H_1}^2\right] = 0. \tag{5.3.12}$$

Finally, we have proved a most important lemma which shows the difference between the method of proving Theorems 5.3.1 and 5.3.2 from the up-till-now proofs of Wong–Zakai theorems in finite and inifinite dimensions. Before, the differences between the approximate and limit equations were examined and by artificially added and subtracted terms, the forms of correction terms were obtained. Now, given the Galerkin approximation and then the Wong–Zakai approximation of the solution to (5.3.3), we have to prove the convergence of this approximate sequence to the limit equation as $m\to\infty$ independently of $n\to\infty$. We have

LEMMA 5.3.4 [112]. *Let* $A^m(t,\cdot)$ *and* $B^m(t,\cdot)$ *be given by* (5.3.6) *and* (5.3.7), *respectively, under assumptions* ($\tilde{\tilde{\mathrm{A}}}1$)–($\tilde{\tilde{\mathrm{A}}}11$). *Let* $w_{(n)}^m(t)$ *be given by* (5.3.2). *Assume that* $v_{(n)}^m(t)$ *and* $u_{(n)}(t)$ *are solutions to Equations* ($5.3.9_n^m$) *and* ($5.3.4_n$), *respectively. Then, for every* $t\in[0,T]$, $0<T<\infty$, *we have*

$$\lim_{m\to\infty} E\left[\|v_{(n(m))}^m(t)-u_{(n(m))}^m(t)\|_{H_1}^2\right] = 0, \tag{5.3.13}$$

where $\{n(m)\}$ is an arbitrary increasing sequence depending on m.

We have proved the following

THEOREM 5.3.1 (Twardowska [112]). *Let $\widehat{u}(t)$ and $u_{(n)}(t)$ be solutions to Equations (5.3.5) and (5.3.4$_n$), respectively. Assume that $(\tilde{\tilde{\mathrm{A}}}1)$–$(\tilde{\tilde{\mathrm{A}}}11)$ are satisfied. Take approximations $w_{(n)}(t)$ of the Wiener process $w(t)$ given by (5.2.5). Then, for each $t \in [0,T]$, $0 < T < \infty$, we have*

$$\lim_{n\to\infty} E\left[\|u_{(n)}(t) - \widehat{u}(t)\|^2_{H_1}\right] = 0. \tag{5.3.14}$$

Now we consider equations involving sums of operators. Let $A_i(t,\cdot)\colon V_i \to V_i^*$ and $B_i(t,\cdot)\colon V_i \to \mathcal{L}^2(H,H_1)$, $i = 1,\dots,q$, be q families of operators. We have proved ([112]) the approximation theorem of Wong–Zakai-type for the following equation:

$$\begin{aligned} &du(t) + \sum_{i=1}^{q} A_i(t,u(t))\,dt + \sum_{i=1}^{q} B_i(t,u(t))\,dw(t) = f(t)\,dt,\\ &u(0) = u_0. \end{aligned} \tag{5.3.15}$$

Several papers (Aquistapace and Terreni [2], Brzeźniak, Capiński and Flandoli [14], Curtain and Pritchard [22], Twardowska [115]) deal with stochastic evolution equations where the unbounded operator is the infinitesimal generator of a semigroup, as in Section 5.2. In other papers (Bensoussan [10], Bensoussan and Temam [11], Gyöngy [35], Krylov and Rozovskiĭ [50], Pardoux [82, 83], Rozovskiĭ [93]) this assumption is replaced by the coercivity of this operator. From Tanabe's book [105] we know that in some cases, evolution equations of the second form can be transformed to the first form. Sometimes, the two assumptions hold simultaneously. For example, the Laplacian is both coercive and is the infinitesimal generator of a semigroup.

On the other hand, these results here are in a sense more general than those in the author's papers [114] and [115], because the operator $A(t,\cdot)$ is allowed to be nonlinear. A nonlinear operator A cannot be used in the semilinear model. Moreover, the model presented here admits consideration of operators depending explicitly on a random event $\omega \in \Omega$, that is, the model considered by Krylov and Rozovskiĭ in [51].

We omit here the strong assumptions (A5), (A3)(i) and (A3)(ii) from Section 5.2 used in [114, 115] to ensure a boundedness property of the composition of the operators, and to ensure certain Lipschitz and growth conditions. Those assumptions were introduced in the papers for technical reasons in the proof of a approximation procedure of the Wong–Zakai-type. However, in [114, 115] we do not assume the monotonicity of operators.

Therefore, the present version of the Wong–Zakai approximation theorem in infinite dimensions is in the above sense more general than in [114, 115].

5.4. STOCHASTIC NAVIER–STOKES EQUATIONS

We consider in this section an approximation theorem (see [113]) for stochastic Navier–Stokes differential equations.

Similar equations were already studied, e.g., by Bensoussan [9], Bensoussan and Temam [12], Capiński [16], Capiński and Cutland [17]. In the above papers, mainly existence and uniqueness theorems were given.

We keep all the notation from Section 5.3. We also consider real separable Hilbert spaces V and W which are continuously and densely embedded in the Hilbert space H_1. Moreover, the inclusion $V \to H_1$ is compact. Then, identifying H_1 with its dual space H_1^* (by the scalar product in H_1) we have, denoting by V^* and W^* the dual spaces to V and W, respectively,

$$W \subset V \subset H_1 = H_1^* \subset V^* \subset W^*.$$

The embeddings are continuous with dense ranges. The above spaces are endowed with the norms $\|\cdot\|_W$, $\|\cdot\|_V$, $\|\cdot\|_{H_1}$ and $\|\cdot\|_{V*}$, $\|\cdot\|_{W*}$, respectively. The pairing between V and V^* (as well as between W and W^*) is denoted by $\langle\cdot,\cdot\rangle$.

DEFINITION 5.4.1. Denote by $\mathcal{P}$ the σ-algebra of sets on $[0,T]\times\Omega$ generated by all $\mathcal{F}_t$-adapted and left continuous X-valued stochastic processes.

An X-valued stochastic process $(X_t)_{t\in[0,T]}$ is called *predictable* if the mapping $(t,\omega)\to X_t(\omega)$ is $\mathcal{P}$-measurable.

