

## A GENERALIZATION OF MIL'SHTEIN'S THEOREM FOR STOCHASTIC DIFFERENTIAL EQUATIONS

UDC 519.21

GEORGIĬ SHEVCHENKO

ABSTRACT. A theorem on a relationship between local and global rates of convergence for a stochastic differential equation is proved in the paper. This theorem implies the convergence of Euler type approximations for semilinear evolution equations with a coercive operator in a Hilbert space.

### INTRODUCTION

There are many papers devoted to finding an approximate solution for stochastic differential equations. The main idea for the numerical solution of stochastic differential equations is the time discretization, where one considers another (approximating) equation with constant functions instead of the coefficients of the original equation on the intervals of discretization. The theory of approximation of finite systems of stochastic differential equations with regular coefficients can be considered as complete. The most extensive description of this theory is contained in the monographs by Mil'shtein [2] and Kloeden and Platen [1]. One of the best-known results in the theory of approximate solving of stochastic differential equations is the Mil'shtein theorem proved in [6]. This theorem establishes a relationship between local and global rates of approximation. One of the assumptions of this theorem is that the coefficients of the equation are Lipschitz, and this does not allow one to apply it to the case of semilinear stochastic evolution equations. Below we introduce a Lipschitz type approximation and prove a result similar to Mil'shtein's theorem for semilinear evolution stochastic differential equations. A corollary to this theorem obtains the rate of the approximation convergence for semilinear evolution equations.

The paper is organized as follows. Section 1 briefly discusses some results on the existence and uniqueness of solutions of the stochastic differential equation under consideration as well as their basic properties. Section 2 contains definitions of the convergence of approximation schemes. A generalization of the Mil'shtein theorem for semilinear stochastic evolution differential equations is proved in Section 3. Its corollary on the convergence of approximations for the semilinear equation of evolution type is also given in Section 3.

### 1. SEMILINEAR STOCHASTIC EVOLUTION EQUATIONS

Let  $\mathcal{H}$  be a separable Hilbert space,  $\mathcal{L}(\mathcal{H})$  be the space of continuous linear operators on  $\mathcal{H}$ , and let  $\mathcal{L}_2(\mathcal{H})$  be the space of Hilbert–Schmidt operators (in what follows we use

---

2000 *Mathematics Subject Classification.* Primary 60H10; Secondary 34G10, 47A50, 47D06.

*Key words and phrases.* Stochastic differential equation, semilinear stochastic evolution equation, time discretization, Mil'shtein theorem.

the abbreviations  $\mathcal{L}$  and  $\mathcal{L}_2$ , respectively). We use the same symbol  $C$  to denote those constants whose values are not important for our purposes.

Consider a semilinear stochastic differential equation on the interval  $[0, T]$  (throughout the paper we deal with the interval  $[0, T]$  unless otherwise specified):

$$(1.1) \quad X(t) = X_0 + \int_0^t (A(s)X(s) + a(s, X(s))) ds + \int_0^t b(s, X(s)) dW(s).$$

Here the coefficients  $a(s, x)$  and  $b(s, x)$  map  $[0, T] \times \mathcal{H}$  to  $\mathcal{H}$  and  $\mathcal{L}^2$ , respectively, and are continuous for all variables,  $A(t)$  are closed linear unbounded operators with a dense domain  $D(A(t))$ , and  $X_0$  is an  $\mathcal{F}_0$ -measurable random variable which usually is assumed to be second order integrable, since we will be mostly interested in the mean square convergence.

Equation (1.1) is not easy to study in the general case. We shall assume (and this is a quite natural assumption) that the operators  $A(s)$  are defined on a dense set  $D \subset \mathcal{H}$  and that they generate a strongly continuous *evolution family*  $\{U(t, s), 0 \leq s \leq t \leq T\}$ , that is, a strongly continuous two-parameter semigroup of linear continuous operators such that

- $U(t, t) = I$  for all  $t \in [0, T]$ , where  $I$  is the identity operator,
- $U(t, s)U(s, r) = U(t, r)$  for all  $t, s$ , and  $r$  such that  $0 \leq r \leq s \leq t \leq T$ ,
- $\frac{d}{dt}U(t, s)x = A(t)U(t, s)x$  for all  $x \in D(A(s))$ ,
- $\frac{d}{ds}U(t, s)x = -U(t, s)A(s)x$  for all  $x \in D(A(s))$ .

