

MEAN-SQUARE NUMERICAL METHODS FOR STOCHASTIC DIFFERENTIAL EQUATIONS WITH SMALL NOISES*

G. N. MILSTEIN[†] AND M. V. TRET'YAKOV[‡]

Abstract. A new approach to the construction of mean-square numerical methods for the solution of stochastic differential equations with small noises is proposed. The approach is based on expanding the exact solution of the system with small noises in powers of time increment and small parameter. The theorem on the mean-square estimate of method errors is proved. Various efficient numerical schemes are derived for a general system with small noises and for systems with small additive and small colored noises. The proposed methods are tested by calculation of Lyapunov exponents and simulation of a laser Langevin equation with multiplicative noises.

Key words. stochastic differential equations, small noises, computer simulation

AMS subject classifications. 60H10, 65L99, 65U05

PII. S1064827594278575

1. Introduction. The stochastic approach has found wide application in physics [1], [2]. Usually analytic solutions for stochastic dynamical systems are hardly available. In this case the importance of numerical methods is obvious. In previous works various mean-square and weak numerical methods were derived for a general system of stochastic differential equations and for some specific systems such as systems with additive and colored noises, etc. (see [3], [4], [5], [6], [7]). In the general case some difficulties arise with realizing numerical methods for stochastic differential equations. For instance, there are no efficient Runge–Kutta schemes. As for mean-square approximation, there are no sufficiently constructive methods to simulate multiple Ito integrals.

But often fluctuations, which affect a physical system, are sufficiently small. Fortunately, as shown in this paper, for a stochastic system with small noises it is possible to construct special numerical methods which are more effective and easier than in the general case.

Herein for the first time numerical integration of stochastic differential equations with small noises is systematically considered. In this paper mean-square approximation is investigated. Weak methods will be the subject of a separate paper.

The system of Ito stochastic differential equations with small noises may be written in the form

$$(1.1) \quad dX = a(t, X) dt + \varepsilon \sum_{r=1}^q \sigma_r(t, X) dW_r, \quad X(t_o) = X_o, \quad t \in [t_o, T], \quad 0 \leq \varepsilon \leq \varepsilon_o,$$

where X , $a(t, X)$, and $\sigma_r(t, X)$, are n -dimensional vectors, W_r , $r = 1, \dots, q$, are uncorrelated standard Wiener processes, ε is a small parameter, and ε_o is a positive number.

If the parameter ε tends to zero, we have a deterministic system for which various effective numerical methods exist. One can believe that if parameter ε is sufficiently

*Received by the editors December 12, 1994; accepted for publication (in revised form) November 30, 1995.

<http://www.siam.org/journals/sisc/18-4/27857.html>

[†]Weierstraß Institut für Angewandte Analysis und Stochastik, Mohrenstr, 39, D-10117 Berlin, Germany (milstein@wias-berlin.de).

[‡]Department of Mathematics, Ural State University, Lenin St., 51, 620083 Ekaterinburg, Russia (michael.tretyakov@usu.ru).

small, i.e., the system (1.1) is sufficiently close to the deterministic one, it is also possible to obtain effective methods taking into account that ε is small.

In this paper, for the system (1.1) an approach to construction of effective numerical methods is proposed, a theorem on the estimate of mean-square global error is stated, and various mean-square methods with small errors are constructed including explicit, implicit, and Runge–Kutta schemes. The derived methods are efficient for the simulation of the used random variables. In the case of a general system, i.e., $\varepsilon = 1$, only schemes of the mean-square order $1/2$ can be efficient. And first-order methods already require calculation of complicated multiple Ito integrals which is a difficult and laborious problem [6], [7]. Using the approach developed in this paper, one can obtain for the system (1.1) numerical schemes with mean-square global errors which are usually estimated by $O(h^p + \varepsilon^k h^q)$, $q < p$. The order of such a method is equal to q which is not large by the reason of efficiency. However, the method error becomes sufficiently small due to large p and the factor ε^k at h^q , and the method reaches high exactness. As a result this idea leads us to construction of new methods with small errors which are efficient with respect to both simulation of the used random variables and calculation expenses. As we believe, these methods would be useful for many physical applications.

Our approach is based on expanding the exact solution of the system (1.1) in powers of time increment h and regrouping of expansion terms according to their factors $h^i \varepsilon^j$ (h - ε approach). One could think that to expand the exact solution first in powers of small parameter ε and then in powers of time increment h (ε - h approach) is more natural. But the ε - h approach suffers from a serious shortcoming because if constructed in this way, methods diverge. For details see section 3.2.

The organization of the paper is as follows. In section 2 we propose our approach to constructing one-step approximations for the solution of the system (1.1). The theorem on the mean-square estimate of method error is stated in section 3. Various efficient numerical schemes for Ito systems with small noises are presented in section 4. Section 5 is devoted to stochastic systems with small noises in Stratonovich interpretation. In sections 6 and 7 one can find numerical methods for systems with small additive and colored noises. Numerical tests of the proposed methods are presented in section 8.

The methods of sections 4–7 are given without derivations because their derivations and, especially, rigorous analysis of errors require a lot of space. But we hope that the content of section 2, where the main ideas for constructing numerical methods for a system with small noises are demonstrated, and the theorem on error estimate from section 3 are sufficient for a reader to understand our results.

We try to give a number of methods of various types and with various errors to ensure the opportunity to select the appropriate method in a concrete situation. It may not be out of place to remind the reader that a considerable number of numerical methods are used for simulating the simpler object, namely, ordinary differential equations. All our methods arise as the result of the strict selection carried out by us. For instance, we restrict ourselves to the methods which are efficient as to the simulation of the used random variables. We select only methods with $p \leq 4$ among methods with the errors $O(h^p + \varepsilon^k h^q)$, $0 < q < p$. We prefer methods with the smallest number of operators, etc. Each of the proposed methods is ready for use in practice.

2. One-step approximation. Let us introduce a discretization Δ_N of the interval $[t_o, T]$: $\Delta_N = \{t_k : k = 0, 1, \dots, N; t_o < t_1 < \dots < t_N = T\}$; the time increment

$h = t_{k+1} - t_k$; the approximation X_k or $\bar{X}(t_k)$ of the exact solution $X(t_k)$ of the system (1.1); the mean value $E\xi$ of a random variable ξ ; operators

$$L = L_1 + \varepsilon^2 L_2, \quad L_1 = \frac{\partial}{\partial t} + \left(a, \frac{\partial}{\partial x} \right) = \frac{\partial}{\partial t} + \sum_{i=1}^n a^i \frac{\partial}{\partial x^i},$$

$$L_2 = \frac{1}{2} \sum_{r=1}^q \left(\sigma_r, \frac{\partial}{\partial x} \right)^2 = \frac{1}{2} \sum_{r=1}^q \sum_{i,j=1}^n \sigma_r^i \sigma_r^j \frac{\partial^2}{\partial x^i \partial x^j},$$

$$\Lambda_r = \left(\sigma_r, \frac{\partial}{\partial x} \right) = \sum_{i=1}^n \sigma_r^i \frac{\partial}{\partial x^i},$$

and Ito integrals

$$I_{i_1, \dots, i_j}(F, t, h) = \int_t^{t+h} dW_{i_j}(\theta) \int_t^\theta dW_{i_{j-1}}(\theta_1) \int_t^{\theta_1} \dots \int_t^{\theta_{j-2}} F(\theta_{j-1}) dW_{i_1}(\theta_{j-1}),$$

where i_1, \dots, i_j are from the set of numbers $\{0, 1, \dots, q\}$ and $dW_o(\theta_r)$ designates $d\theta_r$, $F(\theta)$ is a deterministic (for simplicity continuous) function; $I_{i_1, i_2, \dots, i_j}(t, h) \equiv I_{i_1, i_2, \dots, i_j}(1(\cdot), t, h)$, where $1(\theta)$ is the function which is everywhere equal to one.

It is known [6], [7] that the Ito integrals have the properties

$$E I_{i_1, \dots, i_j} = 0 \text{ if at least one of the indices } i_k \neq 0,$$

$$E I_{i_1, \dots, i_j} = O(h^j) \text{ if all indices } i_k = 0,$$

$$(2.1) \quad (E (I_{i_1, \dots, i_j})^2)^{1/2} = O(h^r), \quad r = l_1 + l_2/2,$$

where l_1 is the number of zero indices i_k and l_2 is the number of nonzero indices i_k .

We assume that restrictions on the coefficients of the system (1.1) ensure the existence and uniqueness of the solution on the whole time interval $[t_o, T]$. Moreover, for construction of high-order methods the coefficients must be sufficiently smooth functions. Note that an initial value X_o of the system (1.1) may be equal to a deterministic value or a random variable which does not depend on the Wiener process.