We consider the stochastic nonlinear differential equation

$$\begin{aligned}&\mathrm{d}u(t) + A(t)u(t)\,\mathrm{d}t + G(u(t))\,\mathrm{d}t + B(t,u(t))\,\mathrm{d}w(t) = f(t)\,\mathrm{d}t,\\ &u(0) = u_0,\end{aligned} \tag{5.4.1}$$

where $(u(t))_{t\in[0,T]}$ is an H_1-valued stochastic process and

(B1) u_0 is an H_1-valued square integrable F_0-measurable random variable, that is, $u_0 \in L^2(\Omega, F_0, P; H_1)$.

For every $n\in\mathbb{N}$ we consider the approximation equation

$$\begin{aligned}&\mathrm{d}u_{(n)}(t) + A(t)u_{(n)}(t)\,\mathrm{d}t + G(u_{(n)}(t))\,\mathrm{d}t+\\ &\quad+B(t,u_{(n)}(t))\,\mathrm{d}w_{(n)}(t) = f(t)\,\mathrm{d}t, u_{(n)}(0) = u_0,\end{aligned} \tag{5.4.2$_n$}$$

where $w_{(n)}(t)$ is given by (5.2.5).

Moreover, we consider the equation

$$\begin{aligned}&\mathrm{d}\widehat{u}(t) + A(t)\widehat{u}(t)\,\mathrm{d}t + G(\widehat{u}(t))\,\mathrm{d}t + B(t\widehat{u}(t))\,\mathrm{d}w(t)+\\ &\quad+\tfrac{1}{2}\widetilde{\mathrm{tr}}(QDB(t,\widehat{u}(t))B(t,\widehat{u}(t)))\,\mathrm{d}t = f(t)\,\mathrm{d}t,\ \widehat{u}(0) = u_0,\end{aligned} \tag{5.4.3}$$

with $\widetilde{\mathrm{tr}}(QDB(t,\widehat{u}(t))B(t,\widehat{u}(t)))$ as described before.

Remark 5.4.1. Assumption (B7) ensures the correctness of the definition of $DB(t,h_1) \circ B(t,h_1) \in L(H, L(H, H_1))$ for $h_1 \in H_1$ because $DB(t,h_1)\colon H_1 \supset V \to L(H, H_1)$ is bounded on V in the norm of H_1.

We assume that the family of operators $A(t) \in L(V, V^*)$ defined for (a.e.) $t \in (0,T)$ has the following properties:

(B2) growth restriction: there exists a constant β such that

$$\|A(t)u\|_{V^*} \leqslant \beta \|u\|_V \quad \text{for every } u \in V \text{ and for a.e. } t,$$

(B3) coercivity: there exist constants $\alpha > 0$ and λ, $\widetilde{\nu}$ such that

$$2\langle A(t)u, u\rangle + \lambda \|u\|_{H_1}^2 + \widetilde{\nu} \geqslant \alpha \|u\|_V^2 + \|B(t,u)\|_{\mathrm{HS}}^2$$

for every $u \in V$ and a.e. t,

(B4) measurability: the mapping

$$(0,T) \ni t \to \langle A(t)u, v\rangle \in \mathbb{R}$$

is Lebesgue measurable for every $u, v \in V$.

The family of operators $B(t,\cdot)\colon V \to \mathcal{L}^2(H, H_1)$ defined for a.e. $t \in (0,T)$ satisfies the following assumptions:

(B5) boundedness: there exists a constant $\tilde{L}$ such that

$$\|B(t,u)\|_{\mathrm{HS}} \leqslant \tilde{L}$$

for all $u \in V$,

(B6) $B(t,\cdot) \in C_{\mathrm{b}}^1$, i.e., is of class C^1 with bounded derivative and this derivative is assumed to be globally Lipschitzian,

(B7) boundedness of $DB(t,\cdot)$ on V in the sense of the norm in H_1: there exists a constant $\tilde{\tilde{L}}$ such that

$$\|DB(t,u)k\|_{\mathrm{HS}} \leqslant \tilde{\tilde{L}} \|k\|_{H_1}$$

for all $u \in V$, $k \in H_1$,

(B8) measurability: for every $u \in V$ the mapping

$$(0,T) \ni t \to B(t,u) \in \mathcal{L}^2(H, H_1)$$

is Lebesgue measurable.

The bilinear continuous mapping $G\colon V \times V \to W^*$ satisfies the following assumptions:

(B9) $\langle G(u,v), v\rangle = 0$ for every $u \in V$ and $v \in W$,

(B10) boundedness: there exists a constant $\widetilde{C}$ such that

$$\|G(u,v)\|_{W^*} \leqslant \widetilde{C} \|u\|_{H_1}^{1/2} \|v\|_{H_1}^{1/2} \|u\|_V^{1/2} \|v\|_V^{1/2}$$

for all $u, v \in V$.

Finally, we assume

(B11) $f \in L^2((0,T) \times \Omega; V^*)$ and f is nonanticipating.

Put $G(u) = G(u,u)$ for every $u \in V$.

The solutions to Equations (5.4.1)–(5.4.3) are understood to be like in Section 5.3.

Denote by $\mathcal{O}$ an open bounded subset of $\mathbb{R}^2$ with a regular boundary $\partial\mathcal{O}$. Let $H^s(\mathcal{O})$ be the Sobolev space of functions y which are in $L^2(\mathcal{O})$ together with all their derivatives of order $\leqslant s$; $s > 2$ ($2 = 2/n + 1$, where n is the dimension of $\mathbb{R}^2$, i.e. 2). Further, $H_0^1(\mathcal{O}) \subset H^1(\mathcal{O})$ is the Hilbert subspace of functions vanishing on $\partial\mathcal{O}$. We also introduce the product Hilbert spaces $(L^2(\mathcal{O}))^2$, $(H_0^1(\mathcal{O}))^2$, $(H^s(\mathcal{O}))^2$, with the appropriate scalar products.

We consider the set $\mathcal{V}(\mathcal{O})$ of functions from C^∞ with a compact support in $\mathcal{O}$.