We additionally assume that the scalar product  $(\cdot, \cdot)_D$  can be defined in the space  $D$  such that the embedding operator of  $D$  into  $\mathcal{H}$  is continuous in the norm generated by this product, and that

$$A(t) \in \mathcal{L}(D, \mathcal{H})$$

for all  $t \in [0, T]$ . The latter assumption is quite natural. For example, in the homogeneous case of  $A(t) \equiv A$ , one can use the *graph norm*  $\|x\|_D^2 = \|x\|^2 + \|Ax\|^2$ .

Moreover, we assume that

$$U(t, s)D \subset D.$$

This condition holds in the homogeneous case, too. In the inhomogeneous case, it is a part of the standard Kato–Tanabe conditions [7] for the existence of an evolution family.

In what follows we assume that the family of operators  $\{A(t), t \in [0, T]\}$  satisfies all the conditions listed above; we shall use the properties of the operators without mentioning them. Note that the strong continuity of the family  $U(t, s)$  implies that the norms  $\|U(t, s)x\|$  are bounded (with respect to  $s, t$ ) for all  $x \in \mathcal{H}$ . Thus the principle of uniform boundedness gives us  $\|U(t, s)\|_{\mathcal{L}} \leq C$ ; that is, the family  $U(t, s)$  is bounded in the operator norm. Similarly,

$$(1.2) \quad \|U(t, s)|_D\|_{\mathcal{L}(D)} \leq C.$$

Since the operators  $A(s)$  are linear and generate an evolution family, one can use the variation of constants method and obtain another equation for  $X(t)$ , namely

$$(1.3) \quad X(t) = U(t, 0)X_0 + \int_0^t U(t, s)(a(s, X(s)) ds + b(s, X(s)) dW(s)).$$

This equation and its solution are called the *variation of constants equation* and *mild solution*, respectively, for equation (1.1). The coefficients of the variation of constants equation are continuous and defined everywhere. We shall use this equation throughout the article, since it is easier to work with.

Now we turn to the question of the existence and uniqueness of a solution of equation (1.1). First we give conditions for the existence and uniqueness of the mild solution.

**Theorem 1.1** ([5]). *Assume that the coefficients of equation (1.1) are such that*

$$(1.4) \quad \begin{aligned} & \|a(t, x)\| + \|b(t, x)\| \leq C(1 + \|x\|), \\ & \|a(t, x) - a(t, y)\| + \|b(t, x) - b(t, y)\| \leq C\|x - y\|. \end{aligned}$$

*Then equation (1.3) has a unique solution; that is, equation (1.1) has a unique mild solution.*

Note that the uniqueness of a solution of equation (1.3) implies that the strong solution of equation (1.1) is unique if it exists. However, we will always emphasize that a solution is unique. Note also that, under the assumptions of Theorem 1.1, it is possible to evaluate a solution using the method of successive approximations.

**Theorem 1.2** ([5]). *Assume that equation (1.1) has a unique soft solution. Moreover let*

- $X_0 \in D$  almost surely,
- $U(t, s)a(s, x) \in D, U(t, s)b(s, x)h \in D$  for all  $s \leq t \leq T$  and  $x, h \in \mathcal{H}$ ,
- $\int_0^T \int_0^t (\|A(t)U(t, s)a(s, x)\|^2 + \|A(t)U(t, s)b(s, x)\|_{\mathcal{L}^2}^2) ds dt < \infty$ .

*Then equation (1.1) has a unique strong solution.*

The following theorem describes the main properties of the solutions of equation (1.1).

**Theorem 1.3.** *Assume that the assumptions of Theorem 1.1 hold. Then the mild solution of equation (1.1) is such that*

$$\begin{aligned} \mathbb{E} \left[ \|X(t)\|^2 \right] & \leq C \left( 1 + \mathbb{E} \left[ \|X_0\|^2 \right] \right), \\ \mathbb{E} \left[ \|X(t) - X(s)\|^2 \right] & \leq C |t - s| \left( 1 + \mathbb{E} \left[ \|X_0\|^2 \right] \right). \end{aligned}$$

*If the assumptions of Theorem 1.2 hold and  $\mathbb{E}[\|X_0\|_D^2] < \infty$ , then*

$$\mathbb{E} \left[ \|X(t)\|_D^2 \right] \leq C \left( 1 + \mathbb{E} \left[ \|X_0\|_D^2 \right] \right).$$