Let us consider the one-step approximation with the local order 2 [6], [7]:

$$\begin{aligned} \bar{X}(t+h) &= X(t) + \varepsilon \sum_{r=1}^q \sigma_r(t, X(t)) I_r(t, h) + a(t, X(t)) h \\ &+ \varepsilon^2 \sum_{i,r=1}^q \Lambda_r \sigma_i(t, X(t)) I_{ri}(t, h) + \varepsilon \sum_{r=1}^q L_1 \sigma_r(t, X(t)) I_{or}(t, h) \\ &+ \varepsilon^3 \sum_{r=1}^q L_2 \sigma_r(t, X(t)) I_{or}(t, h) + \varepsilon \sum_{r=1}^q \Lambda_r a(t, X(t)) I_{ro}(t, h) \\ &+ \varepsilon^3 \sum_{s,i,r=1}^q \Lambda_s \Lambda_i \sigma_r(t, X(t)) I_{sir}(t, h) + L_1 a(t, X(t)) h^2/2 \\ (2.2) \quad &+ \varepsilon^2 L_2 a(t, X(t)) h^2/2. \end{aligned}$$

The remainder $\rho = X(t+h) - \bar{X}(t+h)$ of this approximation has the form

$$\begin{aligned}
 \rho = & \varepsilon^4 \sum_{r,i,s,j=1}^q I_{risj}(\Lambda_r \Lambda_i \Lambda_s \sigma_j, t, h) + \varepsilon^2 \sum_{i,r=1}^q I_{oir}(L_1 \Lambda_i \sigma_r, t, h) \\
 & + \varepsilon^4 \sum_{i,r=1}^q I_{oir}(L_2 \Lambda_i \sigma_r, t, h) + \varepsilon^2 \sum_{i,r=1}^q I_{ior}(\Lambda_i L_1 \sigma_r, t, h) \\
 & + \varepsilon^4 \sum_{i,r=1}^q I_{ior}(\Lambda_i L_2 \sigma_r, t, h) + \varepsilon^2 \sum_{i,r=1}^q I_{iro}(\Lambda_i \Lambda_r a, t, h) \\
 & + \varepsilon^3 \sum_{r,i,s=1}^q I_{osir}(L_1 \Lambda_s \Lambda_i \sigma_r, t, h) + \varepsilon^5 \sum_{r,i,s=1}^q I_{osir}(L_2 \Lambda_s \Lambda_i \sigma_r, t, h) \\
 & + \varepsilon \sum_{r=1}^q I_{oor}(L_1^2 \sigma_r, t, h) + \varepsilon^3 \sum_{r=1}^q I_{oor}((L_1 L_2 + L_2 L_1) \sigma_r, t, h) \\
 & + \varepsilon^5 \sum_{r=1}^q I_{oor}(L_2^2 \sigma_r, t, h) + \varepsilon \sum_{r=1}^q I_{oro}(L_1 \Lambda_r a, t, h) \\
 & + \varepsilon^3 \sum_{r=1}^q I_{oro}(L_2 \Lambda_r a, t, h) + \varepsilon \sum_{r=1}^q I_{roo}(\Lambda_r L_1 a, t, h) \\
 & + \varepsilon^3 \sum_{r=1}^q I_{roo}(\Lambda_r L_2 a, t, h) + I_{ooo}(L_1^2 a, t, h) \\
 (2.3) \quad & + \varepsilon^2 I_{ooo}((L_1 L_2 + L_2 L_1) a, t, h) + \varepsilon^4 I_{ooo}(L_2^2 a, t, h).
 \end{aligned}$$

It is not difficult to obtain

$$E\rho = O(h^3), \quad (E\rho^2)^{1/2} = O(h^3 + \varepsilon^2 h^2).$$

The Ito integrals I_{ri} and I_{sir} of the method (2.2) cannot be easily simulated. But these integrals are multiplied by ε^2 and ε^3 , respectively. That is why they may be transferred to the remainder and the error of the approximation would still not be large. Further, if we transfer from (2.2) not only the terms with complicated Ito integrals but also the terms which are sufficiently small, we obtain the reduced one-step approximation

$$\begin{aligned}
 \bar{X}(t+h) = & X(t) + \varepsilon \sum_{r=1}^q \sigma_r(t, X(t)) I_r(t, h) + a(t, X(t)) h + \varepsilon \sum_{r=1}^q L_1 \sigma_r(t, X(t)) I_{or}(t, h) \\
 (2.4) \quad & + \varepsilon \sum_{r=1}^q \Lambda_r a(t, X(t)) I_{ro}(t, h) + L_1 a(t, X(t)) h^2/2,
 \end{aligned}$$

the remainder ρ_1 of which is equal to

$$\begin{aligned}
 \rho_1 = & \rho + \varepsilon^2 \sum_{i,r=1}^q \Lambda_r \sigma_i(t, X(t)) I_{ri}(t, h) + \varepsilon^3 \sum_{r=1}^q L_2 \sigma_r(t, X(t)) I_{or}(t, h) \\
 (2.5) \quad & + \varepsilon^3 \sum_{s,i,r=1}^q \Lambda_s \Lambda_i \sigma_r(t, X(t)) I_{sir}(t, h) + \varepsilon^2 L_2 a(t, X(t)) h^2/2,
 \end{aligned}$$

where ρ is taken from (2.3).

One can obtain

$$E \rho_1 = O(h^3 + \varepsilon^2 h^2),$$

$$(2.6) \quad (E\rho_1^2)^{1/2} = O(h^3 + \varepsilon h^{5/2} + \varepsilon^2 h^2 + \varepsilon^3 h^{3/2} + \varepsilon^2 h) = O(h^3 + \varepsilon^2 h).$$

The terms $\varepsilon h^{5/2}$, $\varepsilon^2 h^2$, and $\varepsilon^3 h^{3/2}$ of the second expression are omitted because they are not greater than $O(h^3 + \varepsilon^2 h)$. Of course, the order of the approximation (2.4) is less (the order is equal to one due to the term $\varepsilon^2 \sum_{i,r=1}^q \Lambda_r \sigma_i I_{ri}$ in ρ_1) than the order of the approximation (2.2), but the error of the approximation (2.4) has the small factor ε^2 at h . Thus, we obtain the one-step approximation (2.4) which has sufficiently small mean-square local error and is efficient as to the simulation of the used random variables.

Using (2.3)–(2.5), we construct a new approximation by transferring a part of the remainder to the approximation. In this connection we expand the term

$$\begin{aligned} I_{ooo}(L_1^2 a, t, h) &= \int_t^{t+h} \left(\int_t^\theta \left(\int_t^{\theta_1} L_1^2 a(\theta_2, X(\theta_2)) d\theta_2 \right) d\theta_1 \right) d\theta \\ &= L_1^2 a(t, X(t)) h^3 / 6 + \rho', \end{aligned}$$

where

$$E\rho' = O(h^4 + \varepsilon^2 h^3), \quad E[(\rho')^2]^{1/2} = O(h^4 + \varepsilon h^{7/2} + \varepsilon^2 h^3) = O(h^4 + \varepsilon^2 h^3).$$

New approximation has the form

$$(2.7) \quad \tilde{X}(t+h) = \bar{X}(t+h) + L_1^2 a(t, X(t)) h^3 / 6,$$

where $\bar{X}(t+h)$ is taken from (2.4). The remainder $\tilde{\rho}$ of the approximation (2.7) can be obtained from ρ_1 if we substitute ρ' instead of $h^3 I_{ooo}(L_1^2 a, t, h)$. It is clear that

$$E\tilde{\rho} = O(h^4 + \varepsilon^2 h^2), \quad E(\tilde{\rho}^2)^{1/2} = O(h^4 + \varepsilon^2 h).$$

Of course, the approximation (2.7) may be derived by the other way, for instance, from an approximation with local order 3, but the suggested way is the simplest.

In this section we have demonstrated the basic idea of the paper. In contrast to the general case smallness of terms of an approximation for a system with small noise and of its remainder depends not only on time increment h but also on small parameter ε . This circumstance, as shown above, allows us to construct new numerical methods by excluding complicated terms, for instance, multiple Ito integrals, from a method and including them in its remainder. New methods are efficient as to simulation of the used random variables and have small mean-square errors in the sense of product $\varepsilon^i h^j$. Moreover, the methods contain fewer terms with operators than the corresponding schemes for a general system.

3. The theorem on mean-square global estimate. Usually after reducing, estimates of the remainder of a concrete one-step approximation are sufficiently simple and often contain only two terms (for instance, see (2.6)). However, it may not be the case. For instance, the sum $h^3 + \varepsilon h^{3/2} + \varepsilon^2 h$ cannot be reduced. Detailed analysis of possible errors gives the form of estimates which is exactly used in Theorem 3.1; i.e., the conditions of Theorem 3.1 are natural. Let us give the notation: ρ is a local error

of a method, R is a global error (we also call it as method error, mean-square error, or error if it does not lead to misunderstanding), r_o is a natural number, S_1 is either an empty set or a subset of positive integers p which are less than r_o , and S_2 is either an empty set or a subset of positive integers and semi-integers q which are less than r_o , i.e.,

$$S_1 \subset \{p : 0 < p < r_o, p \text{ is an integer}\},$$

$$S_2 \subset \{q : 0 < q < r_o, q \text{ is either an integer or semi-integer}\}.$$

Below in Theorem 3.1 the sum $\sum_{p \in S_1} (\sum_{q \in S_2})$ must be replaced by zero if S_1 (S_2) is an empty set.

