Теорія Ймовір. та Матем. Статист.
Вип. 69, 2003

Theor. Probability and Math. Statist.
No. 69, 2004, Pages 187–199
S 0094-9000(05)00625-3
Article electronically published on February 9, 2005

RATE OF CONVERGENCE OF DISCRETE APPROXIMATE SOLUTIONS OF STOCHASTIC DIFFERENTIAL EQUATIONS IN A HILBERT SPACE

UDC 519.21

G. SHEVCHENKO

Abstract. We consider discrete-time approximations for stochastic differential equations in a Hilbert space. The rate of convergence of approximations is established for equations with Lipschitz continuous coefficients and for semilinear evolution type equations with an unbounded drift. As an auxiliary result, the rate of convergence of approximations is obtained for Itô–Volterra equations in a Hilbert space.

Introduction

There are several applications of numerical solutions of stochastic differential equations. A classical example of such an application is numerical solution of partial differential equations. For example, a solution of the heat equation

$$\begin{cases} \frac{\partial u}{\partial t}(t,x) = \frac{1}{2}(b(t,x)\nabla_x, \nabla_x)u(t,x) + (a(t,x),\nabla_x)u(t,x), \quad t>0,\ x\in\mathbb{R}^d, \\ u(0,x) = \varphi(x), \end{cases} \tag{0.1}$$

can be represented as

$$u(t,x) = \mathsf{E}_x\,\varphi(\xi(t))$$

where $\xi(t)$ is a solution of the associated stochastic differential equation

$$\xi(t) = x + \int_0^t a(s,\xi(s))\,ds + \int_0^t b^{1/2}(s,\xi(s))\,dW(s).$$

One can approximate ξ first, and then get an approximate solution of (0.1) using Monte-Carlo simulations. The advantage of this approach as compared with deterministic methods of numerical solution of (0.1) is that it requires less computations for large dimensions d. Another example of applications is the modelling of processes in finance, insurance, etc.

The first monograph on numerical solution of stochastic differential equations is the book [9] by G. N. Milstein. The current literature on this topic is rather rich; see the books [4, 7] and the paper [11] for surveys of results and other examples of applications. There are several papers devoted to numerical solution of stochastic differential equations involving semimartingles and Lévy processes; see [5, 10] and references therein. Only a few papers are devoted to approximation of stochastic differential equations for the case of infinite dimension, namely to the evolution type equations and Itô–Volterra equations. The papers [1, 8] contain results on the strong convergence of approximations; the rate

2000 *Mathematics Subject Classification.* Primary 60H35; Secondary 60H10, 60H20, 65C30.
Key words and phrases. Stochastic differential equations in a Hilbert space, discrete-time approximations, equations of the Itô–Volterra type.

of convergence, however, is not studied there. The weak convergence of approximations of Itô–Volterra equations is proved in [6], and a theorem on the relation between local and global rates of convergence for Itô–Volterra equation is obtained in [12].

In this paper we consider approximation of stochastic differential equations in a Hilbert space. Section 1 is devoted to the simplest case of an equation with Lipschitz continuous coefficients. Semilinear evolution type equations with an unbounded drift are considered in Section 2. Conditions for the existence of a solution and approximations (of a mild solution) are given. Section 3 contains results on the convergence of approximations for Itô–Volterra type equations. These results are applied to semilinear equations discussed in Section 2.

1. Equations with continuous coefficients

This section is somewhat independent of the rest of the paper. Here we prove that the classical result about the rate of convergence of approximation for finite-dimensional equations with Lipschitz continuous coefficients remains true in the case of infinite dimension. This result follows from the theorems of the next section; nevertheless we provide an independent proof of it for the sake of completeness. On the other hand, the argument in this simple case exhibits how it will work in the general case.

Let X and H be separable Hilbert spaces, $(\Omega, \mathcal{F}, \mathsf{P})$ a probability space with filtration $(\mathcal{F}_t, t \in [0,T])$, and $W(t)$ an $\mathcal{F}_t$-adapted Q-Wiener process in H, that is, the covariance operator of $W(t)$ equals Q where $\operatorname{tr} Q < \infty$.

Consider the following equation on $[0,T]$:

$$X(t) = X_0 + \int_0^t a(s, X(s))\, ds + \int_0^t b(s, X(s))\, dW(s) \tag{1.1}$$

where a and b are measurable maps from $[0,T] \times X$ to X and $\mathcal{L}(H,X)$, respectively ($\mathcal{L}(H,X)$ is the space of linear operators), and X_0 is an $\mathcal{F}_0$-measurable random variable. The solution of this equation exists and is unique if

$$\mathsf{E}\, \|X_0\|_X^{2p} < \infty, \tag{1.2a}$$

$$\|a(t,x)\|_X + \|b(t,x)\|_{\mathcal{L}(H,X)} \le C(1 + \|x\|_X), \tag{1.2b}$$

$$\|a(t,x) - a(t,y)\|_X + \|b(t,x) - b(t,y)\|_{\mathcal{L}(H,X)} \le C\, \|x - y\|_X\,, \tag{1.2c}$$

$$\|a(t,x) - a(s,x)\|_X + \|b(t,x) - b(s,x)\|_{\mathcal{L}(H,X)} \le C(1 + \|x\|_X)\, |t-s|^{1/2} \tag{1.2d}$$

for some $p > 0$ (see, for example, [2, Theorem 7.2.1]) where $x, y \in X$, $t, s \in [0,T]$, and C is a constant independent of x, y, s, and t. In what follows we write $\|\cdot\|$ instead of $\|\cdot\|_X$ and $\mathcal{L}$ instead of $\mathcal{L}(H,X)$. All constants whose exact values are not essential will be denoted by C.

If conditions (1.2) hold, then the solution $X(t)$ of (1.1) is such that

$$\begin{gathered} \sup_{0 \le t \le T} \mathsf{E}\, \|X(t)\|^{2p} < \infty, \\ \mathsf{E}\, \|X(t) - X(s)\|^{2p} \le C\, |t-s|^p\,, \qquad s, t \in [0,T] \end{gathered} \tag{1.3}$$

(see Theorem 7.2.1 in [2]).