*Proof.* The first assertion is standard (see [4]). The method of the proof of the second assertion is very similar: one uses estimate (1.2) and the Gronwall lemma.  $\square$

## 2. APPROXIMATION SCHEMES. MIL'SHTEIN THEOREM

Consider a finite- or infinite-dimensional stochastic differential equation

$$(2.1) \quad X(t) = X_0 + \int_0^t a(s, X(s)) ds + \int_0^t b(s, X(s)) dW(s)$$

on the interval  $[0, T]$ .

The key idea of the time discretization is as follows. Let  $N$  be a natural number. Consider the partition of the interval  $[0, T]$  into segments of equal size  $\delta = T/N$  generated by the points (nodes)  $\tau_n = n\delta, n = 0, \dots, N$ . The explicit approximation scheme is constructed in the following way. Set an initial value  $X_0^\delta$  of the approximate solution (usually we take  $X_0^\delta = X_0$ ) and define approximations recurrently in the following way:

$$(2.2) \quad X^\delta(\tau_{n+1}) = F(\tau_n, X^\delta(\tau_n), \mathbf{W}_{\tau_n}), \quad n = 0, \dots, N - 1,$$

where  $F$  is a certain nonrandom measurable function and

$$\mathbf{W}_{\tau_n} := \{W(t) - W(\tau_n), t \in [\tau_n, \tau_{n+1}]\}$$

is the path of the Wiener process in the interval  $[\tau_n, \tau_{n+1}]$ . The scheme (2.2) is suitable to approximate the solution of the equation at the nodes of the partition. To approximate the solution between the nodes, an interpolation procedure is usually used.

We denote by  $X^\delta(\tau_{n+1}, \tau_n, x)$  the value of the approximation at the node  $\tau_{n+1}$  given that its value at the node  $\tau_n$  is  $x$ . We also denote by  $X(t, s, x)$  the value at a moment  $t$  of the solution of the following equation:

$$Y(r) = x + \int_s^r a(u, Y(u)) du + \int_s^r b(u, Y(u)) dW(u).$$

It is clear that the terms on the right hand side depend on a trajectory of the Wiener process. Nevertheless, we do not indicate this dependence in the notation; it means that the same trajectory is involved on both sides of an equality if those terms appear there.

The convergence of approximations can be understood in various ways: in probability, almost surely, in  $L_p$ , or in a weak sense. We are mostly concerned with the  $L_2$  convergence, since it fits most applications using the approximate solving of stochastic differential equations.

The quality of an approximate scheme of a stochastic differential equation is defined in the following way.

**Definition 2.1.** We say that an approximation scheme (2.2) has a rate  $l > 0$  of the  $L_2$  convergence (or, simply, the “rate of convergence”  $l$ ) if

$$\mathbb{E} \left[ \|X(\tau_n) - X^\delta(\tau_n)\|^2 \right] \leq C \left( 1 + \mathbb{E} [\|X_0\|^2] \right) \delta^{2l},$$

where  $C$  is a constant that is independent of both  $\delta$  and  $n$ .

The rate of convergence is not uniquely defined. Indeed, any positive number smaller than  $l$  is also a rate of convergence.

We also need the notion of the local rate of convergence.

**Definition 2.2.** We say that the approximation scheme (2.2) has a local ( $L_2$ ) rate of convergence  $l > 0$  if

$$\mathbb{E} \left[ \|X(\tau_{n+1}, \tau_n, x) - X^\delta(\tau_{n+1}, \tau_n, x)\|^2 \right] \leq C \left( 1 + |x|^2 \right) \delta^{2l},$$

where the constant  $C$  is independent of  $\delta$  and  $k$ .

If

$$\|\mathbb{E} [X(\tau_{n+1}, \tau_n, x) - X^\delta(\tau_{n+1}, \tau_n, x)]\|^2 \leq C \left( 1 + |x|^2 \right) \delta^{2l}$$

for the approximation scheme (2.2), then we say that the scheme is locally stable in the mean with the rate  $l$  (we also say the scheme has the “local rate of stability  $l$ ”).