THEOREM 3.1. *If the inequalities*

$$(3.1) \quad |E\rho| = |E(X_{t,x}(t+h) - \bar{X}_{t,x}(t+h))| \leq K_1(1 + |x|^2)^{1/2} \\ \times \left(h^{r_o} + \sum_{p \in S_1} h^p \varepsilon^{2J_1(p)} \right),$$

$$(3.2) \quad (E\rho^2)^{1/2} = (E|X_{t,x}(t+h) - \bar{X}_{t,x}(t+h)|^2)^{1/2} \leq K_2(1 + |x|^2)^{1/2} \\ \times \left(h^{r_o} + \sum_{q \in S_2} h^q \varepsilon^{J_2(q)} \right),$$

where $\bar{X}_{t,x}(t+h)$ is an approximation of the exact solution $X_{t,x}(t+h)$ of the system (1.1) with initial condition $X(t) = \bar{X}(t) = x$, $J_1(p)$ and $J_2(q)$ are decreasing functions with natural values, and K_1 and K_2 are constants, are fulfilled, then

$$(3.3) \quad (E|X_{t_o, X_o}(t_k) - \bar{X}_{t_o, X_o}(t_k)|^2)^{1/2} \leq K_3(1 + E|X_o|^2)^{1/2} \\ \times \left(h^{r_o-1} + \sum_{p \in S_1} h^{p-1} \varepsilon^{2J_1(p)} + \sum_{q \in S_2} h^{q-1/2} \varepsilon^{J_2(q)} \right),$$

where the constant K_3 does not depend on discretization step h , parameter ε , $0 \leq \varepsilon \leq \varepsilon_o$, and $k = 1, \dots, N$; i.e., the method, corresponding to the one-step approximation $\bar{X}(t+h)$, gives the mean-square error evaluated by

$$(3.4) \quad (ER^2)^{1/2} = O \left(h^{r_o-1} + \sum_{p \in S_1} h^{p-1} \varepsilon^{2J_1(p)} + \sum_{q \in S_2} h^{q-1/2} \varepsilon^{J_2(q)} \right).$$

The proof of Theorem 3.1 is similar to the proof of the mean-square convergence theorem for a general system [6], [8], and here it is omitted.

According to Theorem 3.1, the method, which, for example, is based on the one-step approximation (2.4), has the mean-square global error estimated by $O(h^2 + \varepsilon^2 h + \varepsilon^2 h^{1/2}) = O(h^2 + \varepsilon^2 h^{1/2})$. Its error is sufficiently small because of small factor ε^2 at $h^{1/2}$.

From Theorem 3.1 it follows that if $r_o > 1$ and the set S_1 is either empty or every number p of S_1 is greater than one and the set S_2 is either empty or every number q of S_2 is greater than $1/2$, then the corresponding method converges. However, the primary meaning of Theorem 3.1 is not that it gives convergence order of a method but that it gives a method error in terms of h and ε .

3.1. Selection of time increment h depending on parameter ε . In practice the parameter ε is small but fixed, and we can usually choose only the step h . Nevertheless, asymptotic behavior of method error under $\varepsilon \rightarrow 0$, when h is chosen depending on ε , is interesting in many respects. And the inequality (3.3) makes such an analysis possible.

Let us choose time increment h so that $h = C\varepsilon^\alpha$, $\alpha > 0$. Then the error of a method can be estimated in powers of small parameter ε

$$(E|X_{t_o, X_o}(t_k) - \bar{X}_{t_o, X_o}(t_k)|^2)^{1/2} = O(\varepsilon^\beta),$$

where $\beta = \min\{\alpha(r_o - 1), \min_{p \in S_1}(\alpha(p - 1) + 2J_1(p)), \min_{q \in S_2}(\alpha(q - 1/2) + J_2(q))\}$. The parameter α and a method may be such that a certain term of this method is smaller than $O(\varepsilon^\beta)$. Such a term may be omitted and, in spite of this, the order of the method error does not change with respect to ε .

Let us analyze the method which is based on the one-step approximation (2.4). If $h = C\varepsilon^\alpha$, the mean-square global error of the method (2.4) is estimated by $O(\varepsilon^{2\alpha} + \varepsilon^{2+\alpha/2})$. Let us choose α to be equal to one. In this case the method error is estimated by $O(\varepsilon^2)$, the order of the terms $\varepsilon L_1 \sigma_r I_{or}$ and $\varepsilon \Lambda_r a I_{ro}$ is equal to $O(\varepsilon^{5/2})$, and their omission gives $O(\varepsilon^2)$ to the mean-square error. So, in the case of $\alpha = 1$ these terms may be omitted, and that does not lead to substantial increasing of the error. Thus, we obtain a new method

$$(3.5) \quad X_{k+1} = X_k + \varepsilon \sum_{r=1}^q (\sigma_r I_r)_k + a_k h + L_1 a_k h^2/2,$$

$$(ER^2)^{1/2} = O(h^2 + \varepsilon h + \varepsilon^2 h^{1/2}),$$

where $\sigma_{r_k} = \sigma_r(t_k, X_k)$, $a_k = a(t_k, X_k)$, $(I_r)_k = I_r(t_k, h)$. It is clear that if $h = C\varepsilon^\alpha$, where $\alpha \leq 1$ or $\alpha \geq 2$, the errors of the methods (2.4) and (3.5) have the same order with respect to ε . But if $h = C\varepsilon^\alpha$, $1 < \alpha < 2$, the method (3.5) has the lower order with respect to ε than the method (2.4).

3.2. $h - \varepsilon$ approach versus $\varepsilon - h$ approach. In the paper we construct numerical methods by $h - \varepsilon$ approach; for instance, see the methods (2.4) and (3.5). According to $h - \varepsilon$ approach, we expand the exact solution $X(t)$ of the system (1.1) in powers of time increment h and obtain an expansion which is similar to the stochastic Taylor-type expansion [4], [6], [7]. Then we regroup terms of the expansion with respect to their $h^i \varepsilon^j$ factors and decide which terms must be included in a method. Such a decision depends on the desired mean-square error of the method and on calculation complexity of an expansion term, especially on complexity of simulation of the used random variables.

$\varepsilon - h$ approach is based on another idea. At first, the exact solution of the system (1.1) is expanded in powers of small parameter ε , for instance,

$$(3.6) \quad \bar{X}(t) = X^o(t) + \varepsilon X^1(t),$$

$$R = X(t) - \bar{X}(t) = O(\varepsilon^2),$$

where $X^o(t)$ and $X^1(t)$ are found as the solutions of the original system under $\varepsilon = 0$ and its system of the first approximation

$$(3.7) \quad dX^o = a(t, X^o) dt, \quad X^o(0) = X_o,$$

$$(3.8) \quad dX^1 = a'_x(t, X^o) X^1 dt + \sum_{r=1}^q \sigma_r(t, X^o) dW_r, \quad X^1(0) = 0.$$

The system (3.7) is the system of deterministic differential equations for which, as is generally known, efficient high-order numerical methods exist, for example,

$$(3.9) \quad \begin{aligned} X_{k+1}^o &= X_k^o + a_k h + (a a'_x + a'_t)_k h^2 / 2, \\ X_o^o &= X_o, R_o = O(h^2), \end{aligned}$$

where $a_k = a(t_k, X_k^o)$, the $n \times n$ matrix $(a'_x)_k$ is equal to $\partial a(t_k, X_k^o) / \partial x$, n -vector $(a'_t)_k$ is equal to $\partial a(t_k, X_k^o) / \partial t$, and R_o is the error of the method. The system (3.8) is the system of stochastic differential equations with additive noises [6], [7]. The Euler method for the system (3.8) has the form

$$(3.10) \quad \begin{aligned} X_{k+1}^1 &= X_k^1 + \sum_{r=1}^q (\sigma_r I_r)_k + (a'_x X^1)_k h, \\ X_o^1 &= 0, \quad (E(R_1)^2)^{1/2} = O(h), \end{aligned}$$

where $\sigma_{r_k} = \sigma_r(t_k, X_k^o)$, $(a'_x)_k = \partial a(t_k, X_k^o) / \partial x$, and R_1 is the error of the method. So, we obtain the method (3.6), (3.9), (3.10) for numerical solution of the system (1.1) with the error $O(h^2 + \varepsilon^2)$.

One can see that $h - \varepsilon$ approach and $\varepsilon - h$ approach are essentially different. If time increment h tends to zero, a method, constructed by $\varepsilon - h$ approach, does not converge to the exact solution and converges to $X^o(t) + \varepsilon X^1(t)$. In contrast to $\varepsilon - h$ approach $h - \varepsilon$ approach gives a method which always converges to the exact solution of the system (1.1) in the case of $h \rightarrow 0$. Our aim is to derive numerical methods for solution of the system (1.1) with small but fixed parameter $\varepsilon > 0$. That is why $h - \varepsilon$ approach is more preferable than $\varepsilon - h$.

4. Some methods for Ito systems with small noises. Our aim is to construct methods with small mean-square errors (provided that ε is a small parameter) and with simply simulated random variables. Herein we restrict ourselves to the methods which contain the following Ito integrals:

$$(4.1) \quad \begin{aligned} I_r &= h^{1/2} \xi_r, \\ I_{ro} &= h^{3/2} (\eta_r / \sqrt{3} + \xi_r) / 2, \\ I_{or} &= h I_r - I_{ro}, \\ J_r &= \int_0^h \vartheta W_r(\vartheta) d\vartheta = h^{5/2} (\xi_r / 3 + \eta_r / (4\sqrt{3}) + \zeta_r / (12\sqrt{5})), \\ I_{roo} &= h I_{ro} - J_r, \\ I_{oro} &= 2 J_r - h I_{ro}, \\ I_{oor} &= h^2 I_r / 2 - J_r, \end{aligned}$$

where ξ_r, η_r, ζ_r are independent normally distributed $N(0, 1)$ random variables with zero mean and unit standard deviation. The used random variables (Ito integrals) of all proposed methods are simulated at each step according to the formulas (4.1).