Put

$$\mathcal{E} = \left\{ y = (y_1, y_2)\colon\ y_i \in \mathcal{V}(\mathcal{O}),\ \operatorname{div} y = \sum_{i=1}^{2} \frac{\partial y_i}{\partial x_i} = 0 \right\}$$

and

$$\begin{aligned} H_1 &= \text{ the closure of } \mathcal{E} \text{ in } (L^2(\mathcal{O}))^2, \\ V &= \text{ the closure of } \mathcal{E} \text{ in } (H_0^1(\mathcal{O}))^2, \\ W &= \text{ the closure of } \mathcal{E} \text{ in } (H^s(\mathcal{O}))^2. \end{aligned}$$

These spaces have the structure of the Hilbert spaces induced by $(L^2(\mathcal{O}))^2$, $(H_0^1(\mathcal{O}))^2$, $(H^s(\mathcal{O}))^2$, that is,

$$\begin{aligned} ((y,z))_{H_1} &= \sum_{i=1}^{2} (y_i, z_i)_{L^2(\mathcal{O})}, \\ ((y,z))_V &= \sum_{i=1}^{2} (y_i, z_i)_{H_0^1(\mathcal{O})}, \\ ((y,z))_W &= \sum_{i=1}^{2} (y_i, z_i)_{H^s(\mathcal{O})}. \end{aligned}$$

It is obvious that W, V and H_1 have all the properties as in the abstract model (5.4.1).

Let $\nu > 0$ be fixed. We define the family of operators $A(t) \in L(V, V^*)$ by

$$\langle A(t)y, z\rangle = \nu((y,z))_V$$

for all $y, z \in$V. Therefore, assumptions (B2), (B3), and (B4) (for $\alpha = \nu$, $\lambda = 0$, $\widetilde{\nu} = 0$) are satisfied. Further, we consider a trilinear form

$$b(y,z,w) = \sum_{i,j=1}^{2} \int_{\mathcal{O}} y_i(x) \frac{\partial z_j(x)}{\partial x_i} w_j(x)\, \mathrm{d}x$$

defined and continuous on $V \times V \times W$. We recall (Lions [59, p. 67 and p. 71]) that for a positive constant C_1 we have

$$|b(y, y, w)| \leqslant C_1 \|y\|^2_{(L^4(\mathcal{O}))^2} \|w\|_{(H^1(\mathcal{O}))^2} \tag{5.4.4}$$

for all $y, w \in V$. Next, we define a bilinear continuous operator G: $V \times V \to W^*$ by

$$\langle G(y, z), w\rangle = b(y, z, w)$$

for all $y, z \in V$ and $w \in$W.

It is easy to check that assumptions (B9) and (B10) are satisfied.

We consider the following stochastic Navier–Stokes equation

$$\begin{aligned} &\mathrm{d}u - \nu\Delta u\,\mathrm{d}t + (u\cdot\nabla)u\,\mathrm{d}t + \nabla p\,\mathrm{d}t + B(u)\,\mathrm{d}w(t) = f(t)\,\mathrm{d}t,\\ &u = 0 \quad \text{on } \Sigma = [0, T] \times \partial\mathcal{O},\\ &u(0) = u_0 \quad \text{in } \mathcal{O},\\ &\operatorname{div} u = 0 \quad \text{in } [0, T] \times \mathcal{O}, \end{aligned} \tag{5.4.5}$$

where $u = u(t, x)$ is the velocity field of a fluid and $p = p(t, x)$ is the pressure.

The reduction to the abstract form (5.4.1) is completely classical and we omit it. Further, we shall understand Equation (5.4.1) as the above Navier–Stokes equation for $\mathcal{O} \subset \mathbb{R}^2$.

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

The existence and uniqueness of our solutions under assumptions (B1)–(B4), (B8)–(B12) follows from the following modification of Theorem 6.3 in the paper of Capiński and Cutland [17]. Namely, we omit the assumption on the periodic boundary condition that we only need to prove the uniqueness of the solution. The uniqueness is obtained from the paper of Capiński [16].

We proved in [113] the following theorem similarly to Theorem 5.3.1 using analogs of Lemmas 5.3.1–5.3.4. Some difficulties arose in the analog of Lemma 5.3.3.

THEOREM 5.4.1 (Twardowska [113]). *Let $\widehat{u}(t)$ and $u_{(n)}(t)$ be solutions to Equations* (5.4.3) *and* $(5.4.2_n)$, *respectively. Assume that assumptions* (B1)–(B10) *are satisfied. Take approximations $w_{(n)}(t)$ of the Wiener process $w(t)$ given by* (5.2.5). *Then, for each $t \in [0, T], 0 < T < \infty$*

$$\lim_{n\to\infty} E[\|u_{(n)}(t) - \widehat{u}(t)\|^2_{H_1}] = 0. \tag{5.4.6}$$

6. Applications

6.1. APPLICATION DERIVING THE RELATION BETWEEN THE ITÔ AND STRATONOVICH INTEGRALS

Firstly, we consider the Stratonovich integral with integrands of delayed argument on a finite time interval. The relation between this integral and the Itô stochastic integral with the same integrand is given in the paper of Dawidowicz and Twardowska [25].

An additional term occurring when the Itô integral is changed to the Stratonovich integral is the same as the correction term in an approximation theorem of the Wong–Zakai-type in [115, 116].

We consider the following stochastic integral equation with delayed argument:

$$\begin{aligned} X^i(t,w) &= X_0^i(w) + \int_0^t b^i(s, X_s(\cdot,w))\,\mathrm{d}s + \\ &\quad + \sum_{p=1}^{m} \int_0^t \sigma^{ip}(s, X_s(\cdot,w))\,\mathrm{d}w^p(s) \end{aligned} \tag{6.1.1}$$

for $i = 1, \ldots, d$. The second integral in (6.1.1) is the Itô integral. Besides (6.1.1), we consider the equation

$$\begin{aligned} X^i(t,w) &= X_0^i(w) + \int_0^t b^i(s, X_s(\cdot,w))\,\mathrm{d}s + \\ &\quad + \sum_{p=1}^{m} \int_0^t \sigma^{ip}(s, X_s(\cdot,w))\,\mathrm{d}w^p(s) + \\ &\quad + \frac{1}{2}\sum_{p=1}^{m}\sum_{j=1}^{d} \int_0^t \widetilde{D}_j\sigma^{ip}(s, X_s(\cdot,w))\sigma^{jp}(s, X_s(\cdot,w))\,\mathrm{d}s \end{aligned} \tag{6.1.2}$$

for $i = 1, \ldots, d$.