Mil’shtein [6] proved a result linking the local and global rates of convergence. Mil’shtein [6] assumes the global Lipschitz property and linear growth of coefficients, namely that the coefficients  $a$  and  $b$  are continuous in all their arguments and satisfy inequality (1.4), where the constant  $C$  is independent of  $t$ ,  $x$ , and  $y$ . The result itself reads as follows.

**Theorem 2.3** ([6]). *Let an approximation scheme have a local rate of convergence  $l_1$  and local rate of stability  $l_2$ . Assume that  $1 \leq l_1 + 1/2 \leq l_2$ . Then this scheme has the rate of convergence  $l_1 - 1/2$ .*

### 3. CONVERGENCE OF APPROXIMATIONS FOR SEMILINEAR EVOLUTION EQUATIONS

We turn to our basic model. Recall that we consider a semilinear stochastic differential equation (1.1) in a Hilbert space  $\mathcal{H}$  and the operator  $A(s)$  is unbounded and generates a strongly continuous evolution family  $\{U(t, s), 0 \leq s \leq t \leq T\}$ .

There are certain generalizations of Theorem 2.3 weakening conditions (1.4). For example, one can assume only the local Lipschitz property or one can allow the constant  $C$  to depend on  $t$  in the condition of linear growth (see [3]). Unfortunately, these conditions

or even their weaker versions are not satisfied in most of the cases. This, in particular, concerns equation (1.1). Thus we need a stronger version of the Mil'shtein theorem. We consider Lipschitz approximations instead of imposing the Lipschitz condition on the coefficients of equation (2.1).

**Definition 3.1.** An approximation  $X^\delta$  is called a Lipschitz scheme if there exists an equivalent norm  $\|\cdot\|_0$  in the space  $H$  such that

$$\mathbb{E} \left[ \left\| X^\delta(\tau_{n+1}, \tau_n, x) - X^\delta(\tau_{n+1}, \tau_n, y) \right\|_0^2 \right] \leq e^{C\delta} \|x - y\|_0^2.$$

*Remark 3.2.* The latter definition uses an equivalent norm, and this might seem strange at first glance. Nevertheless it is quite natural. Were the definition to use the norm  $\|\cdot\|$ , it would cause an instability in the sense that a Lipschitz scheme may turn into a non-Lipschitz scheme under an equivalent norm or, vice versa, may turn a non-Lipschitz scheme into a Lipschitz one. It is clear that both Definitions 2.1 and 2.2 are stable in this sense.

**Theorem 3.3.** Assume that equation (1.1) has a unique strong solution  $X(t)$  and that  $X^\delta$  is an approximation of this solution. Let the second moment of this solution be bounded, that is,

$$\mathbb{E} \left[ \|X(t)\|^2 \right] \leq C \left( 1 + \mathbb{E} \left[ \|X_0\|^2 \right] \right),$$

and let the scheme  $X^\delta$  be of the Lipschitz type and have the local rate of convergence  $l > \frac{1}{2}$ . Then the scheme  $X^\delta$  has the rate of convergence  $l - \frac{1}{2}$ .

*Proof.* Let  $\|\cdot\|_0$  be the norm involved in the definition of the Lipschitz scheme. Due to the equivalency, the scheme has the local rate of convergence  $l$ . It is clear that if the scheme has the global rate of convergence  $l - \frac{1}{2}$  for the norm  $\|\cdot\|_0$ , then the desired property follows for the norm  $\|\cdot\|$ . Thus, without loss of generality and for the sake of simplicity, we assume that  $\|\cdot\| = \|\cdot\|_0$ . Put

$$Z_n = \mathbb{E} \left[ \|X(\tau_n) - X^\delta(\tau_n)\|^2 \right].$$

Then

$$\begin{aligned} Z_{n+1} &\leq (1 + \delta) \mathbb{E} \left[ \|X^\delta(\tau_{n+1}, \tau_n, X^\delta(\tau_n)) - X^\delta(\tau_{n+1}, \tau_n, X(\tau_n))\|^2 \right] \\ &\quad + (1 + 1/\delta) \mathbb{E} \left[ \|X(\tau_{n+1}, \tau_n, X(\tau_n)) - X^\delta(\tau_{n+1}, \tau_n, X(\tau_n))\|^2 \right] \\ &\leq e^\delta \mathbb{E} \left[ \mathbb{E} \left[ \|X^\delta(\tau_{n+1}, \tau_n, X^\delta(\tau_n)) - X^\delta(\tau_{n+1}, \tau_n, X(\tau_n))\|^2 \middle| \mathcal{F}_{\tau_n} \right] \right] \\ &\quad + (1 + 1/\delta) \mathbb{E} \left[ \mathbb{E} \left[ \|X(\tau_{n+1}, \tau_n, X(\tau_n)) - X^\delta(\tau_{n+1}, \tau_n, X(\tau_n))\|^2 \middle| \mathcal{F}_{\tau_n} \right] \right]. \end{aligned}$$