The Euler approximations for the solution of (1.1) are constructed in the following way. For a given number $N \in \mathbb{N}$ set $\delta = T/N$ and let $\tau_n = n\delta$, $n = 0, \dots, N$, be a uniform partition of $[0,T]$. The initial value X_0 is approximated by an $\mathcal{F}_0$-measurable random variable Y_0^δ, and then the approximations are defined recursively at the nodes of the partition, namely

$$Y_{n+1}^\delta = Y_n^\delta + a\left(\tau_n, Y_n^\delta\right) \delta + b\left(\tau_n, Y_n^\delta\right) (W(\tau_{n+1}) - W(\tau_n)), \qquad n \ge 0.$$

To approximate the values of the solution between the nodes τ_n, we put

$$Y^\delta(t) = Y^\delta_{n_t} + \int_{\tau_{n_t}}^t a\left(\tau_{n_t}, Y^\delta_{n_t}\right)\, ds + \int_{\tau_{n_t}}^t b\left(\tau_{n_t}, Y^\delta_{n_t}\right)\, dW(s)$$

where

$$n_s = \max\left\{n \colon \tau_n < s\right\}.$$

This equation can be rewritten in a more convenient form:

$$Y^\delta(t) = Y^\delta_0 + \int_0^t a\left(\tau_{n_s}, Y^\delta(\tau_{n_s})\right)\, ds + \int_0^t b\left(\tau_{n_s}, Y^\delta(\tau_{n_s})\right)\, dW(s). \tag{1.4}$$

The piecewise linear approximation seems to be more natural; however it also has some disadvantages. First, it is not easy to compute it; second, it is anticipating with respect to $\mathcal{F}_t$, and, third, the piecewise linear approximation is the worse if we measure the quality of an approximation in the sense of minimum of

$$\mathsf{E} \sup_{[0,T]} \left\| X(t) - Y^\delta(t) \right\|.$$

Note, however, that the quality of this approximation is the same as that of the others if we measure the quality in the sense of the minimum of

$$\mathsf{E} \left\| X(t) - Y^\delta(t) \right\|^2$$

at a given point t.

Theorem 1.1. *Assume that conditions* (1.2) *hold and*

$$\mathsf{E} \left\| X_0 - Y^\delta_0 \right\|^{2p} < C\delta^p$$

for some $p \geq 1$. *Then the approximations* $Y^\delta(t)$ *converge to the solution* $X(t)$ *of* (1.1) *and moreover*

$$\mathsf{E} \sup_{0\le t\le T} \left\| X(t) - Y^\delta(t) \right\|^{2p} \leq K_p \delta^p. \tag{1.5}$$

Proof. Put

$$Z(t) = \mathsf{E} \sup_{0\le s\le t} \left\| X(s) - Y^\delta(s) \right\|^{2p}.$$

We have

$$Z(t) \leq C\left(\mathsf{E} \left\| X_0 - Y^\delta_0 \right\|^{2p} + A + B \right)$$

where

$$A = \mathsf{E} \sup_{0\le s\le t} \left\| \int_0^s \left(a(u, X(u)) - a\left(\tau_{n_u}, Y^\delta(\tau_{n_u})\right)\right) du \right\|^{2p}$$

and

$$B = \mathsf{E} \sup_{0\le s\le t} \left\| \int_0^s \left(b(u, X(u)) - b\left(\tau_{n_u}, Y^\delta(\tau_{n_u})\right)\right) dW(u) \right\|^{2p}.$$

We estimate only B; A is estimated similarly. We have

$$B \leq C(B_1 + B_2 + B_3).$$

To estimate the first term we use the Doob–Burkholder inequality and (1.2):

$$B_1 = \mathsf{E} \sup_{0\le s\le t} \left\| \int_0^s \left(b(\tau_{n_u}, X(\tau_{n_u})) - b\left(\tau_{n_u}, Y^\delta(\tau_{n_u})\right)\right) dW(u) \right\|^{2p}$$
$$\le C \int_0^t \mathsf{E} \left\| b(\tau_{n_u}, X(\tau_{n_u})) - b\left(\tau_{n_u}, Y^\delta(\tau_{n_u})\right)\right\|_{\mathcal{L}}^{2p} du$$
$$\le C \int_0^t \mathsf{E} \left\| X(\tau_{n_u}) - Y^\delta(\tau_{n_u})\right\|^{2p} du \le C \int_0^t Z(u)\, du.$$

We proceed in the same way with the second and third terms:

$$B_2 = \mathsf{E} \sup_{0\le s\le t} \left\| \int_0^s \left(b(\tau_{n_u}, X(\tau_{n_u})) - b(\tau_{n_u}, X(u))\right) dW(u) \right\|^{2p}$$
$$\le C \int_0^t \mathsf{E} \left\| X(\tau_{n_u}) - X(u)\right\|^{2p} du \le C\cdot \delta^p,$$
$$B_3 = \mathsf{E} \sup_{0\le s\le t} \left\| \int_0^s \left(b(\tau_{n_u}, X(u)) - b(u, X(u))\right) dW(u) \right\|^{2p}$$
$$\le C \int_0^t \mathsf{E} \left|\tau_{n_u} - u\right|^p \left(1 + \|X(u)\|^{2p}\right) du \le C\cdot \delta^p.$$

Note that $B_i < \infty$ since

$$\sup_{[0,T]} \mathsf{E} \|X(t)\|^{2p} < \infty, \qquad \sup_{[0,T]} \mathsf{E} \left\|Y^\delta(t)\right\|^{2p} < \infty.$$

Applying the Gronwall lemma, we get

$$Z(t) \le C \left(\mathsf{E} \left\| X_0 - Y_0^\delta \right\|^{2p} + \delta^p \right),$$

and Theorem 1.1 follows. □

Remark 1.2. It is possible to construct approximations for equation (1.1) and to obtain a similar result in the case of a Q-Wiener process $W(t)$ such that $\operatorname{tr} Q = \infty$ (in other words, in the case of a cylindric Wiener process). The norm $\|\cdot\|_{\mathcal{L}(H,X)}$ in conditions (1.2) should be replaced with the Hilbert–Schmidt norm $\|\cdot\|_{\mathcal{L}_2(H,X)}$ (in what follows we write $\mathcal{L}_2$ instead of $\mathcal{L}_2(H, X)$). The proof of estimate (1.5) in this case follows the argument of Theorem 1.1.