We assume $(\tilde{\mathrm{A}}1)$–$(\tilde{\mathrm{A}}5)$ for finite $r \in \mathbb{R}_+$ from Section 5.1 and also

(A*6) there exists a constant $M > 0$ such that for every $s, t \in I$ and $\varphi \in \mathcal{C}_-$ we have

$$|\sigma(s,\varphi) - \sigma(t,\varphi)|_L \leqslant M|s-t|,$$

(A*7) the process X_0 satisfies $E[|X_0(\theta)|^4] < \infty$ for every $\theta \in [-r, 0]$, where

$$|X_0(\theta,w)| = \sum_{i=1}^{d} |X_0^i(\theta,w)|.$$

We have

DEFINITION 6.1.1. Given a function $f\colon\ [0,T]\times\mathcal{C}_- \to \mathbb{R}$, consider the following limit:

$$\begin{aligned} &\lim_{n\to\infty} S_n \\ &\quad = \lim_{n\to\infty} \sum_{i=1}^{n} [w(t_i^n) - w(t_{i-1}^n)] f\Big(\tfrac{1}{2}(t_i^n + t_{i-1}^n), \tfrac{1}{2}(X_{t_i^n} + X_{t_{i-1}^n})\Big), \end{aligned} \tag{6.1.3}$$

where $w(t)$ is the one-dimensional Wiener process. The limit is taken in the mean-square sense and $0 = t_0^n < t_1^n < \cdots < t_n^n = T$ is a partition of the interval $[0, t]$. We assume that the sequence of partitions is normal, that is, $\max(t_{i-1}^n - t_i^n) \to 0$ as $n \to \infty$. If this limit exists and does not depend on the choice of partition, it is called the *Stratonovich integral* and is denoted by (S) $\int_0^T f(t, X_t)\,\mathrm{d}w(t)$.

We recall the definition of the Itô integral:

$$\text{(I)} \quad \int_0^T f(t, X_t)\,\mathrm{d}w(t) = \lim_{n\to\infty} I_n = \lim_{n\to\infty} \sum_{i=1}^{n} [w(t_i^n) - w(t_{i-1}^n)] f(t_{i-1}^n, X_{t_{i-1}^n}), \tag{6.1.4}$$

with the same assumptions as in Definition 6.1.1.

We proved in [25] the following

THEOREM 6.1.1 (Dawidowicz and Twardowska [25]). *Let $f\colon [0,T]\times\mathcal{C}_- \to \mathbb{R}$ be continuous in the first variable and differentiable in both variables with bounded derivative $\partial f/\partial t$, $t \in [0,T]$, and with a continuous Fréchet derivative in the second variable. Moreover, let*

$$\int_0^T E[|f(s, X_s)|]^2\,\mathrm{d}s < \infty, \qquad E\left[\int_0^T \left|\frac{\mathrm{d}}{\mathrm{d}s} f(s, X_s)\right|^2 \mathrm{d}s\right] < \infty, \tag{6.1.5}$$

where X is the strong solution to stochastic differential Equation (6.1.1) *and $w(t)$ is the m-dimensional Wiener process. Moreover, assume that conditions* ($\tilde{\mathrm{A}}$1)–($\tilde{\mathrm{A}}$5) *and* (A*6), (A*7) *are satisfied. Then there exists the Stratonovich integral* (6.1.3) *and*

$$\begin{aligned} \text{(S)} \sum_{p=1}^{m} \int_0^T f(t, X_t)\,\mathrm{d}w^p(t) &= \text{(I)} \sum_{p=1}^{m} \int_0^T f(t, X_t)\,\mathrm{d}w^p(t) + \\ &\quad + \frac{1}{2} \sum_{p=1}^{m} \sum_{j=1}^{d} \int_0^T \widetilde{D}_j f(s, X_s) \sigma^{jp}(s, X_s)\,\mathrm{d}s. \end{aligned} \tag{6.1.6}$$

Remark 6.1.1. In particular, if $f(t, X_t) = \sigma(X_t)$, then the correction term in (6.1.6) has the form

$$\frac{1}{2} \sum_{p=1}^{m} \sum_{j=1}^{d} \int_0^T \widetilde{D}_j \sigma^{ip}(s, X_s) \sigma^{jp}(s, X_s)\,\mathrm{d}s.$$

It is the same correction term as in an approximation theorem of the Wong–Zakai-type in Section 5.1.

Now we mention that a similar result for the model described in Section 5.2, where semilinear stochastic differential equations in Hilbert spaces are considered, was proved by the author in [117]. The correction term introduced in (4.1.10) occurring in an approximation theorem of the Wong–Zakai-type is the same as in the course of the transition from the Itô to Stratonovich integral in [117], which should have been expected, like in the one-dimensional case.

6.2. APPLICATION TO SUPPORT THEOREMS

The second important application of Wong–Zakai theorems is that they constitute an important part of the proofs of the theorems on the support of measures connected with the solutions of the appropriate stochastic differential equations. They are also an indication for computations of Itô formulas for the stochastic delay differential equations, as well as in the more general case of stochastic semilinear evolution equations in Hilbert spaces.