Note that

$$X(\tau_{n+1}, \tau_n, x) = G(x, \mathbf{W}_{\tau_n})$$

for some  $G$  that is a measurable functional of a trajectory of the Wiener process  $\mathbf{W}_{\tau_n}$  on an interval  $[\tau_n, \tau_{n+1}]$ ; we treat this trajectory as a random element in the space of all continuous functions on this interval. On the other side, the increments

$$\{W(s) - W(\tau_n), s \geq \tau_n\}$$

(and thus, the path  $\mathbf{W}_{\tau_n}$  as well) do not depend on  $\mathcal{F}_{\tau_n}$ . Using a well-known property of the conditional mathematical expectation, we have the following almost sure equalities:

$$\begin{aligned} & \mathbb{E} \left[ \left\| X^\delta(\tau_{n+1}, \tau_n, X^\delta(\tau_n)) - X^\delta(\tau_{n+1}, \tau_n, X(\tau_n)) \right\|^2 \middle| \mathcal{F}_{\tau_n} \right] \\ &= \mathbb{E} \left[ \left\| X^\delta(\tau_{n+1}, \tau_n, x) - X^\delta(\tau_{n+1}, \tau_n, y) \right\|^2 \middle|_{x=X^\delta(\tau_n), y=X(\tau_n)} \right], \\ & \mathbb{E} \left[ \left\| X(\tau_{n+1}, \tau_n, X(\tau_n)) - X^\delta(\tau_{n+1}, \tau_n, X(\tau_n)) \right\|^2 \middle| \mathcal{F}_{\tau_n} \right] \\ &= \mathbb{E} \left[ \left\| X(\tau_{n+1}, \tau_n, x) - X^\delta(\tau_{n+1}, \tau_n, x) \right\|^2 \middle|_{x=X(\tau_n)} \right]. \end{aligned}$$

Recalling the assumption on the rate of convergence of approximations, boundedness of the moments, and the Lipschitz property of the scheme, we get

$$\begin{aligned} Z_{n+1} &\leq e^{C\delta} \mathbb{E} \left[ \left\| X(\tau_n) - X^\delta(\tau_n) \right\|^2 \right] + C\delta^{2l-1} \left( 1 + \mathbb{E} \left[ \left\| X(\tau_n) \right\|^2 \right] \right) \\ &\leq e^{C\delta} Z_n + C\delta^{2l-1} \left( 1 + \mathbb{E} \left[ \left\| X_0 \right\|^2 \right] \right). \end{aligned}$$

Using an induction we prove that

$$Z_n \leq C e^{C\delta n} \delta^{2l-1} \left( 1 + \mathbb{E} \left[ \left\| X_0 \right\|^2 \right] \right) \leq C e^{C\delta N} \delta^{2l-1} \left( 1 + \mathbb{E} \left[ \left\| X_0 \right\|^2 \right] \right),$$

whence the desired inequality follows:

$$Z_n \leq C\delta^{2l-1} \left( 1 + \mathbb{E} \left[ \left\| X_0 \right\|^2 \right] \right). \quad \square$$

*Remark 3.4.* The assumption that the constant in Definition 2.2 is of linear growth with respect to the initial condition  $x$  can be replaced by another assumption allowing a more complex dependence of  $x$ . In this case, Theorem 3.3 remains valid if we assume the existence of the mathematical expectations of those constants for which the value of a solution substitutes  $x$  instead of the finiteness of moments of the solution. It is clear that the constant in Definition 2.1 may grow faster than a linear function with respect to the initial conditions.