2. Semilinear equations of evolution type

Consider a typical semilinear stochastic differential equation of evolution type:

$$X(t) = X_0 + \int_0^t \left(A(s)X(s) + a(s, X(s))\right) ds + \int_0^t b(s, X(s))\, dW(s). \tag{2.1}$$

Here $A(s)\colon D \to X$ are closed, linear, and in general unbounded operators with a common domain $D \subset X$; W is a cylindrical Wiener process; a and b are such that

$$\|a(t,x)\| + \|b(t,x)\|_{\mathcal{L}_2} \le C\left(1 + \|x\|\right), \tag{2.2a}$$
$$\|a(t,x) - a(t,y)\| + \|b(t,x) - b(t,y)\|_{\mathcal{L}_2} \le C\,\|x - y\|; \tag{2.2b}$$

X_0 is an $\mathcal{F}_0$-measurable square integrable random variable assuming values in D almost surely. We try to construct approximations for this equation similarly to (1.4):

$$Y^\delta(\tau_n) = Y_0^\delta + \int_0^t \left(A(\tau_{n_s})Y^\delta(\tau_{n_s}) + a\left(\tau_{n_s}, Y^\delta(\tau_{n_s})\right)\right) ds + \int_0^t b(\tau_{n_s}) Y^\delta(\tau_{n_s})\, dW(s). \tag{2.3}$$

This leads to an additional assumption that

$$Y_0^\delta \in C^\infty(A).$$

It turns out that in general such approximations do not converge. Here is a simple example of divergence even in the case of a deterministic equation.

Example 2.1. Let $H = X = L_2(\mathbb{R})$, $a = b \equiv 0$, and let A be the derivative operator. In this case equation (2.1) takes the following simple form:

$$X(t,z) = X_0(x) + \int_0^t \frac{\partial}{\partial z} X(s,z)\,ds.$$

Its solution is easy to write, namely, $X(t,z) = X_0(t+z)$. It is straightforward to check that approximations (2.3) are of the form

$$Y^\delta(\tau_n, z) = Y_n^\delta(z) = \left(I + \delta \frac{d}{dz}\right)^n Y_0^\delta(z).$$

Now let

$$Y_0^\delta(z) = X_0(z) = \begin{cases} \exp\left((1-z^2)^{-1}\right), & z \in (-1,1), \\ 0, & |z| \geq 1. \end{cases}$$

For any $\delta > 0$ and $n \geq 0$ we have $Y^\delta(t,-1) = 0$. Since the function $Y_0^\delta(z)$ is analytic for $z \in (-1,1)$, we have for $t \in (0, 1-z)$ that

$$Y^\delta(t,z) \to X_0(t+z), \qquad \delta \to 0.$$

In particular,

$$Y^\delta\left(1, -1+n^{-1}\right) \to X_0\left(n^{-1}\right) \quad \text{as } \delta \to 0.$$

Thus the pointwise limit $\lim_{\delta\to 0} Y^\delta(z)$ is even discontinuous. At the same time the solution itself is continuous.

The above arguments suggest an idea on how to construct approximations for equation (2.1) in a different way. Suppose that the operators $A(t)$ generate a strongly continuous evolution family of operators $U(t,s)$, $0 \leq s \leq t$. Then the strong solution of (2.1), if it exists, is the so-called mild solution, that is, the one that satisfies the equation

$$X(t) = U(t,0)X_0 + \int_0^t U(t,s)a(s,X(s))\,ds + \int_0^t U(t,s)b(s,X(s))\,dW(s). \tag{2.4}$$

This equation is a particular case of the Itô–Volterra equation considered in the next section. One can consider the Euler type approximations for equation (2.4) by discretizing the time as we did in the preceding section for equation (1.1). These approximations are better in the sense that they converge to the solution of the equation.

Now we state our assumption about the coefficients of (2.1).

(1) The operators $A(s)$ generate a strongly continuous evolution family of operators $\{U(t,s), 0 \leq s \leq t \leq T\}$.
(2) $\ker A(s)$ is independent of s.
(3) The following conditions are satisfied:

$$\left\|A(t)U(t,s)A^{-1}(s)\right\|_{\mathcal{L}} < C,$$
$$\left\|A(t)A^{-1}(s)\right\|_{\mathcal{L}} < C$$

(this is the so-called "hyperbolic" case).
(4) The coefficients a and b satisfy conditions (2.2) and moreover

$$\|A(t)a(t,x)\| \leq C(1+\|x\|),$$
$$\|A(t)b(t,x)\|_{\mathcal{L}_2} \leq C(1+\|x\|).$$

The initial value X_0 is supposed to be an $\mathcal{F}_0$-measurable square integrable random variable that belongs to D almost surely. These assumptions are somewhat milder than those used in [2, Theorem 7.2.3] (where the last condition is that x is bounded). The existence, uniqueness, continuity, and square integrability of the solution of (2.4) can be found in Theorem 7.2.1 in [2] and the remark that follows it.

Proposition 2.2. *Under the assumptions* 1)–4) *on the coefficients of equation* (2.4), *its solution* $X(t)$ *satisfies* (2.1).

Proof. The proof follows the idea of the proof of Theorem 7.2.3 in [2].

Lemma 2.3. *The solution* $X(t)$ *belongs to* D *almost surely for* $t \in [0,T]$.