Suppose now that our aim is to describe the topological support $S_1 = \sigma(P_{\hat{z}})$ of the measure $P_{\hat{z}}$ (that is, of the probability law of the solution $\hat{z}(t)$ to (5.2.3)) in the space G of H_1-valued continuous functions (of the trajectories of these solutions). We know that this support is the intersection of all closed sets $A \subseteq G$ such that $P(\hat{z} \in A) = 1$. Let

$$S = \{v \in C^1([0,T], H_1)\colon\ v(0) = 0\} \tag{6.2.1}$$

denote the set of smooth disturbances of the Wiener process $w(t)$. We introduce the following deterministic equation for $v \in S$ (see the notation of Section 5.2):

$$\begin{aligned} u^*(t) &= S(t)z_0 + \int_0^t S(t-s)\mathcal{C}(u^*(s))\,\mathrm{d}s + \\ &\quad + \int_0^t S(t-s)\mathcal{B}(u^*(s))v(s)\,\mathrm{d}s, \quad u^*(0) = z_0. \end{aligned} \tag{6.2.2}$$

Define

$$S_2 = S_{u^*} = \mathrm{cl}\{u^*\colon\ u^* \text{ is the solution to (6.2.2) for some } v \in S\}. \tag{6.2.3}$$

The support theorem ([118]) reads as follows:

THEOREM 6.2.1. *Consider Equations* (5.2.3), (5.2.7$_n$). *Suppose that assumptions* (A1)–(A5) *from Section* 5.2 *are satisfied. Then* $S_1 = S_2$.

Similarly, we obtain the support theorem for the model (5.3.5). Let S_1 denote the topological support of the measure $P_{\hat{u}}$ in the space $G = C(0,T;H_1) \cap$

$L^p(0,T;V)$. The set S is given by (6.2.1) and S_2 is defined in a way similar to (6.2.3). Then $S_1 = S_2$.

6.3. APPLICATIONS TO ENGINEERING AND PHYSICAL SCIENCES

Important questions arise in engineering and physical sciences when stochastic differential equations are used in applications and the appropriate version (Itô or Stratonovich) of an equation should be chosen. The interpretation of stochastic integrals and stochastic differential equations results from the property of the sample paths of a Wiener process that they are not differentiable or even of bounded variation. This is not a problem if we remember that such equations arise from adding to ordinary differential equations random fluctuations described by a Gaussian white noise that generalizes the derivative of a Wiener process. We know that the white noise processes are meant in engineering to be idealizations of real coloured noise processes with arbitrarily small autocorrelations at different time instants. Our answer as to which kind of equation should be chosen depends on the intention of exactly how the white noise processes are to approximate and how a stochastic differential equation approximates the real events.

Physically realizable processes are often smooth with a small degree of autocorrelation. Consider a random differential equation that is an ordinary differential equation in each of its sample paths:

$$\mathrm{d}X^{(n)}(t) = m(X^{(n)}(t),t)\,\mathrm{d}t + \sigma(X^{(n)}(t),t)\,\mathrm{d}R^{(n)}(t), \tag{6.3.1^n}$$

where $R^{(n)}(t)$ is a close process to the Wiener process. This equation in its integral form involves Riemann–Stieltjes integrals which can be computed by the methods of classical calculus. Since these methods are also valid for the Stratonovich calculus, a Stratonovich interpretation of equations is more appropriate for the limit equation obtained by replacing the Wiener process $w(t)$ by the real process $R^{(n)}(t)$. We have seen, e.g., in (1.8) the importance of knowing the Wong–Zakai correction terms in numerical schemes.

Theorems of the Wong–Zakai-type establish the same result under some restrictive assumptions on the coeffitients m, σ and smooth approximation noise process $R^{(n)}(t)$, therefore we see that the Stratonovich interpretation is more appropriate when the white noise is used to idealize a real noise process. The Itô interpretation is more useful in other cases where, for example, we need to have the martingale properties of solutions and the appearance of the correction term in the limit equation is not of great discomfort.

There are many problems in engineering and physics that can be modelled by an ordinary differential equation with random fluctuations which have been deduced from phenomological or physical laws, e.g. biological systems in genetics and population growth, environmental sciences, and so on.

References

1. Akhiezer, N. I.: *Approximation Theory*, PWN, Warsaw, 1957.
2. Aquistapace, P. and Terreni, B.: An approach to Itô linear equations in Hilbert spaces by approximation of white noise with coloured noise, *Stochastic Anal. Appl.* **2** (1984), 131–186.
3. Arnold, L.: *Stochastic Differential Equations: Theory and Applications*, Wiley, New York, 1974.
4. Artem'ev, S. S.: Numerical solution of stochastic differential equations, in A. K'nchev (ed.), *Proc. Third Conf. Differential Equations and Applications*, Ruse 1985, VTU, 1985, pp. 15–18 (in Russian).
5. Bally, V.: Approximation for the solutions of stochastic differential equations, III. Jointly weak convergence, *Stochastics Stochastics Rep.* **30** (1990), 171–191.
6. Banks, H. T. and Kappel, F.: Spline approximation for functional differential equations, *J. Differential Equations* **34** (1979), 496–522.
7. Bell, D. R. and Mohammed, S. E. A.: On the solution of stochastic ordinary differential equations via small delays, *Stochastics Stochastics Rep.* **28** (1989), 293–299.
8. Bellman, R. E. and Roth, R. S.: *Methods in Approximation*, D. Reidel, Dordrecht, 1986.
9. Bensoussan, A.: A model of stochastic differential equations applicable to Navier–Stokes equation in dimension, II, in: E. Mayer-Wolf, E. Merzbach and Schwartz (eds), *Stochastic Analysis, Liber Amicorum for Moshe Zakai*, Academic Press, 1991, pp. 51–73.
10. Bensoussan, A.: Some existence results for stochastic partial differential equations, in: *Pitman Research Notes in Math.* 268, Longman, London, 1992, pp. 37–53.
11. Bensoussan, A. and Temam, R.: Equations aux *d* érivées partielles stochastiques nonlinéaires (1), *Israel J. Math.* **11** (1972), 95–129.
12. Bensoussan, A. and Temam, R.: Equations stochastiques *du* type Navier–Stokes, *J. Functional Anal.* **13** (1973), 195–222.
13. Blum, E. K.: *Numerical Analysis and Computation Theory in Practice*, Addison-Wesley, Reading, Mass., 1972.
14. Brzeźniak, Z., Capiński, M., and Flandoli, F.: A convergence result for stochastic partial differential equations, *Stochastics* **24** (1988), 423–445.
15. Butzer, P. L. and Berens, H.: *Semigroups of Operators and Approximations*, Springer, Berlin, 1967.
16. Capiński, M.: A note on uniqueness of stochastic Navier–Stokes equations, *Univ. Iagel. Acta Math.* **30** (1993), 219–228.
17. Capiński, M. and Cutland, N.: Stochastic Navier–Stokes equations, *Acta Appl. Math.* **25** (1991), 59–85.
18. Chen, H. F. and Gao, A. J.: Robustness analysis for stochastic approximation algorithms, *Stochastics Stochastics Rep.* **26** (1989), 3–20.
19. Chojnowska-Michalik, A.: Representation theorem for general stochastic delay equations, *Bull. Acad. Polon. Sci.* **26**(7) (1978), 635–642.
20. Clark, J. M. C.: An efficient approximation for a class of stochastic differential equations, in: W. Fleming and L. G. Gorostiza (eds), *Advances in Filtering and Optimal Stochastic Control*, Proc. IFIP Working Conf., Cocoyoc, Mexico, 1982, Lecture Notes in Control and Inform. Sci. 42, Springer, Berlin, 1982, pp. 69–78.
21. Clark, J. M. C. and Cameron, R. J.: The maximum rate of convergence of discrete approximations for stochastic differential equations, in: B. Grigelionis (ed.), *Stochastic Differential Systems-Filtering and Control*, Proc. IFIP Working Conf., Vilnius, Lithuania, 1978, Lecture Notes in Control and Inform. Sci. 25, Springer, Berlin, 1980, pp. 162–171.
22. Curtain, R. F. and Pritchard, A. J.: *Infinite Dimensional Linear Systems Theory*, Springer, Berlin, 1978.
23. Da Prato, G.: Stochastic differential equations with noncontinuous coefficients in Hilbert space, *Rend. Sem. Math. Univers. Politecn. Torino*, Numero Speciale (1982), 73–85.
24. Da Prato, G. and Zabczyk, J.: *Stochastic Equations in Infinite Dimensions*, Cambridge Univ. Press, 1991.
25. Dawidowicz, A. L. and Twardowska, K.: On the Stratonovich and Itô integrals with integrands of delay argument, *Demonstratio Math.* **28**(2) (1995), 465–478.