Recall that any strong solution of equation (1.1) is also a mild solution, that is, a solution of the variation of constant equation (1.3). Discretizing the time in this equation we obtain the following equation for the approximate solution:

$$(3.1) \quad X^\delta(t) = U(t, 0)X_0 + \int_0^t U(t, t_s) \left( a(t_s, X^\delta(t_s)) ds + b(t_s, X^\delta(t_s)) dW(s) \right).$$

After some simple algebra and taking into account the evolution property, we determine these approximations at the nodes step-by-step by

$$(3.2) \quad X^\delta(\tau_{n+1}) = U(\tau_n, \tau_{n-1}) \left( X^\delta(\tau_n) + a(\tau_n, X^\delta(\tau_n)) \delta + b(\tau_n, X^\delta(\tau_n)) \Delta W_n \right).$$

**3.1. Approximation of solutions of the equation with an evolution family of the uniform contractive type.** Assume that the coefficients  $a$  and  $b$  of equation (1.1) have the Lipschitz property and satisfy the condition of linear growth (1.4). Assume further that there exists a constant  $\beta \in \mathbb{R}$  such that

$$(3.3) \quad (A(t)x, x) \leq \beta \|x\|^2$$

for all  $x \in D$  and  $t \in [0, T]$ . In this case, the operator  $A(t)$  is called *coercive*, while the evolution family generated by  $A(t)$  is said to be of the *uniform contractive type*.

The following lemma clarifies this notion.

**Lemma 3.5.** *Let condition (3.3) hold. Then*

$$\|U(t, s)\|_{\mathcal{L}} \leq e^{\beta(t-s)}.$$

*Proof.* The definition of an evolution family implies that

$$\begin{aligned} \|U(t, s)x\|^2 &= \|U(s, s)x\|^2 + \int_s^t 2(A(r)U(r, s)x, U(r, s)x) dr \\ &\leq \|x\|^2 + 2\beta \int_0^t \|U(r, s)x\|^2 dx \end{aligned}$$

for  $x \in D$  and  $s \in [0, T]$ . Using the Gronwall lemma we get

$$\|U(t, s)x\| \leq e^{\beta(t-s)} \|x\|$$

for  $x \in D$ . Taking into account that  $D \in \mathcal{H}$  is dense everywhere and that the operator  $U(t, s)$  is continuous we complete the proof of the lemma.  $\square$

There are many examples of uniformly contractive families of operators, say, contractive semigroups. We shall show below that the convergence of the approximations for such equations follows from weaker assumptions as compared to those in Theorem 1.2 imposed on the coefficients  $a$  and  $b$  in equation (1.1).

First we prove the Lipschitz property of the approximation scheme.

**Proposition 3.6.** *Let (1.4) and (3.3) hold. Then the approximation scheme (3.2) is of the Lipschitz type.*

*Proof.* For  $h \in \mathcal{H}$  and  $B \in \mathcal{L}^2$ , we have

$$(3.4) \quad \mathbb{E} \left[ \|h + B\Delta W_n\|^2 \right] = \mathbb{E} \left[ \|h\|^2 + 2(h, B\Delta W_n) + \|B\Delta W_n\|^2 \right] = \|h\|^2 + \delta \|B\|_{\mathcal{L}^2}^2.$$

Setting  $h = (x - y) + (a(\tau_n, x) - a(\tau_n, y))\delta$  and  $B = (b(\tau_n, x) - b(\tau_n, y))$  and using Lemma 3.5 we obtain

$$\begin{aligned} &\mathbb{E} \left[ \|X^\delta(\tau_{n+1}, \tau_n, x) - X^\delta(\tau_{n+1}, \tau_n, y)\|^2 \right] \\ &\leq \|U(\tau_{n+1}, \tau_n)\|_{\mathcal{L}}^2 \\ &\quad \times \left( \|(x - y) + (a(\tau_n, x) - a(\tau_n, y))\delta\|^2 + \delta \|(b(\tau_n, x) - b(\tau_n, y))\|_{\mathcal{L}^2}^2 \right) \\ &\leq e^{2\beta\delta} \left( (1 + C\delta) \|x - y\|^2 + \delta^2(1 + 1/\delta) \|a(\tau_n, x) - a(\tau_n, y)\|^2 \right) \\ &\leq e^{2\beta\delta} (1 + C\delta) \|x - y\|^2 \leq e^{C\delta} \|x - y\|^2. \end{aligned} \quad \square$$