Proof. It suffices to show that the integrals on the right-hand side of

$$\begin{aligned} A(t)X(t) &= A(t)U(t,0)X_0 \\ &\quad + \int_0^t A(t)U(t,s)a(s,X(s))\,ds + \int_0^t A(t)U(t,s)b(s,X(s))\,dW(s) \end{aligned}$$

exist. Note that

$$A(t)U(t,s)x = A(t)U(t,s)A^{-1}(s)A(s)x.$$

Indeed,

$$A^{-1}(s)A(s)x - x \in \ker A(s) = \ker A(u)$$

and therefore

$$U(t,s)(A^{-1}(s)A(s)x - x) = A^{-1}(s)A(s)x \in \ker A(t).$$

Thus, using the bounds

$$\begin{aligned} \|A(t)U(t,s)a(s,x)\| &\le \left\|A(t)U(t,s)A^{-1}(s)\right\|_{\mathcal{L}} \|A(s)a(s,x)\| \le C\big(1+\|x\|\big), \\ \|A(t)U(t,s)b(s,x)\|_{\mathcal{L}_2} &\le \left\|A(t)U(t,s)A^{-1}(s)\right\|_{\mathcal{L}} \|A(s)b(s,x)\|_{\mathcal{L}_2} \le C\big(1+\|x\|\big), \end{aligned}$$

and taking into account that $\mathsf{E}\,\|X(t)\|^2 \le C$, we prove that the above integrals exist. □

Put

$$Y(t) = \int_0^t a(s,X(s))\,ds + \int_0^t b(s,X(s))\,dW(s), \qquad t \in [0,T].$$

It is easy to see that

$$X(t) = U(t,s)X(s) + \int_s^t U(t,u)\,dY(u).$$

Then

$$\begin{aligned} r(t,s) &:= X(t) - X(s) - (Y(t)-Y(s)) - (U(t,s)-I)X(s) \\ &= \int_s^t (U(t,u)-I)\,dY(u). \end{aligned}$$

Moreover,

$$\begin{aligned} \|\mathsf{E}\,r(t,s)\| &= \left\|\mathsf{E}\int_s^t (U(t,u)-I)a(u,X(u))\,du\right\| \\ &\le \mathsf{E}\left\|\int_s^t\int_u^t A(z)U(z,u)A^{-1}(u)A(u)a(u,X(u))\,dz\,du\right\| \\ &\le C\,|t-s|^2\,(1+\mathsf{E}\,\|X(u)\|) \le C\,|t-s|^2\,. \end{aligned}$$

Further,

$$\begin{aligned}\|r(t,s)\|^2 &\le 2\Bigg(\mathsf{E}\left\|\int_s^t (U(t,u)-I)a(u,X(u))\,du\right\|^2 \\ &\qquad + \mathsf{E}\left\|\int_s^t (U(t,u)-I)b(s,X(s))\,dW(s)\right\|^2\Bigg) \\ &\le C\,|t-s| \int_s^t \mathsf{E}\left\|\int_u^t A(z)U(z,u)A^{-1}(u)A(u)a(u,X(u))\,dz\right\|^2 du \\ &\quad + C\int_s^t \mathsf{E}\left\|\int_u^t A(z)U(z,u)A^{-1}(u)A(u)b(u,X(u))\,dz\right\|_{\mathcal{L}_2}^2 du \\ &\le C\,|t-s|^4 + C\,|t-s|^3 \le C\,|t-s|^3 .\end{aligned}$$

This implies that

$$\mathsf{E}\,\|r(t,s)-\mathsf{E}\,r(t,s)\|^2 \le C\,|t-s|^3 .$$

Consider the uniform partition $\{\tau_n = t/N, n=0,\dots,N\}$ of the segment $[0,t]$ and prove that

$$r_N = \sum_{n=0}^{N-1} r(\tau_{n+1},\tau_n) \to 0, \qquad N\to\infty,$$

in probability. Indeed,

$$\|\mathsf{E}\,r_N\| = O\left(N^{-1}\right), \qquad N\to\infty,$$

and

$$\mathsf{E}\,\|r_N - \mathsf{E}\,r_N\|^2 = \sum_{n=0}^{N-1}\mathsf{E}\,\|r(\tau_{n+1},\tau_n)-\mathsf{E}\,r(\tau_{n+1},\tau_n)\|^2 \le C\sum_{n=0}^{N-1}|\tau_{n+1}-\tau_n|^3 = \frac{C}{N^2}.$$

This implies the convergence in probability:

$$X(t)-X(0)-(Y(t)-Y(0)) = \lim_{N\to\infty}\sum_{n=0}^{N-1}(U(\tau_{n+1},\tau_n)-I)X(\tau_n).$$

Estimating as before, we get

$$\begin{aligned}&\mathsf{E}\left\|\int_{\tau_n}^{\tau_{n+1}} A(s)X(s)\,ds - (U(\tau_{n+1},\tau_n)-I)X(\tau_n)\right\|^2 \\ &\quad = \mathsf{E}\left\|\int_{\tau_n}^{\tau_{n+1}}\int_{\tau_n}^{s} A(u)U(s,u)\big(a(u,X(u))\,du + b(u,X(u))\,dW(u)\big)\,ds\right\|^2 \\ &\quad \le C\,|\tau_{n+1}-\tau_n| \int_{\tau_n}^{\tau_{n+1}}\int_{\tau_n}^{s}\left(1+\mathsf{E}\,\|X(u)\|^2\right)du\,ds \le C\,|\tau_{n+1}-\tau_n|^3 ,\end{aligned}$$

whence

$$\lim_{N\to\infty}\sum_{n=0}^{N-1}(U(\tau_{n+1},\tau_n)-I)X(\tau_n) = \int_0^t A(u)X(u)\,du.$$