4.1. Taylor-type numerical methods.

4.1.1. Method $O(h + \dots)$. The simplest numerical method is the Euler one:

$$(4.2) \quad \begin{aligned} X_{k+1} &= X_k + \varepsilon \sum_{r=1}^q (\sigma_r I_r)_k + a_k h, \\ E\rho &= O(h^2), \quad (E\rho^2)^{1/2} = O(h^2 + \varepsilon^2 h), \\ (ER^2)^{1/2} &= O(h + \varepsilon^2 h^{1/2}). \end{aligned}$$

4.1.2. Methods $O(h^2 + \dots)$ and $O(h^3 + \dots)$. These methods are based on the one-step approximations which have been written down above in sections 2 and 3. The mean-square error of the method (2.4) is equal to $O(h^2 + \varepsilon^2 h^{1/2})$. The mean-square error of the method (3.5) is estimated by $O(h^2 + \varepsilon h + \varepsilon^2 h^{1/2})$. The method, which is based on the one-step approximation (2.7), has the error $(ER^2)^{1/2} = O(h^3 + \varepsilon^2 h^{1/2})$.

4.1.3. Methods $O(h^4 + \dots)$. The following method is obtained:

$$\begin{aligned}
 X_{k+1} = & X_k + \varepsilon \sum_{r=1}^q (\sigma_r I_r)_k + a_k h + \varepsilon \sum_{r=1}^q (L_1 \sigma_r I_{or})_k + \varepsilon \sum_{r=1}^q (\Lambda_r a I_{ro})_k \\
 & + L_1 a_k h^2 / 2 + \varepsilon \sum_{r=1}^q (L_1^2 \sigma_r I_{oor})_k + \varepsilon \sum_{r=1}^q (L_1 \Lambda_r a I_{oro})_k \\
 (4.3) \quad & + \varepsilon \sum_{r=1}^q (\Lambda_r L_1 a I_{roo})_k + L_1^2 a_k h^3 / 6 + L_1^3 a_k h^4 / 24, \\
 E\rho = & O(h^5 + \varepsilon^2 h^2), \quad (E\rho^2)^{1/2} = O(h^5 + \varepsilon^2 h), \\
 (ER^2)^{1/2} = & O(h^4 + \varepsilon^2 h^{1/2}).
 \end{aligned}$$

In some cases the derived methods may be improved due to special properties of a concrete system. For instance, let us consider the commutative case, i.e., $\Lambda_i \sigma_r = \Lambda_r \sigma_i$, or a system with one noise ($q = 1$). For these systems we obtain

$$\begin{aligned}
 X_{k+1} = & A(t_k, X_k, h; (\xi, \eta, \zeta)_k) + \varepsilon^2 \sum_{i=1}^{q-1} \sum_{r=i+1}^q (\Lambda_i \sigma_r I_i I_r)_k \\
 (4.4) \quad & + \varepsilon^2 \sum_{i=1}^q (\Lambda_i \sigma_i (I_i^2 - h) / 2)_k + \varepsilon^2 L_2 a_k h^2 / 2, \\
 E\rho = & O(h^5 + \varepsilon^2 h^3), \quad (E\rho^2)^{1/2} = O(h^5 + \varepsilon^2 h^2 + \varepsilon^3 h^{3/2}), \\
 (ER^2)^{1/2} = & O(h^4 + \varepsilon^2 h^{3/2} + \varepsilon^3 h),
 \end{aligned}$$

where $A(t_k, X_k, h; (\xi, \eta, \zeta)_k)$ is equal to the right side of (4.3).

Note that for the system with one noise ($q = 1$) the term $\varepsilon^2 \sum_{i=1}^{q-1} \sum_{r=i+1}^q (\Lambda_i \sigma_r I_i I_r)_k$ is neglected. One can see that the error of the method (4.4) is smaller than the error of the scheme (4.3). Moreover, the mean-square order of the method (4.4) is equal to one, while the mean-square order of the method (4.3) is equal to one-half.

4.2. Runge–Kutta methods. To reduce calculations of derivatives in the methods of section 4.1 we propose Runge–Kutta schemes (of course, they are not genuine Runge–Kutta methods because they need calculation of some derivatives).

4.2.1. Method $O(h^2 + \dots)$. The following method is obtained:

$$\begin{aligned}
 X_{k+1} = & X_k + \varepsilon \sum_{r=1}^q (\sigma_r I_r)_k + \left(a \left(t_k + h, X_k + \varepsilon \sum_{r=1}^q (\sigma_r I_r)_k + a_k h \right) + a_k \right) h / 2 \\
 (4.5) \quad & + \varepsilon \sum_{r=1}^q (L_1 \sigma_r I_{or})_k + \varepsilon \sum_{r=1}^q (\Lambda_r a (I_{ro} - I_r h / 2))_k, \\
 E\rho = & O(h^3), \quad (E\rho^2)^{1/2} = O(h^3 + \varepsilon^2 h), \\
 (ER^2)^{1/2} = & O(h^2 + \varepsilon^2 h^{1/2}).
 \end{aligned}$$

4.2.2. Method $O(h^3 + \dots)$. The following Runge–Kutta method is obtained:

$$\begin{aligned}
 X_{k+1} &= X_k + (K_1 + 4K_2 + K_3)/6 + \varepsilon \sum_{r=1}^q (\sigma_r I_r)_k \\
 (4.6) \quad &+ \varepsilon \sum_{r=1}^q (L_1 \sigma_r I_{or})_k + \varepsilon \sum_{r=1}^q (\Lambda_r a I_{ro})_k, \\
 &(ER^2)^{1/2} = O(h^3 + \varepsilon^2 h^{1/2}),
 \end{aligned}$$

where

$$K_1 = ha(t_k, X_k), \quad K_2 = ha(t_k + h/2, X_k + K_1/2), \quad K_3 = ha(t_{k+1}, X_k - K_1 + 2K_2).$$

4.2.3. Methods $O(h^4 + \dots)$. The following Runge–Kutta method is obtained:

$$\begin{aligned}
 X_{k+1} &= X_k + (K_1 + 2K_2 + 2K_3 + K_4)/6 + \varepsilon \sum_{r=1}^q (\sigma_r I_r)_k + \varepsilon \sum_{r=1}^q (L_1 \sigma_r I_{or})_k \\
 (4.7) \quad &+ \varepsilon \sum_{r=1}^q (\Lambda_r a I_{ro})_k + \varepsilon \sum_{r=1}^q (L_1^2 \sigma_r I_{oor})_k + \varepsilon \sum_{r=1}^q (L_1 \Lambda_r a I_{oro})_k + \varepsilon \sum_{r=1}^q (\Lambda_r L_1 a I_{roo})_k, \\
 &(ER^2)^{1/2} = O(h^4 + \varepsilon^2 h^{1/2}),
 \end{aligned}$$

where

$$\begin{aligned}
 K_1 &= ha(t_k, X_k), & K_2 &= ha(t_k + h/2, X_k + K_1/2), \\
 K_3 &= ha(t_k + h/2, X_k + K_2/2), & K_4 &= ha(t_{k+1}, X_k + K_3).
 \end{aligned}$$

In the commutative case the method (4.7) can be improved as in section 4.1.3.

The simpler method

$$(4.8) \quad X_{k+1} = X_k + (K_1 + 2K_2 + 2K_3 + K_4)/6 + \varepsilon \sum_{r=1}^q (\sigma_r I_r)_k,$$

where K_i are calculated as in (4.7), has the greater error in comparison with (4.7).

$$(ER^2)^{1/2} = O(h^4 + \varepsilon h + \varepsilon^2 h^{1/2}).$$

Note that the error of the method (4.8) with respect to ε is greater than the error of the scheme (4.7) under $h = C\varepsilon^\alpha$, $1/3 < \alpha < 2$, and otherwise they have the same order.

4.3. Implicit methods. Implicit methods are useful for stiff stochastic systems.

4.3.1. Methods $O(h + \dots)$. The one-parameter family of implicit Euler schemes is written in the form [6], [7]

$$\begin{aligned}
 (4.9) \quad X_{k+1} &= X_k + \varepsilon \sum_{r=1}^q (\sigma_r I_r)_k + \alpha ha_k + (1 - \alpha) ha_{k+1}, \\
 0 \leq \alpha \leq 1, & \quad (ER^2)^{1/2} = O(h + \varepsilon^2 h^{1/2}).
 \end{aligned}$$

4.3.2. Methods $O(h^2 + \dots)$. As in the monograph [6], the two-parameter family of implicit schemes is constructed:

$$\begin{aligned}
 X_{k+1} &= X_k + \varepsilon \sum_{r=1}^q (\sigma_r I_r)_k + \alpha h a_k + (1 - \alpha) h a_{k+1} + \varepsilon \sum_{r=1}^q (\Lambda_r a(I_{r_o} - (1 - \alpha) I_r h))_k \\
 (4.10) \quad &+ \varepsilon \sum_{r=1}^q (L_1 \sigma_r I_{or})_k + \beta(2\alpha - 1) L_1 a_k h^2 / 2 + (1 - \beta)(2\alpha - 1) L_1 a_{k+1} h^2 / 2, \\
 &0 \leq \alpha \leq 1, \quad 0 \leq \beta \leq 1, \\
 E\rho &= O(h^3 + \varepsilon^2 h^2), \quad (E\rho^2)^{1/2} = O(h^3 + \varepsilon^2 h), \\
 (ER^2)^{1/2} &= O(h^2 + \varepsilon^2 h^{1/2}).
 \end{aligned}$$

If $\alpha = 1/2$, we obtain the trapezoidal method which is the simplest of the family (4.10):

$$\begin{aligned}
 X_{k+1} &= X_k + \varepsilon \sum_{r=1}^q (\sigma_r I_r)_k + h(a_k + a_{k+1})/2 \\
 (4.11) \quad &+ \varepsilon \sum_{r=1}^q (L_1 \sigma_r I_{or})_k + \varepsilon \sum_{r=1}^q (\Lambda_r a(I_{r_o} - I_r h/2))_k, \\
 (ER^2)^{1/2} &= O(h^2 + \varepsilon^2 h^{1/2}).
 \end{aligned}$$

In the commutative case or in the case of one noise the methods (4.10)–(4.11) can be improved as the method (4.3) in section 4.1.