Now we consider the local rate of convergence. Since  $t_s$  replaces  $s$  in the coefficients  $a$  and  $b$ , we need a certain regularity of these coefficients with respect to the first argument, namely

$$(3.5) \quad \|a(t, x) - a(s, x)\| + \|b(t, x) - b(s, x)\|_{\mathcal{L}^2} \leq C |t - s|^{1/2} (1 + \|x\|).$$

**Theorem 3.7.** *Let the conditions (1.4), (3.3), and (3.5) hold and let equation (1.1) have a strong solution. Assume that  $U(t, s)a(s, x) \in D$  and  $U(t, s)b(s, x)h \in D$  for all  $x \in D$ ,  $h \in \mathcal{H}$ , and  $s \leq t \leq T$ . Then the rate of convergence of the approximation scheme (3.2) to the solution of equation (1.1) is equal to 1/2.*

*Proof.* Theorem 1.3 implies that

$$\mathbb{E} \left[ \|X(t)\|^2 \right] \leq C \left( 1 + \mathbb{E} \left[ \|X_0\|^2 \right] \right).$$

By Theorem 3.3, it is sufficient to prove that the local rate of convergence of the approximation scheme is equal to 1. To do this, it is necessary to check that the definition of

the local rate of convergence holds for all  $x \in D$ , since both  $X(t)$  and  $X^\delta(t)$  are continuous with respect to the initial conditions and since  $D$  is everywhere dense in  $\mathcal{H}$ . For  $t \in [\tau_n, \tau_{n+1}]$ , let

$$X^\delta(t, \tau_n, x) = U(t, \tau_n) [x + a(\tau_n, x)(t - \tau_n) + b(\tau_n, x)(W(t) - W(\tau_n))].$$

Note that this definition is compatible with (3.2). Using the properties of an evolution family one can check that

$$X^\delta(t, \tau_n, x) = x + \int_{\tau_n}^t \left[ (AX^\delta(s, \tau_n, x) + a(\tau_n, x)) ds + b(\tau_n, x) dW(s) \right].$$

Put  $Z(t) = X(t, \tau_n, x) - X^\delta(t, \tau_n, x)$ . We derive from Itô's formula that

$$\begin{aligned} \|Z(t)\|^2 &= 2 \int_{\tau_n}^t \left( Z(s), A(s)(X(s, \tau_n, x) - X^\delta(s, \tau_n, x)) \right) ds \\ &\quad + 2 \int_{\tau_n}^t \left( Z(s), a(s, X(s, \tau_n, x)) - a(\tau_n, x) \right) ds \\ &\quad + 2 \int_{\tau_n}^t \left( Z(s), (b(s, X(s, \tau_n, x)) - b(\tau_n, x)) dW(s) \right) \\ &\quad + \int_{\tau_n}^t \|b(s, X(s, \tau_n, x)) - b(\tau_n, x)\|_{\mathcal{L}^2}^2 ds. \end{aligned}$$

The mathematical expectation of the penultimate term is zero. We estimate the mathematical expectation of the last term as follows:

$$\begin{aligned} &\int_{\tau_n}^t \mathbb{E} \left[ \|b(s, X(s, \tau_n, x)) - b(\tau_n, x)\|_{\mathcal{L}^2}^2 \right] ds \\ &\leq C \int_{\tau_n}^t \mathbb{E} \left[ \|b(s, X(s, \tau_n, x)) - b(s, x)\|_{\mathcal{L}^2}^2 + \|b(s, x) - b(\tau_n, x)\|_{\mathcal{L}^2}^2 \right] ds \\ &\leq C \int_{\tau_n}^t \mathbb{E} \left[ \|X(s, \tau_n, x) - x\|^2 \right] ds + C \int_{\tau_n}^t |s - \tau_n| (1 + \|x\|^2) ds \\ &\leq C \int_{\tau_n}^t \mathbb{E} \left[ \|X(s, \tau_n, x) - X(\tau_n, \tau_n, x)\|^2 \right] ds + C\delta^2 (1 + \|x\|^2) \\ &\leq C\delta^2 (1 + \|x\|^2). \end{aligned}$$

The latter inequality follows from Theorem 1.3. Using the Cauchy–Bunyakovskiĭ inequality for the second term above and following a similar reasoning we get

$$\begin{aligned} &\mathbb{E} \left[ \int_{\tau_n}^t (Z(s), a(s, X(s, \tau_n, x)) - a(\tau_n, x)) ds \right] \\ &\leq C\delta^2 (1 + \|x\|^2) + C \int_{\tau_n}^t \mathbb{E} \left[ \|Z(s)\|^2 \right] ds. \end{aligned}$$