□

3. Approximations of Itô–Volterra type equations

Consider an abstract Itô–Volterra equation of the form

$$X(t) = m(t) + \int_0^t a(t,s,X(s))\,ds + \int_0^t b(t,s,X(s))\,dW(s), \qquad t\in[0,T], \tag{3.1}$$

where $a\colon S\times X\to X$ and $b\colon S\times X\to \mathcal{L}_2$ are measurable maps,

$$S=\left\{(t,s)\in[0,T]^2\colon s\le t\right\},$$

and $m(t)$ is an $\mathcal{F}_t$-adapted continuous square integrable process. Assume that the coefficients and the process $m(t)$ satisfy the following conditions:

$$\|a(t,s,x)-a(t,s,y)\|_X+\|b(t,s,x)-b(t,s,y)\|_{\mathcal{L}_2}\le C\,\|x-y\|\,, \tag{3.2a}$$

$$\|a(t,s,x)\|+\|b(t,s,x)\|_{\mathcal{L}_2}\le C(1+\|x\|), \tag{3.2b}$$

$$\begin{aligned}&\|a(t,s,x)-a(t,u,x)\|+\|b(t,s,x)-b(t,u,x)\|_{\mathcal{L}_2}\\&\qquad\le C(1+\|x\|)\,|s-u|^{1/2}\,,\end{aligned} \tag{3.2c}$$

$$\begin{aligned}&\|a(t,u,x)-a(s,u,x)\|+\|b(t,u,x)-b(s,u,x)\|_{\mathcal{L}_2}\\&\qquad\le C(1+\|x\|)\,|t-s|^{1/2}\,,\end{aligned} \tag{3.2d}$$

$$\mathsf{E}\,\|m(t)-m(s)\|^2\le|t-s| \tag{3.2e}$$

for $0\le u\le s\le t\le T$ and $x\in X$.

Proposition 3.1. *If conditions* (3.2) *hold, then equation* (3.1) *has a unique solution, and moreover*

$$\begin{aligned}\mathsf{E}\,\|X(t)\|^2\le C,&\qquad t\in[0,T],\\ \mathsf{E}\,\|X(t)-X(s)\|^2\le C\,|t-s|\,,&\qquad t,s\in[0,T].\end{aligned} \tag{3.3}$$

Proof. The existence, uniqueness, and square integrability of the solution are proved in [2] (Theorem 7.2.1 and remarks thereafter). Further,

$$\begin{aligned}&\mathsf{E}\,\|X(t)-X(s)\|^2\\&\quad\le C\left(\mathsf{E}\,\|m(t)-m(s)\|^2+\mathsf{E}\int_0^t\left\|\left(a(t,u,X(u))-a(s,u,X(u))\right)\right\|^2du\right.\\&\qquad\qquad\left.+\,\mathsf{E}\int_0^t\left\|\left(b(t,u,X(u))-b(s,u,X(u))\right)\right\|^2_{\mathcal{L}_2}du\right)\\&\quad\le C\,|t-s|\left(1+\sup_{t\in[0,T]}\mathsf{E}\,\|X(t)\|^2\right)\le C\,|t-s|\end{aligned}$$

for $s\le t$. □

Approximations for equation (3.1) are defined by

$$Y_n^\delta=m^\delta(\tau_n)+\sum_{i=0}^{n-1}\left(a\left(\tau_n,\tau_i,Y_i^\delta\right)\delta+b\left(\tau_n,\tau_i,Y_i^\delta\right)\left(W(\tau_{i+1})-W(\tau_i)\right)\right),\qquad n\ge1,$$

where $m^\delta(t)$ is a certain approximation of $m(t)$. In contrast to ordinary equations, for which the value of the approximate solution at the next point is determined by the value at the previous one and by the increment of the Wiener process, now we must know all the preceding values.

Similarly to (1.4), setting $Y^\delta(\tau_n)=Y_n^\delta$, it is possible to construct the continuous interpolation:

$$Y^\delta(t)=m^\delta(t)+\int_0^t a\left(t,\tau_{n_s},Y^\delta(\tau_{n_s})\right)ds+\int_0^t b\left(t,\tau_{n_s},Y^\delta(\tau_{n_s})\right)dW(s). \tag{3.4}$$

Theorem 3.2. *If the coefficients of* (3.1) *satisfy assumptions* (3.2) *and*

$$
\begin{aligned}
&\mathsf{E}\sup_{t\in[0,T]}\left\|m^\delta(t)-m(t)\right\|^2\le C\delta,\\
(3.5)\qquad &\left\|\partial_t b(t,s,x)-\partial_t b(t,s,y)\right\|_{\mathcal{L}_2}\le C\left\|x-y\right\|,\\
&\left\|\partial_t b(t,s,x)-\partial_t b(t,u,x)\right\|_{\mathcal{L}_2}\le C\left|s-u\right|^{1/2}(1+\left\|x\right\|),
\end{aligned}
$$

then approximations (3.4) *converge to the solution of* (3.1) *in the sense that*

$$
\mathsf{E}\sup_{0\le t\le T}\left\|X(t)-Y^\delta(t)\right\|^2\le K\delta. \tag{3.6}
$$

Remark 3.3. Here $\partial_t b(t,s,x)\colon S\times X\to\mathcal{L}_2$ is a measurable function such that

$$
b(u,s,x)=b(v,s,x)+\int_v^u\partial_t b(r,s,x)\,dr.
$$

At first glance it might seem that one can repeat the arguments of Theorem 1.1 under the conditions (3.2) and prove an inequality similar to (1.5). But now the integrals $\int\cdots dW(t)$ in (3.1) and (3.4) are not martingales, thus the Doob–Burkholder inequality does not apply. This is the reason for introducing the extra condition of the existence and Lipschitz continuity of the derivatives.