26. Doss, H.: Liens entre équations différentielles stochastiques et ordinaires, *Ann. Inst. H. Poincaré XIII* **2** (1977), 99–125.
27. Dziadyk, W. K.: *Approximation Methods of Solving Differential and Integral Equations*, Naukova Dumka, Kiev, 1988 (in Russian).
28. Ferreyra, G.: A Wong–Zakai type theorem for certain discontinuous semimartingales, *J. Theoret. Probab.* **2**(3) (1989), 313–323.
29. Funaro, D.: *Polynomial Approximations of Differential Equations*, Springer, Berlin, 1992.
30. Gorostiza, L. G.: Rate of convergence of a approximate solution of stochastic differential equations, *Stochastics* **3** (1980), 267–276.
31. Greiner, A. and Stritttmatter, W.: Numerical integration of stochastic differential equations, *J. Statist. Phys.* **51**(1–2) (1988), 95–108.
32. Gyöngy, I.: On the approximation of stochastic differential equations, *Stochastics* **23** (1988), 331–352.
33. Gyöngy, I.: On the approximation of stochastic partial differential equations, Part I, *Stochastics* **25** (1988), 59–85; Part II, *Stochastics* **26** (1989), 129–164.
34. Gyöngy, I.: The approximation of stochastic parial differential equations and applications in nonlinear filtering, *Comput. Math. Appl.* **19**(1) (1990), 47–63.
35. Gyöngy, I.: The stability of stochastic partial differential equations and applications. Theorems on supports, Lecture Notes in Math. 1390, Springer, Berlin, 1989, pp. 91–118.
36. Hale, J.: *Theory of Functional Differential Equations*, Springer, Berlin, 1977.
37. Heble, H. P.: *Approximation Problems in Analysis and Probability*, Elsevier, Amsterdam, 1989.
38. Herman, E. A.: *Convergence, Approximation and Differential Equations*, Wiley, New York, 1986.
39. Ikeda, N., Nakao, S., and Yamato, Y.: A class of approximations of Brownian motion, *Publ. RIMS Kyoto Univ.* **13** (1977), 285–300.
40. Ikeda, N. and Watanabe, S.: *Stochastic Differential Equations and Diffusion Processes*, North-Holland, Amsterdam, 1981.
41. Itô, K. and Nisio, M.: On stationary solutions of a stochastic differential equation, *J. Math. Kyoto Univ.* **4**(1) (1964), 1–75.
42. Jakubowski, A., Mémin, J., and Pages, G.: Convergence en loi des suites d'intégrales stochastiques sur l'espace D^1 de Skorokhod, *Probab. Theory Related Fields* **81** (1989), 111–137.
43. Janssen, R.: Difference methods for stochastic differential equations with discontinuous coefficient, *Stochastics* **13** (1984), 199–212.
44. Janssen, R.: Discretization of the Wiener process in difference methods for stochastic differential equations, *Stochastic Process. Appl.* **18** (1984), 361–369.
45. Jerome, J. W.: *Approximation of Nonlinear Evolution Systems*, Academic Press, New York, 1983.
46. Kawabata, S.: On the successive approximation of solutions of stochastic differential equations, *Stochastics Stochastics Rep.* **30** (1990), 69–84.
47. Klauder, J. R. and Petersen, W. P.: Numerical interpretation of multiplicative noise of stochastic differential equations, *SIAM J. Numer. Anal.* **22**(6) (1985), 1153–1166.
48. Kloeden, P. and Platen, E.: *Numerical Solution of Stochastic Differential Equations*, Springer, Berlin, 1992.
49. Koneczny, F.: On the Wong–Zakai approximation of stochastic differential equations, *J. Mult. Anal.* **13** (1983), 605–611.
50. Krylov, N. U. and Rozovskiĭ, B. L.: On stochastic evolution equations, *Itogi Nauki i Techniki, Teor. Veroyatnost.* **14** (1979), 71–146 (in Russian).
51. Kurtz, T. G., Pardoux, E., and Protter, P.: Stratonovich stochastic differential equations driven by general semimartingales, Technical Report No. 91–22, Dept. of Statistics, Purdue University, March 1992.
52. Kurtz, T. G. and Protter, P.: Characterizing the weak convergence of stochastic integrals, in *Stochastic Analysis*, Cambridge Univ. Press, Cambridge, 1991.
53. Kurtz, T. G. and Protter, P.: Weak limit theorems for stochastic integrals and stochastic differential equations, *Ann. Probab.* **19** (1991), 1035–1070.