The first term is estimated by using condition (3.3). Finally we obtain

$$\mathbb{E} \left[ \|Z(t)\|^2 \right] \leq C\delta^2 (1 + \|x\|^2) + C \int_{\tau_n}^t \mathbb{E} \left[ \|Z(s)\|^2 \right] ds,$$

whence, according to the Gronwall lemma, we get

$$\mathbb{E} \left[ \|Z(t)\|^2 \right] \leq C\delta^2 (1 + \|x\|^2),$$

and this is what was to be proved.  $\square$



*Remark 3.8.* Theorem 3.7 remains valid for approximations of the following form:

$$\begin{aligned}\tilde{X}^\delta(\tau_{n+1}) &= \tilde{X}^\delta(\tau_{n+1}, \tau_n, \tilde{X}^\delta(\tau_n)) \\ &= U_n^\delta(\tilde{X}^\delta(\tau_n) + a(\tau_n, \tilde{X}^\delta(\tau_n))\delta + b(\tau_n, \tilde{X}^\delta(\tau_n))\Delta W_n)\end{aligned}$$

and for the operator  $U_n^\delta$  such that

$$(3.6) \quad \|U_n^\delta - U(\tau_{n+1}, \tau_n)\|_{\mathcal{L}} \leq C\delta.$$

Indeed, in this case,

$$\mathbb{E} \left[ \left\| \tilde{X}^\delta(\tau_{n+1}, \tau_n, x) - X^\delta(\tau_{n+1}, \tau_n, x) \right\|^2 \right] \leq C\delta^2 (1 + \|x\|^2).$$

Thus the local rate of convergence for approximations  $\tilde{X}$  is equal to 1. By (3.6) and Lemma 3.5, we prove the Lipschitz property for this scheme. Now Proposition 3.3 implies that the rate of convergence is  $1/2$ , which is exactly the same as for scheme (3.2).

*Remark 3.9.* Condition (3.5) means that the coefficients  $a$  and  $b$  satisfy the Hölder condition of order  $1/2$  with respect to the first argument. If the Hölder order is  $\gamma < 1/2$ , then one can see that the rate of convergence in Theorem 3.7 is  $\gamma$ .

#### BIBLIOGRAPHY

1. P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, Springer, Berlin, 1992. MR1214374 (94b:60069)
2. G. N. Mil'shtein, *Numerical Integration of Stochastic Differential Equations*, Ural'skiĭ Universitet, Sverdlovsk, 1988; English transl., Kluwer, Dordrecht, 1995. MR955705 (90k:65018); MR1335454 (96e:65003)
3. H. Schurz, *Stability, Stationarity, and Boundedness of Some Implicit Numerical Methods for Stochastic Differential Equations and Applications*, Logos-Verlag, Berlin, 1997. MR1991701
4. G. Da Prato and J. Zabczyk, *Stochastic Equations in Infinite Dimensions*, Encyclopedia of Mathematics and its Applications, vol. 44, Cambridge University Press, Cambridge, 1992. MR1207136 (95g:60073)
5. W. Grecksch and C. Tudor, *Stochastic Evolution Equations*, Mathematical Research, vol. 85, Akademie-Verlag, Berlin, 1995. MR1353910 (96m:60130)
6. G. N. Mil'shtein, *A theorem on the order of convergence of mean square approximations of solutions of stochastic differential equations*, Teor. Veroyatnost. i Primenen. **32** (1987), no. 4, 809–811; English transl. in Theory Probab. Appl. **32** (1988), no. 4, 738–741. MR927268 (89d:60104)
7. H. Tanabe, *Equations of Evolution*, Monographs and Studies in Mathematics, vol. 6, Pitman, Boston, MA, 1979. MR533824 (82g:47032)

DEPARTMENT OF PROBABILITY THEORY AND MATHEMATICAL STATISTICS, FACULTY OF MECHANICS AND MATHEMATICS, KIEV NATIONAL TARAS SHEVCHENKO UNIVERSITY, 64 VOLODYMYRSKA STREET, 01033 KIEV, UKRAINE

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

Received 9/JAN/2007

Translated by S. KVASKO