Proof. As in the proof of Theorem 1.1, put

$$
Z(t)=\mathsf{E}\sup_{0\le s\le t}\left\|X(s)-Y^\delta(s)\right\|^2
$$

and write

$$
Z(t)\le 3\left(\mathsf{E}\sup_{t\in[0,T]}\left\|m^\delta(t)-m(t)\right\|^2+A+B\right)
$$

where

$$
\begin{aligned}
A&=\mathsf{E}\sup_{0\le s\le t}\left\|\int_0^s\left(a(s,u,X(u))-a\left(s,\tau_{n_u},Y^\delta(\tau_{n_u})\right)\right)du\right\|^2,\\
B&=\mathsf{E}\sup_{0\le s\le t}\left\|\int_0^s\left(b(s,u,X(u))-b\left(s,\tau_{n_u},Y^\delta(\tau_{n_u})\right)\right)dW(u)\right\|^2.
\end{aligned}
$$

Now we estimate B as follows:

$$
\begin{aligned}
&\mathsf{E}\sup_{0\le s\le t}\left\|\int_0^s\left(b(s,u,X(u))-b\left(s,\tau_{n_u},Y^\delta(\tau_{n_u})\right)\right)dW(u)\right\|^2\\
&\quad\le 3\Bigg(\mathsf{E}\sup_{0\le s\le t}\left\|\int_0^s\left(b(s,u,X(u))-b(s,u,X(\tau_{n_u}))\right)dW(u)\right\|^2\\
&\qquad+\mathsf{E}\sup_{0\le s\le t}\left\|\int_0^s\left(b(s,u,X(\tau_{n_u}))-b(s,\tau_{n_u},X(\tau_{n_u}))\right)dW(u)\right\|^2\\
&\qquad+\mathsf{E}\sup_{0\le s\le t}\left\|\int_0^s\left(b(s,\tau_{n_u},X(\tau_{n_u}))-b\left(s,\tau_{n_u},Y^\delta(\tau_{n_u})\right)\right)dW(u)\right\|^2\Bigg).
\end{aligned}
$$

All three terms in the brackets are estimated in the same way: the difference $(b(s,\dots)-b(s,\dots))$ in every integral is replaced by

$$
b(u,\dots)-b(u,\dots)+\int_u^s\partial_t(b(r,\dots)-b(r,\dots))\,dr.
$$

Then the first term is estimated similarly to the proof of Theorem 1.1. In the second term we change the order of integration by using the stochastic Fubini theorem (see

Theorem 4.18 in [3]). Then the norm of the exterior (nonrandom) integrals is estimated by the integral of the norm, and the supremum of the latter is attained at $s = t$. Now we can estimate the first term:

$$B_1 \le C \int_0^t \mathsf{E} \left\| \int_0^r \left(\partial_t b(r,u,X(u)) - \partial_t b(r,u,X(\tau_{n_u}))\right) dW(u) \right\|^2 dr$$
$$\le C \int_0^t \int_0^r \mathsf{E} \left\| \left(\partial_t b(r,u,X(u)) - \partial_t b(r,u,X(\tau_{n_u}))\right) \right\|^2 du\, dr$$
$$\le C \int_0^t \int_0^r \mathsf{E} \left\| X(u) - X(\tau_{n_u}) \right\|^2 du\, dr \le C\delta.$$

The second and the third terms are estimated in a similar way:

$$B_2 \le C \int_0^t \int_0^r \mathsf{E} \left\| \partial_t b(r,u,X(\tau_{n_u})) - \partial_t b(r,\tau_{n_u},X(\tau_{n_u})) \right\|^2 du\, dr$$
$$\le C \int_0^t \int_0^r |u - \tau_{n_u}|\, \mathsf{E} \left(1 + \|X(\tau_{n_u})\|^2\right) du\, dr \le C\delta,$$
$$B_3 \le C \int_0^t \int_0^r \mathsf{E} \left\| \partial_t b(s,\tau_{n_u},X(\tau_{n_u})) - \partial_t b\left(s,\tau_{n_u},Y^\delta(\tau_{n_u})\right) \right\|^2 du\, dr$$
$$\le C \int_0^t \int_0^r \mathsf{E} \left\| X(\tau_{n_u}) - Y^\delta(\tau_{n_u}) \right\|^2 du\, dr \le C \int_0^t Z(u)\, du.$$

The integrals above are finite because $X(t)$ and $Y^\delta(t)$ are square integrable and continuous. Thus,

$$Z(t) \le C\left(\delta + \int_0^t Z(u)\, du\right),$$

whence $Z(T) \le K\delta$ by the Gronwall lemma. □

Assumptions (3.5) are restrictive in the sense that they are redundant if we want only to estimate the rate of convergence at a fixed point. In other words, the following result is true.

Proposition 3.4. *If conditions* (3.2) *hold and*

$$\mathsf{E} \left\| m^\delta(t) - m(t) \right\| \le C\delta, \quad t \in [0,T],$$

then

$$\mathsf{E} \left\| X(t) - Y^\delta(t) \right\|^2 \le K\delta, \qquad t \in [0,T].$$

Proof. The proof follows the lines of that of Theorems 1.1 and 3.2. □

3.1. **Application to semilinear evolution equations.** We apply the results of the preceding section to equation (2.1). The approximations are constructed as follows:

$$Y_n^\delta = U(\tau_n,0)Y_0^\delta + \sum_{i=1}^n \left(U(\tau_n,\tau_i)a\left(\tau_i,Y_{i-1}^\delta\right)\delta + U(\tau_n,\tau_i)b\left(\tau_i,Y_{i-1}^\delta\right)\right)(W(\tau_i) - W(\tau_{i-1})).$$

The approximations can be rewritten in a compact form by using the evolution property:

$$Y_n^\delta = U(\tau_n,\tau_{n-1})\left(Y_{n-1}^\delta + a\left(\tau_n,Y_{n-1}^\delta\right)\delta + b\left(\tau_n,Y_{n-1}^\delta\right)(W(\tau_n) - W(\tau_{n-1}))\right).$$