54. Kurtz, T. G. and Protter, P.: Wong–Zakai corrections, random evolutions and simulation schemes for SDE's. Stochastic analysis, in *Proc. Conf. Honor Moshe Zakai 65th Birthday*, Haifa, Israel, 1991, pp. 331–346.
55. Kushner, H. J.: Jump-diffusion approximations for ordinary differential equations with wide-band random right hand sides, *SIAM J. Control Optim.* **17** (1979), 729–744.
56. Kushner, H. J.: *Probability Methods for Approximations in Stochastic Control and for Elliptic Equations*, Academic Press, New York, 1977.
57. Kushner, H. J. and Dupuis, P. G.: *Numerical Methods for Stochastic Control Problems in Continuous Time*, Springer, Berlin, 1993.
58. Kushner, H. J. and Yin, G.: Stochastic approximation algorithms for parallel and distributed processing, *Stochastics* **22** (1987), 219–250.
59. Lions, J. L.: *Quelques méthodes de résolution des problèmes aux limites non linéaires*, Dunod, Paris, 1969.
60. Liptser, R. and Shiryaev, A.: *Statistics of Random Processes*, Springer, Berlin, 1977.
61. Mackevičius, V.: On polygonal approximation of Brownian motion in stochastic integrals, *Stochastics* **13** (1984), 167–175.
62. Mackevičius, V.: On the support of a solution of stochastic differential equation, *Liet. Mat. Rinkinys* **26**(1) (1986), 91–98.
63. Mackevičius, V.: SP-stability of symmetric stochastic differential equations, *Liet. Mat. Rinkinys* **25**(4) (1985), 72–84.
64. Mackevičius, V.: Symmetric stochastic integrals and their approximations, *Stochastics* **8** (1982), 121–138.
65. Mackevičius, V. and Žibaitis, B.: Gaussian approximations of Brownian motion in stochastic integral, *Liet. Mat. Rinkinys*, **33**(4) (1993), 508–526.
66. Mao, X.: Approximate solutions for a class of delay stochastic differential equations, *Stochastics Stochastics Rep.* **35** (1991), 111–123.
67. Marcus, S.: Modeling and approximation of stochastic differential equations driven by seminartingales, *Stochastics* **4** (1981), 223–245.
68. Maruyama, G.: Continuous Markov processes and stochastic equations, *Rend. Cir. Mat. Palermo* **4** (1955), 48–90.
69. McShane, E. J.: *Stochastic Calculus and Stochastic Models*, Academic Press, New York, 1974.
70. McShane, E. J.: Stochastic differential equations and models of random processes, in *Proc. 6th Berkeley Sympos., Math. Statist. Probab. 3*, University of California Press, Berkeley, 1972, pp. 263–294.
71. Mémin, J. and Słomiński, L.: Condition UT et stabilité en loi des solutions *d'* équations différentielles stochastiques, in *Séminaire de Probabilités XXV*, Lecture Notes in Math. 1485, Springer, Berlin, 1991, pp. 162–177.
72. Métivier, M.: *Semimartingales*, Walter de Gruyter, Berlin, 1982.
73. Millet, A. and Sanz-Solé, M.: A simple proof of the support theorem for diffusion process, Lecture Notes in Math. 1583, Springer, Berlin, 1994, pp. 36–48.
74. Milshtein, G. N.: Approximate integration of stochastic differential equations, *Theor. Veroyatnost. i Primenen.* **19** (1974), 583–588 (in Russian).
75. Milshtein, G. N.: Weak approximation of solutions of systems of stochastic differential equations, *Theor. Veroyatnost. i Primenen.* **30** (1985), 706–721 (in Russian).
76. Mohammed, S. E. A.: *Stochastic Functional Differential Equations*, Pitman, Marshfield, 1984.
77. Mohammed, S. E. A.: *Retarded Functional Differential Equations*, Pitman, London, 1978.
78. Nakao, S. and Yamato, Y.: Approximation theorem on stochastic differential equations, in *Proc. Internat. Sympos. SDE Kyoto, 1976*, Tokyo, 1978, pp. 283–296.
79. Newton, N. J.: An asymptotically efficient difference formula for solving stochastic differential equations, *Stochastics* **19** (1986), 175–206.
80. Newton, N. J.: An efficient approximation for stochastic differential equations on the partition of symmetrical first passage times, *Stochastics* **29** (1990), 227–258.
81. Onicescu, O. and Istratescu, V. I.: Approximations theory for random functions, *Rend. Matem.* **8** (1975), 65–81.