4.4. Remark. Obviously, many other methods may be derived. First, by adding or omitting some terms one can obtain methods that are similar to above but have other mean-square errors, for instance, $O(h^5 + \dots)$, $O(h^6 + \dots)$. Second, it is possible to derive other types of methods, for instance, implicit Runge–Kutta methods. In this section we have restricted ourselves to the set of more common and, in our opinion, useful methods and have illustrated the proposed approach to numerical solution of a stochastic system with small noises.

5. Stratonovich stochastic differential equations with small noises. For some physical applications Stratonovich interpretation of a stochastic system is preferable [1], [2]. It is known that the stochastic system in the Stratonovich sense (marked by “*”)

$$(5.1) \quad dX = a(t, X) dt + \varepsilon \sum_{r=1}^q \sigma_r(t, X) * dW_r, \quad X(t_o) = X_o$$

is equivalent to the following system of the Ito stochastic differential equations:

$$\begin{aligned}
 (5.2) \quad dX &= \left[a(t, X) + \frac{\varepsilon^2}{2} \sum_{r=1}^q \frac{\partial \sigma_r}{\partial x}(t, X) \sigma_r(t, X) \right] dt + \varepsilon \sum_{r=1}^q \sigma_r(t, X) dW_r, \\
 X(t_o) &= X_o.
 \end{aligned}$$

In the general case ($\varepsilon = 1$) numerical methods, constructed for the Ito system, are easily rewritten for the Stratonovich system by adding the term $\frac{1}{2} \sum_{r=1}^q \frac{\partial \sigma_r}{\partial x} \sigma_r$ to the drift [6], [7]. However, in the case of small noises the additional term is multiplied by

small factor ε^2 . So, the Stratonovich system with small noises (5.1) is distinguished from the Ito system $dX = a(t, X)dt + \varepsilon \sum_{r=1}^q \sigma_r(t, X)dW_r$ by the small component in the drift, and constructing a numerical method for the system (5.2), one must take the magnitude of the additional term into account.

Most of the methods for the system (5.1) are obtained from methods for the Ito system by adding the term $\frac{\varepsilon^2}{2} \sum_{r=1}^q (\frac{\partial \sigma_r}{\partial x} \sigma_r)h$. Namely, for the Stratonovich system (5.1) the methods (2.4), (2.7), (3.5), (4.2), (4.3), (4.5), (4.6), (4.7), (4.8) acquire the form

$$(5.3) \quad X_{k+1} = A(t_k, X_k, h; (\xi, \eta, \zeta)_k) + \frac{\varepsilon^2}{2} \sum_{r=1}^q \left(\frac{\partial \sigma_r}{\partial x} \sigma_r \right)_k h,$$

where the expression $A(t_k, X_k, h; (\xi, \eta, \zeta)_k)$ is calculated according to the same rules as the right sides of the corresponding methods.

In the commutative case we obtain

$$(5.4) \quad \begin{aligned} X_{k+1} = & A(t_k, X_k, h; (\xi, \eta, \zeta)_k) + \frac{\varepsilon^2}{2} \sum_{r=1}^q \left(\frac{\partial \sigma_r}{\partial x} \sigma_r \right)_k h + \varepsilon^2 \sum_{i=1}^{q-1} \sum_{r=i+1}^q (\Lambda_i \sigma_r I_i I_r)_k \\ & + \varepsilon^2 \sum_{i=1}^q (\Lambda_i \sigma_i (I_i^2 - h)/2)_k + \varepsilon^2 L_1 \left\{ \sum_{r=1}^q \frac{\partial \sigma_r}{\partial x} \sigma_r \right\}_k h^2/4 + \varepsilon^2 (\tilde{L}_2 a)_k h^2/2, \\ & (ER^2)^{1/2} = O(h^4 + \varepsilon^2 h^{3/2} + \varepsilon^3 h), \end{aligned}$$

where $A(t_k, X_k, h; (\xi, \eta, \zeta)_k)$ is the right side of (4.3) and

$$\tilde{L}_2 = \frac{1}{2} \sum_{r=1}^q \left(\sigma_r, \frac{\partial}{\partial x} \right)^2 + \frac{1}{2} \sum_{r=1}^q \left(\frac{\partial \sigma_r}{\partial x} \sigma_r, \frac{\partial}{\partial x} \right).$$

Analogously, the method for the Stratonovich system, which corresponds to the method (4.7) for the Ito system, can be improved in the commutative case.

The one-parameter family of implicit methods for the Stratonovich system has the form

$$(5.5) \quad X_{k+1} = A(t_k, X_k, h; (\xi)_k) + \frac{\varepsilon^2}{2} \sum_{r=1}^q \left[\alpha \left(\frac{\partial \sigma_r}{\partial x} \sigma_r \right)_k + (1 - \alpha) \left(\frac{\partial \sigma_r}{\partial x} \sigma_r \right)_{k+1} \right] h,$$

$$(ER^2)^{1/2} = O(h + \varepsilon^2 h^{1/2}),$$

where $A(t_k, X_k, h; (\xi)_k)$ is the right side of (4.9).

The two-parameter family of implicit methods for the Stratonovich system has the form

$$(5.6) \quad X_{k+1} = A(t_k, X_k, h; (\xi, \eta)_k) + \frac{\varepsilon^2}{2} \sum_{r=1}^q \left[\alpha \left(\frac{\partial \sigma_r}{\partial x} \sigma_r \right)_k + (1 - \alpha) \left(\frac{\partial \sigma_r}{\partial x} \sigma_r \right)_{k+1} \right] h,$$

$$(ER^2)^{1/2} = O(h^2 + \varepsilon^2 h^{1/2}),$$

where $A(t_k, X_k, h; (\xi, \eta)_k)$ is the right side of (4.10).

6. Numerical methods for systems with small additive noises. One of the important particular cases of the system (1.1) is the system with additive noises

$$(6.1) \quad dX = a(t, X) dt + \varepsilon \sum_{r=1}^q \sigma_r(t) dW_r.$$

Note that in this case the Stratonovich system coincides with the Ito system.

For the system (6.1) we obtain the Taylor-type and Runge–Kutta methods with the errors $O(h^2 + \varepsilon h)$, $O(h^2 + \varepsilon^2 h)$, $O(h^2 + \varepsilon^2 h^{3/2})$, $O(h^3 + \varepsilon^2 h^{3/2})$, $O(h^4 + \varepsilon^2 h^{3/2})$ which are similar to the methods of section 4 but always require additional analysis of the errors. We also obtain the implicit schemes which follow from the schemes of section 4.3. Herein we restrict ourselves to the methods with errors like $O(h^4 + \dots)$ which, from our point of view, are the most interesting for a reader.

The Taylor-type method with the error $O(h^4 + \varepsilon^2 h^{3/2})$ is

$$(6.2) \quad \begin{aligned} X_{k+1} = X_k + \varepsilon \sum_{r=1}^q (\sigma_r I_r)_k + a_k h + \varepsilon \sum_{r=1}^q \left(\frac{d\sigma_r}{dt} I_{or} \right)_k + \varepsilon \sum_{r=1}^q (\Lambda_r a I_{ro})_k \\ + (L_1 + \varepsilon^2 L_2) a_k h^2 / 2 + \varepsilon \sum_{r=1}^q \left(\frac{d^2 \sigma_r}{dt^2} I_{oor} \right)_k + \varepsilon \sum_{r=1}^q (L_1 \Lambda_r a I_{oro})_k \\ + \varepsilon \sum_{r=1}^q (\Lambda_r L_1 a I_{roo})_k + L_1^2 a_k h^3 / 6 + L_1^3 a_k h^4 / 24, \\ E\rho = O(h^5 + \varepsilon^2 h^3), \quad (E\rho^2)^{1/2} = O(h^5 + \varepsilon^2 h^2), \\ (ER^2)^{1/2} = O(h^4 + \varepsilon^2 h^{3/2}). \end{aligned}$$

The Runge–Kutta method with the error $O(h^4 + \varepsilon^2 h^{3/2})$ is

$$(6.3) \quad \begin{aligned} X_{k+1} = X_k + (K_1 + 2K_2 + 2K_3 + K_4) / 6 + \varepsilon \sum_{r=1}^q (\sigma_r I_r)_k + \varepsilon \sum_{r=1}^q \left(\frac{d\sigma_r}{dt} I_{or} \right)_k + \varepsilon \sum_{r=1}^q (\Lambda_r a I_{ro})_k \\ + \varepsilon^2 L_2 a_k h^2 / 2 + \varepsilon \sum_{r=1}^q \left(\frac{d^2 \sigma_r}{dt^2} I_{oor} \right)_k + \varepsilon \sum_{r=1}^q (L_1 \Lambda_r a I_{oro})_k + \varepsilon \sum_{r=1}^q (\Lambda_r L_1 a I_{roo})_k, \\ (ER^2)^{1/2} = O(h^4 + \varepsilon^2 h^{3/2}), \end{aligned}$$

where

$$K_1 = ha(t_k, X_k), \quad K_2 = ha(t_k + h/2, X_k + K_1/2),$$

$$K_3 = ha(t_k + h/2, X_k + K_2/2), \quad K_4 = ha(t_{k+1}, X_k + K_3).$$

As in section 4.2 (see the methods (4.7)–(4.8)), it is possible to obtain the simpler method which in the case of additive noises coincides with the scheme (4.8), but its mean-square error is equal to $(ER^2)^{1/2} = O(h^4 + \varepsilon h)$.