Formally speaking, we approximate the solution of some Itô–Volterra equation; nevertheless the approximations are stepwise. Now we construct the linear interpolation as in (3.4):

$$
\begin{aligned}
Y^\delta(t) &= U(t,0)Y_0^\delta + \int_0^t U(t,\tau_{n_s})\left(a\left(\tau_{n_s}, Y^\delta(\tau_{n_s})\right)\,ds + b\left(\tau_{n_s}, Y^\delta(\tau_{n_s})\right)\,dW(s)\right)\\
&= U(t,\tau_{n_t})\Bigg(Y^\delta(\tau_{n_t}) + \int_{\tau_{n_t}}^t a\left(\tau_{n_t}, Y^\delta(\tau_{n_t})\right)\,ds\\
&\qquad\qquad + \int_{\tau_{n_t}}^t b\left(\tau_{n_t}, Y^\delta(\tau_{n_t})\right)\,dW(s)\Bigg).
\end{aligned}
\tag{3.7}
$$

Assume that the coefficients a and b satisfy (2.2) and

$$
\|a(t,x)-a(s,x)\| + \|b(t,x)-b(s,x)\|_{\mathcal{L}_2} \le C\,|t-s|^{1/2}\,(1+\|x\|)
\tag{3.8}
$$

for $s,t\in[0,T]$ and $x\in X$. We also assume that $A(s)$ is as in Section 2. Then

$$
\begin{gathered}
\|U(t,s)b(s,x)\|_{\mathcal{L}_2} \le \|U(t,s)\|_{\mathcal{L}}\,\|b(s,x)\|_{\mathcal{L}_2} \le C(1+\|x\|),\\
\|U(t,s)b(s,x)-U(t,s)b(s,y)\|_{\mathcal{L}_2} \le \|U(t,s)\|_{\mathcal{L}}\,\|b(s,y)\|_{\mathcal{L}_2}\,\|x-y\|,
\end{gathered}
$$

$$
\begin{aligned}
\|U(t,u)b(u,x)-U(s,u)b(u,x)\| &\le \left\|\int_s^t A(v)U(v,u)A^{-1}(u)A(u)b(u,x)\,dv\right\|_{\mathcal{L}_2}\\
&\le |t-s|\sup_{u,v}\left\|A(v)U(v,u)A^{-1}(u)\right\|_{\mathcal{L}}\sup_u\|A(u)b(u,x)\|_{\mathcal{L}_2}\\
&\le C\,|t-s|^{1/2}\,(1+\|x\|)
\end{aligned}
$$

for $x,y\in X$ and $0\le u\le s\le t\le T$. The same bounds for $U(t,s)a(s,x)$ can be proved in a similar way. Condition (3.2c) is not so easy to check. By the triangle inequality

$$
\begin{aligned}
&\|U(t,s)b(s,x)-U(t,u)b(u,x)\|\\
&\quad\le \|U(t,s)b(u,x)-U(t,u)b(u,x)\| + \|U(t,s)b(u,x)-U(t,s)b(s,x)\|.
\end{aligned}
$$

The estimate for the second term is obvious. The first one is estimated as follows:

$$
\begin{aligned}
\|U(t,s)b(u,x)-U(t,u)b(u,x)\|_{\mathcal{L}_2} &= \left\|U(t,s)\int_u^s A(v)U(v,u)A^{-1}(u)A(u)b(u,x)\,dv\right\|_{\mathcal{L}_2}\\
&\le C\,|s-u|\,(1+\|x\|).
\end{aligned}
$$

The process $m(t)=U(t,0)X_0$ is $\mathcal{F}_t$-adapted (even $\mathcal{F}_0$-measurable). It is square integrable since

$$
\|m(t)\|^2 \le \|U(t,0)\|\,\|X_0\|^2.
$$

Moreover,

$$
\begin{aligned}
\mathsf{E}\,\|m(t)-m(s)\|^2 &\le \mathsf{E}\left\|\int_s^t A(v)U(v,s)U(s,0)X_0\,dv\right\|^2\\
&= \mathsf{E}\left\|\int_s^t A(v)U(v,s)A^{-1}(s)A(s)U(s,0)A^{-1}(0)A(0)X_0\,dv\right\|^2\\
&\le C\,|t-s|
\end{aligned}
$$

for $s\le t$. Therefore all the conditions in (3.2) hold. Now consider some initial value of the approximations Y_0^δ such that $\mathsf{E}\left\|Y_0^\delta - X_0\right\|^2 \le C\delta$. Then

$$
\mathsf{E}\left\|m(t)-U(t,0)Y_0^\delta\right\|^2 \le \mathsf{E}\,\|U(t,0)\|^2\left\|Y_0^\delta - X_0\right\| \le C\delta.
$$

Thus, all the assumptions of Proposition 3.4 hold, that is, the following result is true.

Theorem 3.5. *Assume that the coefficients of* (2.1) *satisfy condition* (3.8)*. Under the assumptions of Section 2 and the condition*

$$\mathsf{E}\left\|Y_0^\delta - X_0\right\|^2 \le C\delta,$$

we have

$$\mathsf{E}\left\|Y^\delta(t) - X(t)\right\|^2 \le K\delta, \qquad t \in [0,T].$$

It is a natural question whether or not the uniform convergence as in Theorem 3.2 can be achieved in this case. Note that the coefficient of $dW(s)$ is differentiable with respect to the first variable:

$$\begin{aligned} U(t,u)b(u,x) - U(s,u)b(u,x) &= \int_s^t A(v)U(v,s)b(s,x)\,dv \\ &= \int_s^t A(v)U(v,s)A^{-1}(s)A(s)b(s,x)\,dv. \end{aligned}$$

The estimate

$$\mathsf{E}\sup_{t\in[0,T]}\left\|U(t)(Y_0^\delta - X(t))\right\|^2 < C\delta$$

is obvious. In order to apply Theorem 3.2 one must impose the following restriction:

$$\begin{aligned} \|A(t)(U(t,s)b(s,x) - U(t,s)b(s,y))\|_{\mathcal{L}_2} &\le C\|x-y\|, \\ \|A(t)(U(t,s)b(s,x) - U(t,u)b(u,x))\|_{\mathcal{L}_2} &\le C|s-u|^{1/2}(1+\|x\|). \end{aligned} \tag{3.9}$$