82. Pardoux, E.: Equations aux dérivées partielles stochastiques non linéaires monotones. Etude de solutions fortes de type Itô, Thèse Doct. Sci. Math. Univ. Paris Sud, 1975.
83. Pardoux, E.: Stochastic partial differential equations and filtering of diffusion processes, *Stochastics* **3** (1979), 127–137.
84. Pardoux, E. and Talay, D.: Discretization and simulation of stochastic differential equations, *Acta Appl. Math.* **3** (1985), 23–47.
85. Pazy, A.: *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer, Berlin, 1983.
86. Pettersson, R.: Approximations for stochastic differential equations with reflecting convex boundaries, submitted to *Stochastic Proc. Appl.*
87. Picard, J.: Approximation of stochastic differential equations and application of the stochastic calculus of variations to the rate of convergence, Lecture Notes in Math. 1316, Springer, Berlin, 1988, pp. 267–287.
88. Platen, E.: An approximation methods for a class of Itô process, *Liet. Mat. Rinkinys* **21**(1) (1981), 121–133.
89. Platen, E.: Approximation of the first exit times of diffusions and approximate solutions of parabolic equations, *Math. Nachr.* **111** (1983), 127–146.
90. Platen, E.: A Taylor formula for semimartingales solving a stochastic equation, in *Lecture Notes in Control and Inform. Sci. 36*, Springer, Berlin, 1981, pp. 157–174.
91. Protter, P.: Approximations of solutions of stochastic differential equations driven by semimartingales, *Ann. Probab.* **13**(3) (1985), 716–743.
92. Protter, P.: Stochastic differential equations with jump reflection at the boundary, *Stochastics* **3** (1980), 193–201.
93. Rozovskiĭ, B. L.: *Stochastic Evolution Systems. Linear Theory and Applications to Non-Linear Filtering*, Kluwer, Dordrecht, 1990.
94. Shimizu, A.: Approximate solutions of stochastic differential equations, *Bull. Nagoya Inst. Tech.* **36** (1984), 105–108.
95. Słomiński, L.: On approximation of solutions of multidimensional SDE's with reflecting boundary conditions, *Stochastic Proc. Appl.* **50** (1994), 197–219.
96. Słomiński, L.: Stability of strong solutions of stochastic differential equations, *Stochastic Proc. Appl.* **31** (1989), 173–202.
97. Sobczyk, K.: *Stochastic Differential Equations with Applications to Physics and Engineering*, Kluwer, Dordrecht, 1991.
98. Stratonovich, R. L.: A new representation for stochastic integrals and equations, *SIAM J. Control Optim.* 4(2) (1966), 362–371.
99. Strook, D. W. and Varadhan, S. R. S.: On the support of diffusion processes with applications to the strong maximum principle, in: *Proc. 6th Berkeley Sympos., Math. Statist. Probab. 3*, University of California Press, Berkeley, 1972, pp. 333–359.
100. Sussmann, H.: On the gap between deterministic and stochastic ordinary differential equations, *Ann. Probab.* **6**(1) (1978), 19–41.
101. Talay, D.: Approximation of upper Lyapunov exponents of bilinear stochastic differential systems, submitted to *INRIA*, Report 965 (1989).
102. Talay, D.: Efficient numerical schemes for the approximation of expectations of functionals of SDE, in: H. Korezlioglu, G. Mazziotto and J. Szpirglas (eds), *Filtering and Control of Random Process*, Proc. ENST–CNET Colloquium, Paris, 1983, Lecture Notes in Control and Inform. Sci. 61, Springer, Berlin, 1984, pp. 294–313.
103. Talay, D.: Résolution trajectorielle et analyse numérique des équations différentielles stochastiques, *Stochastics* **9** (1983), 275–306.
104. Talay, D.: Second-order discretization schemes of stochastic differential systems for the computation of the invariant law, *Stochastics Stochastics Rep.* **29** (1990), 13–36.
105. Tanabe, H.: *Equations of Evolution*, Monographs and Studies in Math. 6, Pitman, London, 1979.
106. Trotter, H. F.: Approximation of semi-groups of operators, *Pacific J. Math.* **8** (1958), 887–919.
107. Tudor, C.: On stochastic evolution equations driven by continuous semimartingales, *Stochastics* **24** (1988), 179–195.

108. Tudor, C.: On the successive approximations of solutions of delay stochastic evolution equations, *Ann. Univ. Bucureşti Math.* **34** (1985), 70–86.
109. Tudor, C.: On weak solutions of Volterra equations, *Boll. Univ. Math. Italy B* **1**(7) (1987), 1033–1054.
110. Tudor, C.: Some properties of mild solutions of delay stochastic evolution equations, *Stochastics* **17** (1986), 1–18.
111. Tudor, M.: Approximation schemes for two-parametric stochastic equations, *Probab. Math. Statist.* **13**(2) (1992), 177–189.
112. Twardowska, K.: An approximation theorem of Wong–Zakai type for nonlinear stochastic partial differential equations, Preprint No. 53, Warsaw Technical University, 1993, 1–25, *Stochastic Anal. Appl.* **13**(5) (1995).
113. Twardowska, K.: An approximation theorem of Wong–Zakai type for stochastic Navier–Stokes equations, *Rend. Sem. Math. Univ. Padova* **96** (1996).
114. Twardowska, K.: An extension of Wong–Zakai theorem for stochastic evolution equations in Hilbert spaces, *Stochastic Anal. Appl.* **10**(4) (1992), 471–500.
115. Twardowska, K.: Approximation theorems of Wong–Zakai type for stochastic differential equations in infinite dimensions, *Dissertationes Math.* **325** (1993), 1–54.
116. Twardowska, K.: On the approximation theorem of the Wong–Zakai type for the functional stochastic differential equations, *Probab. Math. Statist.* **12**(2) (1991), 319–334.
117. Twardowska, K.: On the relation between the Itô and Stratonovich integrals in Hilbert spaces, submitted to *Statist. Probab. Lett.* (1992).
118. Twardowska, K.: Wong–Zakai approximations for stochastic differential equations and support theorems, *Proc. Internat. Workshop on Stochastic Partial Diff. Equations*, Center for Stochastic Process, Chapel Hill, U.S.A., 1994, pp. 97–100.
119. Vinter, R. B.: A representation of solutions to stochastic delay equations, Imperial College Report, 1975.
120. Vinter, R. B.: 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.
121. Wagner, E.: Unbiased Monte-Carlo estimators for functionals of weak solutions of stochastic differential equations, *Stochastics Stochastics Rep.* **28** (1989), 1–20.
122. Wong, E.: *Stochastic Processes in Information and Dynamical Systems*, McGraw-Hill, 1971.
123. Wong, E. and Zakai, M.: On the convergence of ordinary integrals to stochastic integrals, *Ann. Math. Statist.* **36** (1965), 1560–1564.
124. Wong, E. and Zakai, M.: Riemann–Stieltjes approximations of stochastic integrals, *Z. Warsch. verw. Gebiete* **12** (1969), 87–97.
125. Zabczyk, J.: On decomposition of generators, *SIAM J. Control Optim.* **16**(4) (1978), 523–534.