7. Numerical methods for systems with small colored noises. It is known that for some physical applications colored noises are more preferable than white ones. In [9], [10] various special numerical methods for the solution of a system with colored noises were derived. Here we present schemes for a system with small colored noises. Thanks to small parameter ε , they are simpler and have smaller errors than in the case of a general system with colored noises.

A system with small colored noises can be written in the form

$$\begin{aligned}
 dY &= f(t, Y) dt + \varepsilon G(t, Y) Z dt, \\
 dZ &= A(t) Z dt + \sum_{r=1}^q b_r(t) dW_r, \\
 Y(t_o) &= Y_o, \quad Z(t_o) = Z_o, \quad t \in [t_o, T],
 \end{aligned}
 \tag{7.1}$$

where Y and $f(t, Y)$ are l -dimensional vectors, Z and $b_r(t)$ are m -dimensional vectors, $A(t)$ is an $m \times m$ matrix, and $G(t, Y)$ is an $l \times m$ matrix, W_r are uncorrelated standard Wiener processes, and ε is a small parameter.

Let us introduce a new variable U :

$$U = \varepsilon Z. \tag{7.2}$$

Then the system (7.1) is rewritten in the convenient form

$$\begin{aligned}
 dY &= f(t, Y) dt + G(t, Y) U dt, \\
 dU &= A(t) U dt + \varepsilon \sum_{r=1}^q b_r(t) dW_r, \\
 Y(t_o) &= Y_o, \quad U(t_o) = \varepsilon Z_o.
 \end{aligned}
 \tag{7.3}$$

The system (7.3) is the particular case of the system with small additive noises (6.1). However, the system (7.3) is simpler than (6.1) because it is linear with respect to U and the first equation of (7.3) does not contain Wiener differentials. These properties allow us to construct special numerical methods for the system (7.3) which are simpler and have smaller errors than the corresponding schemes for the system (6.1).

The operators L_1 , L_2 , and Λ_r for the system (7.3) have the form

$$\begin{aligned}
 L_1 &= \frac{\partial}{\partial t} + \left(f(t, Y) + G(t, Y) U, \frac{\partial}{\partial y} \right) + \left(A(t) U, \frac{\partial}{\partial u} \right), \\
 L_2 &= \frac{1}{2} \sum_{r=1}^q \sum_{i,j=1}^m b_r^i b_r^j \frac{\partial^2}{\partial u^i \partial u^j}, \quad \Lambda_r = \left(b_r, \frac{\partial}{\partial u} \right).
 \end{aligned}
 \tag{7.4}$$

On the base of the methods for a system with additive noises we obtain various methods (Taylor-type and Runge–Kutta, explicit and implicit) with errors from $O(h^2 + \varepsilon h)$ up to $O(h^4 + \varepsilon h^3 + \varepsilon^2 h^{5/2})$ for the system (7.3). But herein we restrict ourselves to the Runge–Kutta schemes with the errors like $O(h^4 + \dots)$.

For the system (7.3) we obtain the Runge–Kutta method

$$Y_{k+1} = Y_k + (K_1 + 2K_2 + 2K_3 + K_4)h/6 + \varepsilon G_k \sum_{r=1}^q (b_r I_{r\alpha})_k + \varepsilon \sum_{r=1}^q ((Gb_r)'_t$$

$$\begin{aligned}
 &+(Gb_r)'_y(f + GU)]I_{oro})_k + \varepsilon \sum_{r=1}^q (\Lambda_r L_1(f + GU)I_{roo})_k + \varepsilon^2 \sum_{r=1}^q ((Gb_r)'_y Gb_r)_k h^3/6, \\
 U_{k+1} &= U_k + \varepsilon \sum_{r=1}^q (b_r I_r)_k + A_k U_k h + \varepsilon \sum_{r=1}^q \left(\frac{db_r}{dt} I_{or} \right)_k + \varepsilon \sum_{r=1}^q (Ab_r I_{ro})_k \\
 &+ (A'_t U + A^2 U)_k h^2/2 + \varepsilon \sum_{r=1}^q \left(\frac{d^2 b_r}{dt^2} I_{oor} \right)_k + \varepsilon \sum_{r=1}^q ((Ab_r)'_t I_{oro})_k \\
 &+ \varepsilon \sum_{r=1}^q ((A'_t b_r + A^2 b_r)I_{roo})_k + (A''_{tt} U + AA'_t U + A^3 U + 2A'_t A U)_k h^3/6 \\
 (7.5) \quad &+ \left\{ \left(\frac{\partial}{\partial t} + \left(Au, \frac{\partial}{\partial u} \right) \right) (A''_{tt} U + AA'_t U + A^3 U + 2A'_t A U) \right\}_k h^4/24, \\
 &(ER^2)^{1/2} = O(h^4 + \varepsilon h^3 + \varepsilon^2 h^{5/2}),
 \end{aligned}$$

where

$$\begin{aligned}
 F(t, Y, U) &= f(t, Y) + G(t, Y)U, \quad A_{k+1/2} = A(t_k + h/2), \\
 K_1 &= hF(t_k, Y_k, U_k), \quad K_2 = hF(t_k + h/2, Y_k + K_1/2, U_k + hA_k U_k/2), \\
 K_3 &= hF(t_k + h/2, Y_k + K_2/2, U_k + hA_{k+1/2}(U_k + hA_k U_k/2)/2), \\
 K_4 &= hF(t_{k+1}, Y_k + K_3, U_k + hA_{k+1/2}[U_k + hA_{k+1/2}(U_k + hA_k U_k/2)/2]).
 \end{aligned}$$

The method (7.5) may be improved up to $(ER^2)^{1/2} = O(h^4 + \varepsilon^2 h^{5/2})$ by adding terms with the order $\varepsilon h^{7/2}$.

The simpler method is

$$\begin{aligned}
 (7.6) \quad Y_{k+1} &= Y_k + (K_1 + 2K_2 + 2K_3 + K_4)h/6, \\
 U_{k+1} &= U_k + \varepsilon \sum_{r=1}^q (b_r I_r)_k + A_k U_k h + (A'_t U + A^2 U)_k h^2/2 \\
 &+ (A''_{tt} U + AA'_t U + A^3 U + 2A'_t A U)_k h^3/6 \\
 &+ \left\{ \left(\frac{\partial}{\partial t} + \left(Au, \frac{\partial}{\partial u} \right) \right) (A''_{tt} U + AA'_t U + A^3 U + 2A'_t A U) \right\}_k h^4/24, \\
 &(ER^2)^{1/2} = O(h^4 + \varepsilon h),
 \end{aligned}$$

where K_i are taken from (7.5). Note that if, for example, $h = C\varepsilon^{1/4}$, the errors of both methods (7.5) and (7.6) are estimated by $O(\varepsilon)$. However, if $h = C\varepsilon$, the method (7.5) gives $(ER^2)^{1/2} = O(\varepsilon^4)$ and the method (7.6) gives only $O(\varepsilon^2)$.

8. Numerical tests.

8.1. Simulation of Lyapunov exponent of a linear system with small noises. The stability problem of a stochastic system is of great importance from physical and engineering points of view. It is known [11], [12] that one can investigate stability of a dynamical stochastic system by Lyapunov exponents. The negativeness of upper Lyapunov exponents is an indication of system stability. It is usually impossible to derive analytical expressions for Lyapunov exponents. In this case numerical approaches are useful. For the first time an algorithm of numerical computation of Lyapunov exponents was proposed by D. Talay [13]. The algorithm is based on weak schemes.

Here we calculate Lyapunov exponent as a convenient example to illustrate the correctness and effectiveness (in comparison with ordinary mean-square schemes) of the proposed methods. Although the weak schemes are usually more efficient than mean-square ones, our approach is interesting in itself because we find the exponent together with the real trajectory.

Let us consider the following two-dimensional linear Ito stochastic system:

$$(8.1) \quad dX = AXdt + \varepsilon \sum_{r=1}^q B_r X dW_r(t),$$

where X is a two-dimensional vector, A and B_r are constant 2×2 matrices, W_r are independent standard Wiener processes, and $\varepsilon > 0$ is a small parameter. In ergodic case the unique Lyapunov exponent λ of the system (8.1) exists [11] and

$$(8.2) \quad \lambda = \lim_{t \rightarrow \infty} \frac{1}{t} E(\ln|X(t)|) = \lim_{t \rightarrow \infty} \frac{1}{t} \ln|X(t)|,$$

where $X(t)$, $t \geq 0$, is a nontrivial solution of the system (8.1). The last equality of (8.2) holds with the probability one. A nontrivial solution of the system (8.1) is asymptotically stable with probability one if and only if the Lyapunov exponent λ is negative [11].