The latter condition holds, for example, if

$$\begin{aligned} \|A(t)(b(t,x) - b(t,y))\|_{\mathcal{L}_2} &\le C\|x-y\|, \\ \|A(t)b(t,x) - A(t)b(s,x)\|_{\mathcal{L}_2} &\le C|t-s|^{1/2}(1+\|x\|), \\ \left\|A^2(t)b(t,x)\right\|_{\mathcal{L}_2} &\le C(1+\|x\|). \end{aligned} \tag{3.10}$$

Theorem 3.6. *Let all the assumptions of Theorem 3.5, as well as* (3.10)*, hold. Then*

$$\mathsf{E}\sup_{t\in[0,T]}\left\|X(t) - Y^\delta(t)\right\|^2 \le K\delta.$$

Proof. It is sufficient to check (3.9). The first estimate follows from

$$\left\|A(t)U(t,s)A^{-1}(s)A(s)b(s,x)\right\|_{\mathcal{L}_2} \le \left\|A(t)U(t,s)A^{-1}(s)\right\|_{\mathcal{L}}\left\|A(s)b(s,x)\right\|_{\mathcal{L}_2}.$$

The second estimate is proved similarly to the proof of (3.2c): first we apply the triangle inequality resulting in two terms; one of the terms is estimated in an obvious way, while for the other one we get

$$\begin{aligned} &\left\|A(t)\big(U(t,s) - U(t,u)\big)b(u,x)\right\|_{\mathcal{L}_2} \\ &\le \left\|\int_u^s A(t)U(t,v)A^{-1}(v)A^2(v)A^{-2}(u)A^2(u)b(u,x)\,dv\right\|_{\mathcal{L}_2} \\ &\le C|s-u|(1+\|x\|). \end{aligned}$$

□

Example 3.7. Consider the following particular case of Example 2.1: $H = X = L_2(\mathbb{R})$ and $A(t) = A$ is the derivative operator. Equation (2.1) becomes

$$X(t,z) = X_0(z) + \int_0^t \Big(\frac{d}{dz}X(s,z) + a(s,X(s))(z)\Big)\,ds + \int_0^t b(s,X(s))(z)\,dW(s). \tag{3.11}$$

The operator A generates a strongly continuous semigroup of shifts $U(t)x(z) = x(t+z)$ on X. The operators A and $U(t)$ commute and all the assumptions on $U(t)$ introduced in Section 2 are obviously true. The assumptions on the coefficients a and b mean in this case that they "smooth" functions $x \in L_2$. (One can take, for instance,

$$a(t,x)(z) = \int_{z-r}^{z} x(u)\,du.)$$

The approximations are constructed in the following way:

$$Y_{n+1}^{\delta}(\cdot) = Y_n^{\delta}(\cdot+\delta) + a\left(\tau_n, Y_n^{\delta}\right)(\cdot+\delta)\cdot\delta + b\left(\tau_n, Y_n^{\delta}\right)(\cdot+\delta)\cdot\big(W(\tau_{n+1}) - W(\tau_n)\big).$$

Such approximations, in contrast to (2.3), converge in the mean-square sense to the true solution.

Bibliography

1. D. Barbu and V. Radu, *Approximations to mild solutions of stochastic semilinear equations*, Novi Sad J. Math. **30** (2000), no. 1, 183–190. MR1835643 (2002c:60104)
2. Yu. L. Daletskiĭ and S. V. Fomin, *Measures and Differential Equations in Infinite-Dimensional Spaces*, "Nauka", Moscow, 1983. (Russian) MR0720545 (86g:46059)
3. G. Da Prato and J. Zabczyk, *Stochastic Equations in Infinite Dimensions*, Cambridge Univ. Press, Cambridge, 1992. MR1207136 (95g:60073)
4. P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, Springer, Berlin, 1992. MR1214374 (94b:60069)
5. A. Kohatsu-Higa and P. Protter, *The Euler scheme for SDE's driven by semimartingales*, Pitman Res. Notes Math. Ser., vol. 310, Longman Sci. Tech., Harlow, 1994, pp. 141–151. MR1415665 (97i:60074)
6. A. Kolodiĭ, *On convergence of approximations of Itô–Volterra equations*, Progr. Systems Control Theory, vol. 23, Birkhäuser Boston, Boston, MA, 1997, pp. 157–165. MR1636834 (99g:60101)
7. D. F. Kuznetsov, *Some Problems in the Theory of Numerical Solution of Itô Stochastic Differential Equations*, S.-Peterburg. Gos. Tekhn. Univ., St. Petersburg, 1998. (Russian) MR1711917 (2000k:60116)
8. H. Lisei, *Approximation by time discretization of special stochastic evolution equations*, Math. Pannon. **12** (2001), no. 2, 245–268. MR1860165 (2002j:60113)
9. G. N. Milstein, *Numerical Integration of Stochastic Differential Equations*, Ural. Gos. Univ., Sverdlovsk, 1988; English transl., Mathematics and Its Applications, vol. 313, Kluwer Academic Publishers Group, Dordrecht, 1995. MR0955705 (90k:65018); MR1335454 (96e:65003)
10. P. Protter and D. Talay, *The Euler scheme for Lévy driven stochastic differential equations*, Ann. Probab. **25** (1997), no. 1, 393–423. MR1428514 (98c:60063)
11. D. Talay, *Simulation and numerical analysis of stochastic differential systems: a review*, Probabilistic Methods in Applied Physics (P. Krée and W. Wedig, eds.), Springer, Berlin, 1995, pp. 54–96.
12. C. Tudor and M. Tudor, *Approximation schemes for Itô–Volterra stochastic equations*, Bol. Soc. Mat. Mexicana **1** (1995), no. 1, 73–85. MR1350639 (96k:60165)

Department of Probability Theory and Mathematical Statistics, Faculty of Mechanics and Mathematics, National Taras Shevchenko University, Academician Glushkov Avenue 6, Kyiv 03127, Ukraine

E-mail address: zhora@univ.kiev.ua

Received 16/DEC/2002

Translated by THE AUTHOR