In [12], [14] the expansion of Lyapunov exponent of the system (8.1) in powers of the small parameter ε was obtained. In the case of

$$(8.3) \quad A = \begin{bmatrix} a & c \\ -c & a \end{bmatrix}, \quad B_r = \begin{bmatrix} b_r & d_r \\ -d_r & b_r \end{bmatrix}, \quad r = 1, \dots, q,$$

the Lyapunov exponent of the system (8.1) is exactly equal to [14]

$$(8.4) \quad \lambda = a + \frac{\varepsilon^2}{2} \sum_{r=1}^q [(d_r)^2 - (b_r)^2].$$

To test the numerical schemes of the present paper we choose the case (8.3) of the system (8.1) with two independent noises ($q = 2$).

We calculate the function $\lambda(t)$

$$(8.5) \quad \lambda(t) = \frac{1}{t} \ln|X(t)| \approx \frac{1}{t} \ln|\bar{X}(t)|$$

which in the limit of large time ($t \rightarrow \infty$) tends to the Lyapunov exponent λ . The approximation $\bar{X}(t)$ of the exact solution $X(t)$ of the system (8.1) is simulated by

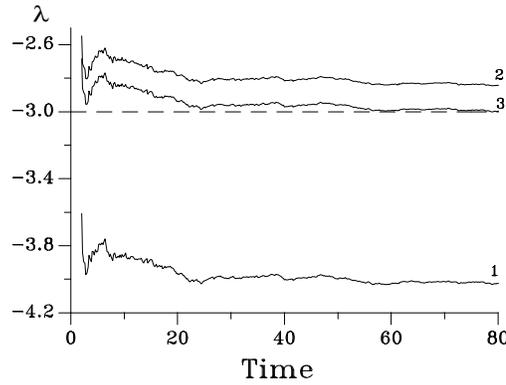


FIG. 1. *Lyapunov exponent. Time dependence of the function $\lambda(t)$, $t \geq 2$, for $a = -3$, $c = 1$, $b_1 = b_2 = 1$, $d_1 = 1$, $d_2 = -1$, $\varepsilon = 0.1$, $X^1(0) = 0$, $X^2(0) = 1$, and time step $h = 0.3$. The solution of the system (8.1)–(8.3) is approximated by (1) the method with the error $O(h)$, (2) the Runge–Kutta method with the error $O(h^2 + \varepsilon h + \varepsilon^2 h^{1/2})$, and (3) the Runge–Kutta method with the error $O(h^4 + \varepsilon h + \varepsilon^2 h^{1/2})$. Dashed line is the exact value of the Lyapunov exponent λ ($\lambda = -3$).*

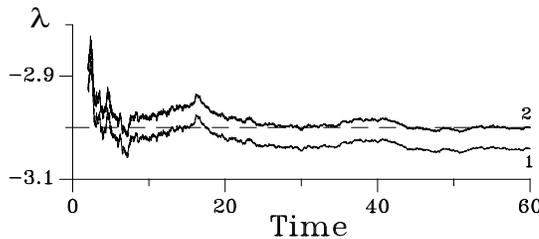


FIG. 2. *Lyapunov exponent. Time dependence of the function $\lambda(t)$, $t \geq 2$, for $h = 0.01$, other parameters are the same as in Fig. 1. The solution of the system (8.1)–(8.3) is approximated by (1) the method with the error $O(h)$ and (2) the Runge–Kutta methods with the errors $O(h^2 + \varepsilon h + \varepsilon^2 h^{1/2})$ and $O(h^4 + \varepsilon h + \varepsilon^2 h^{1/2})$. Dashed line is the exact value of the Lyapunov exponent λ ($\lambda = -3$).*

three mean-square schemes: (i) the first-order method [3], [6], [7] with the error $O(h)$ which in our case is efficient as to simulation of the used random variables due to commutativity of the matrices B_r , $r = 1, \dots, q$; (ii) the simplified version of the Runge–Kutta scheme (4.5) with the error $O(h^2 + \varepsilon h + \varepsilon^2 h^{1/2})$; and (iii) the Runge–Kutta scheme (4.8) with the error $O(h^4 + \varepsilon h + \varepsilon^2 h^{1/2})$.

If one chooses $h = O(\varepsilon^{1/2})$ (see Fig. 1), the method (i) gives the error $O(\varepsilon^{1/2})$, the method (ii) gives the error $O(\varepsilon)$, and the method (iii) gives the error $O(\varepsilon^{3/2})$. For $h = O(\varepsilon^2)$ (see Fig. 2) we have (i) gives $O(\varepsilon^2)$, and (ii), (iii) give $O(\varepsilon^3)$. Analyzing Figures 1 and 2, one can conclude that in the case of small noises new methods may have smaller errors than ordinary methods and permit us to save CPU time.

To simulate Gaussian random numbers we use the procedure GASDEV [15].

8.2. Laser Langevin equation with multiplicative noises. Our second example is devoted to trajectory simulation of the following laser Langevin equation [16], [17]:

$$(8.6) \quad db/dt = [(\alpha + i\beta) - (A + iB)|b|^2]b + \Gamma(t),$$

where α , β , and Γ fluctuate according to

$$\alpha = \alpha_o + \Gamma_\alpha(t), \quad \beta = \beta_o + \Gamma_\beta(t), \quad \Gamma(t) = \Gamma_1(t) + i\Gamma_2(t),$$

$$\langle \Gamma_\alpha \rangle = \langle \Gamma_\beta \rangle = \langle \Gamma_j \rangle = 0, \quad j = 1, 2, \quad \langle \Gamma_\alpha(t) \Gamma_\alpha(t') \rangle = Q_\alpha \delta(t - t'),$$

$$\langle \Gamma_\beta(t) \Gamma_\beta(t') \rangle = Q_\beta \delta(t - t'), \quad \langle \Gamma_\alpha(t) \Gamma_\beta(t') \rangle = Q_{\alpha\beta} \delta(t - t'),$$

$$\langle \Gamma_j(t) \Gamma_l(t') \rangle = Q \delta_{jl} \delta(t - t'), \quad \langle \Gamma_\alpha(t) \Gamma_j(t') \rangle = \langle \Gamma_\beta(t) \Gamma_j(t') \rangle = 0, \quad l, j = 1, 2.$$

Here we use the notation $\langle \rangle$ for mean value. Let us suppose that fluctuations are small. According to the notation of the paper the system (8.6) can be written as the following Stratonovich system:

$$\begin{aligned} b \equiv X &= X^1 + iX^2, \\ dX^1 &= (\alpha_o X^1 - \beta_o X^2 - (AX^1 - BX^2)XX^*)dt \\ &\quad + \varepsilon \left\{ \sum_{i=1}^2 (\alpha_i X^1 - \beta_i X^2) * dW_i + \sigma dW_3 \right\}, \\ dX^2 &= (\beta_o X^1 + \alpha_o X^2 - (BX^1 + AX^2)XX^*)dt \\ (8.7) \quad &\quad + \varepsilon \left\{ \sum_{i=1}^2 (\beta_i X^1 + \alpha_i X^2) * dW_i + \sigma dW_4 \right\}, \end{aligned}$$

where

$$\begin{aligned} \varepsilon^2[(\alpha_1)^2 + (\alpha_2)^2] &= Q_\alpha, \quad \varepsilon^2[(\beta_1)^2 + (\beta_2)^2] = Q_\beta, \\ \varepsilon^2(\alpha_1\beta_1 + \alpha_2\beta_2) &= Q_{\alpha\beta}, \quad \varepsilon\sigma = \sqrt{Q}. \end{aligned}$$

Under $\varepsilon = 0$ the system (8.7) becomes deterministic. In the case of $\alpha_o/A > 0$ it has an asymptotically stable limit cycle $(X^1)^2 + (X^2)^2 = \alpha_o/A$. The radius $\rho = |X|$ under $\varepsilon = 0$ satisfies the equation

$$d\rho/dt = \rho(\alpha_o - A\rho^2)$$

and does not depend on the detuning parameters β_o and B . The value ρ^2 under $\varepsilon \neq 0$ satisfies the Stratonovich equation

$$d\rho^2 = 2\rho^2(\alpha_o - A\rho^2)dt + 2\varepsilon\rho^2(\alpha_1 * dW_1 + \alpha_2 * dW_2) + 2\varepsilon\sigma(X_1 * dW_3 + X_2 * dW_4)$$

and also does not depend on β_o and B . But difference equations, which are the result of applying numerical methods to the system (8.7), essentially depend not only on the choice of a scheme and time step but also on the detuning parameters, and growing of $|\beta_o - B|$ leads to vanishing of stable cycle. Therefore, to solve the system (8.7) one must use high-order schemes or choose a sufficiently small time step. Since the system (8.7) contains multiplicative noises and does not belong to the class of systems with commutative noises, the Euler method is the highest-order scheme among known mean-square methods with easily simulated random variables [6], [7]. The Euler method has the mean-square error $O(h + \varepsilon^2 h^{1/2})$ and in the case of large

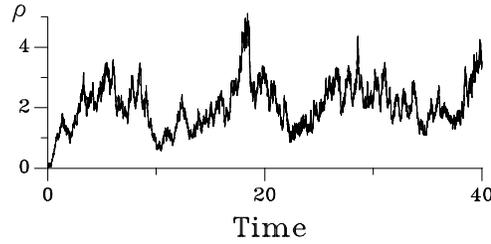


FIG. 3. *Laser Langevin equation. Time dependence of the radius $\rho = |X_k|$ for $\alpha_o = 0.5, \beta_o = 1, A = 0.1, B = 0.4, \varepsilon = 0.3, \alpha_i = \beta_i = \sigma = 1, i = 1, 2, X^1(0) = X^2(0) = 0$, and time step $h = 0.005$. The solution X_k of the system (8.7) is approximated by the Euler method and by the Runge–Kutta method with the error $O(h^4 + \varepsilon h + \varepsilon^2 h^{1/2})$.*

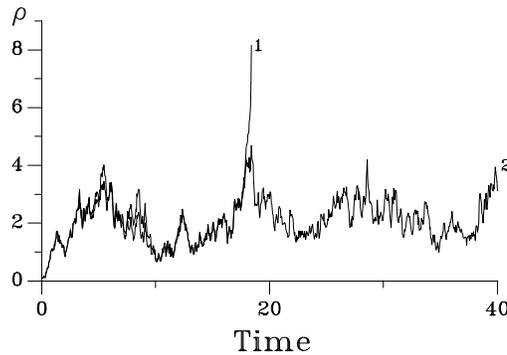


FIG. 4. *Laser Langevin equation. Time dependence of the radius $\rho = |X_k|$ for time step $h = 0.05$, other parameters are the same as in Fig. 3. The solution X_k of the system (8.7) is approximated by (1) the Euler method and (2) the Runge–Kutta method with the error $O(h^4 + \varepsilon h + \varepsilon^2 h^{1/2})$.*

$|\beta_o - B|$ too small step h is required. On the other hand, for instance, the method with the mean-square error $O(h^4 + \varepsilon h + \varepsilon^2 h^{1/2})$ allows us to obtain sufficiently accurate approximations for solutions of the system (8.7) and, particularly, to simulate phase trajectories.

The radius $\rho = |X_k|$ of a typical trajectory is plotted in Figures 3 and 4. Figure 3 demonstrates the radius ρ calculated under the time step $h = 0.005$ by the Euler scheme and by the Runge–Kutta scheme with the error $O(h^4 + \varepsilon h + \varepsilon^2 h^{1/2})$ which corresponds to the method (4.8) for the Ito system. In this case both methods give the same results. As seen in Figure 4, if one chooses greater time step ($h = 0.05$), the Runge–Kutta scheme gives quite good results (compare with Fig. 3), but the Euler method becomes unstable. In all cases we use the same sample paths for the Wiener processes.

Note that the Runge–Kutta method (4.8) and the corresponding method for the Stratonovich system can be improved up to $(ER^2)^{1/2} = O(h^4 + \varepsilon^2 h^{1/2})$ (see the method (4.7)).

The laser Langevin equation is often considered with colored multiplicative pump noises $\Gamma_\alpha(t)$ and $\Gamma_\beta(t)$ [18]. In that case one can simulate this equation by the methods of section 7.

9. Conclusions. Differential equations with small noises are the important cases of a stochastic system because, as mentioned in the Introduction, a dynamical system

is often affected by small fluctuations. For instance, stochastic resonance (see [19] and the references therein), which has the high degree of current interest, arises just in such systems. (Stochastic resonance is a phenomenon which is the result of a cooperative effect of noise and periodic forcing in a stochastic system.)

In this paper the approach to construction of efficient high-exactness mean-square methods for a system with small noises is developed. Thanks to a small parameter ε , new methods may be easier, require less computer time, and have smaller errors than ordinary schemes. Special attention has been paid to constructing methods with efficiently simulated random variables. Accuracy and convergence of a method are analyzed by the theorem on the estimate of mean-square errors. Herein the explicit, implicit, and Runge–Kutta methods with the mean-square errors from $O(h + \varepsilon^2 h^{1/2})$ up to $O(h^4 + \varepsilon^2 h^{1/2})$ are proposed for general Ito and Stratonovich systems with small noises. Moreover, systems with small additive noises and systems with small colored noises are considered. The appropriate methods for these systems have been also derived: for systems with small additive noises—schemes with mean-square errors from $O(h^2 + \varepsilon h)$ up to $O(h^4 + \varepsilon^2 h^{3/2})$ and for systems with small colored noises—schemes with mean-square errors from $O(h^2 + \varepsilon h)$ up to $O(h^4 + \varepsilon h^3 + \varepsilon^2 h^{5/2})$. By the proposed approach it is possible to derive many other numerical schemes for systems with small noises.

Mean-square methods are useful for direct simulation of stochastic trajectories which, for instance, can give an information on qualitative behavior of a stochastic model. They are the basis for construction of weak methods which are important for many practical applications. Weak methods are sufficient for calculation of mean values and solving problems of mathematical physics by Monte Carlo technique, and they are simpler than mean-square ones. Weak methods, constructed for systems with small noises, can be a useful tool for numerical solution of partial differential equations with small parameter at high derivative. Weak methods for systems with small noises and their applications to problems of mathematical physics will be the subject of a separate paper.

REFERENCES

- [1] C. W. GARDINER, *Handbook of Stochastic Methods for Physics, Chemistry and Natural Sciences*, Springer-Verlag, Berlin, 1983.
- [2] F. MOSS AND P. V. E. MCCLINTOCK, EDs., *Noise in Nonlinear Dynamical Systems*, Vol. I, II, III, Cambridge University Press, Cambridge, 1989.
- [3] G. N. MILSTEIN, *Approximate integration of stochastic differential equations*, Theory Probab. Appl., 19 (1974), pp. 557–562; *A method of second-order accuracy integration of stochastic differential equations*, Theory Probab. Appl., 23 (1978), pp. 396–401.
- [4] W. WAGNER AND E. PLATEN, *Approximation of Ito Integral Equations*, Zentralinstitut für Mathematik und Mechanik, Akad. der Wiss. der DDR, Berlin, 1978, preprint.
- [5] D. TALAY, *Efficient numerical schemes for the approximation of expectations of functionals of the solution of an SDE and applications*, Lecture Notes in Control and Inform. Sci., 61 (1984), pp. 294–313; E. PARDOUX AND D. TALAY, *Discretization and simulation of stochastic differential equations*, Acta Appl. Math., 3 (1985), pp. 23–47.
- [6] G. N. MILSTEIN, *Numerical Integration of Stochastic Differential Equations*, Ural University Press, Sverdlovsk, 1988; Kluwer Academic Publishers, Norwell, MA, 1995.
- [7] P. E. KLOEDEN AND E. PLATEN, *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag, Berlin, 1992.
- [8] G. N. MILSTEIN, *A theorem on the order of convergence of mean-square approximations of solutions of systems of stochastic differential equations*, Theory Probab. Appl., 32 (1987), pp. 738–741.
- [9] G. N. MILSTEIN AND M. V. TRET'YAKOV, *Numerical solution of differential equations with colored noise*, J. Statist. Phys., 77 (1994), pp. 691–715.

- [10] R. MANNELA AND V. PALLESCHI, *Fast and precise algorithm for computer simulation of stochastic differential equations*, Phys. Rev. A, 40 (1989), pp. 3381–3386; R. F. FOX, *Second-order algorithm for the numerical integration of colored-noise problems*, Phys. Rev. A, 43 (1991), pp. 2649–2654.
- [11] R. Z. HASMINSKI, *Stochastic Stability of Differential Equations*, Sijthoff & Noordhoff, Groningen, the Netherlands, 1980.
- [12] L. ARNOLD AND V. WIHSTUTZ, EDS., *Lyapunov Exponents*, Lecture Notes in Math., V. 1186, Springer-Verlag, Berlin, 1986; E. PARDOUX AND V. WIHSTUTZ, *Lyapunov exponent and rotation number of two-dimensional linear stochastic systems with small diffusion*, SIAM J. Appl. Math., 48 (1988), pp. 442–457.
- [13] D. TALAY, *Approximation of upper Lyapunov exponents of bilinear stochastic differential equations*, SIAM J. Numer. Anal., 28 (1991), pp. 1141–1164.
- [14] E. I. AUSLENDER AND G. N. MILSTEIN, *Asymptotic expansions of the Lyapunov index for linear stochastic systems with small noises*, Prikl. Mat. Mekh., 46 (1982), pp. 358–365.
- [15] W. H. PRESS, W. T. VETTERLING, S. A. TEUKOLSKY, AND B. P. FLANNERY, *Numerical Recipes*, Cambridge University Press, Cambridge, 1986.
- [16] R. GRAHAM, *Hopf bifurcation with fluctuating control parameter*, Phys. Rev. A, 25 (1982), pp. 3234–3258; A. SCHENZLE AND H. BRAND, *Multiplicative stochastic processes in statistical physics*, Phys. Rev. A, 20 (1979), pp. 1628–1647.
- [17] H. RISKEN, *The Fokker-Planck Equation*, Springer-Verlag, Berlin, 1984.
- [18] S. N. DIXIT AND P. S. SAHNI, *Nonlinear stochastic processes driven by colored noise: Application to dye-laser statistics*, Phys. Rev. Lett., 50 (1983), pp. 1273–1276; R. F. FOX AND R. ROY, *Steady-state analysis of strongly colored multiplicative noise in a dye laser*, Phys. Rev. A, 35 (1987), pp. 1838–1842.
- [19] F. MOSS, A. BULSARA, AND M. SHLESINGER, EDS., *Proceedings of the NATO advanced research workshop: Stochastic resonance in physics and biology*, J. Statist. Phys., 70 (1993), Nos. 1/2.