REFERENCES

[1] D. W. STROOK AND S. R. S. VARADHAN, *Diffusion processes with continuous coefficients*, Comm. Pure Appl. Math., 22 (1969), pp. 345–400.
[2] A. M. IL'IN, A. S. KALASHNIKOV AND O. A. OLEINIK, *Linear equations of the second order of parabolic type*, Russian Mathematical Surveys, 17, 3 (1962), pp. 1–143.
[3] E. B. DYNKIN, *Functionals of trajectories of Markov stochastic processes*, Dokl. Akad. Nauk SSSR, 104, 5 (1955), pp. 691–694. (In Russian.)
[4] I. I. GIKHMAN AND A. V. SKOROKHOD, *Stochastic Differential Equations*, Springer, New York, 1972.

APPROXIMATE INTEGRATION OF STOCHASTIC DIFFERENTIAL EQUATIONS

G. N. MIL'SHTEIN

(Translated by K. Durr)

1. Introduction

We shall consider a stochastic differential equation in the sense of Itô (see [1]–[3])

$$dX = a(t, X)\,dt + \sigma(t, X)\,dw(t), \qquad X(t_0) = x_0, \tag{1.1}$$

where $w(t)$ is a standard Wiener process.

The goal of this work is a numerically realizable mean square approximation of the random variable $X(t_0+T)$.

If the segment $[t_0, t_0+T]$ is divided into m equal parts and one sets $h = T/m$, $t_{k+1} = t_k + h$, $k = 0, 1, \cdots, m-1$, then the known recurrent approximation

$$\bar{X}(t_0) = X(t_0),\ \bar{X}(t_{k+1}) = a(t_k, \bar{X}(t_k))h + \sigma(t_k, \bar{X}(t_k))w_{k+1}, \tag{1.2}$$

where the w_{k+1} are independent random variables which are normally distributed with zero expectation and variance h, yields for the mean square deviation $\mathbf{E}_{t_0,x_0}(X(t_0+T) - \bar{X}(t_0+T))^2$ a value equal to $O(h)$.

In this note we propose a method which uses, as does (1.2), at each stage only w_{k+1}, but whose accuracy is equal to $O(h^2)$. We indicate how a greater accuracy can be achieved when using at each stage other random variables.

2. One-Stage Approximation

Let us form the random variable

$$\bar{X}(t_0+h) = x_0 + c_1 h + c_2 w(h) + c_3 w^2(h), \tag{2.1}$$

where $w(h) = w(t_0+h) - w(t_0)$ from (1.1) and take c_1, c_2 and c_3 so that the square of the mean square deviation $\mathbf{E}(X(t_0+h) - \bar{X}(t_0+h))^2$ has the highest possible order of smallness in h. For this purpose introduce the random process $\{X, w\}$ and its infinitesimal operator L:

$$Lf(h, x, w) = \frac{\partial f}{\partial h} + a(h, x)\frac{\partial f}{\partial x} + \frac{1}{2}\sigma^2(h, x)\frac{\partial^2 f}{\partial x^2} + \sigma(h, x)\frac{\partial^2 f}{\partial x \partial w} + \frac{1}{2}\frac{\partial^2 f}{\partial w^2}. \tag{2.2}$$

Let us adduce an important formula connected with Taylor's expansion of semigroups:

$$\mathbf{E}_{t_0,x_0,w_0}f(t_0+h,\ X(t_0+h),\ w(t_0+h)) \tag{2.3}$$
$$= f(t_0, x_0, w_0) + Lf(t_0, x_0, w_0)h + \cdots + \frac{1}{k!}L^k f(t_0, x_0, w_0)h^k + O(h^{k+1}).$$

Formula (2.3) may be obtained under suitable restrictions on the functions a, σ and f from well-known theorems from the theory of semigroups (see [4], Chapter 11, § 2). Set $\varphi(x, h, w) = x - \bar{X}(t_0+h) = x - x_0 - c_1h - c_2w - c_3w^2$. From (2.3) it is clear that if $L^s\varphi^2(0, x_0, 0) = 0$, $s = 1, \cdots, k$, then $\mathbf{E}_{t_0,x_0,0}(X(t_0+h) - \bar{X}(t_0+h))^2 = O(h^{k+1})$.

By simple calculations we obtain the formula

$$L(f\cdot g) = Lf\cdot g + f\cdot Lg + Sf\cdot Sg, \tag{2.4}$$

where the operator S is given by

$$Sf(h, x, w) = \sigma(h, x)\frac{\partial f}{\partial x} + \frac{\partial f}{\partial w}. \tag{2.5}$$

Using (2.4) we find that

$$L\varphi^2 = 2\varphi\cdot L\varphi + (S\varphi)^2, \tag{2.6}$$

$$L^2\varphi^2 = 2(L\varphi)^2 + 2\varphi L^2\varphi + 2S\varphi(SL\varphi + LS\varphi) + (S^2\varphi)^2. \tag{2.7}$$

Denoting by $(f)_0$ the value of f at $h = 0$, $x = x_0$, $w = 0$, from the relations $(L\varphi^2)_0 = 0$ and $(L^2\varphi^2)_0 = 0$, we obtain

$$(S\varphi)_0 = 0, \qquad (L\varphi)_0 = 0, \qquad (S^2\varphi)_0 = 0. \tag{2.8}$$

From (2.8) we get

$$\begin{aligned} c_1 &= a(t_0, x_0) - \frac{1}{2}\sigma(t_0, x_0)\frac{\partial\sigma}{\partial x}(t_0, x_0),\\ c_2 &= \sigma(t_0, x_0),\\ c_3 &= \frac{1}{2}\sigma(t_0, x_0)\frac{\partial\sigma}{\partial x}(t_0, x_0). \end{aligned} \tag{2.9}$$

Thus the one-stage approximation (2.1) with coefficients (2.9) has order h^3 for the square of the mean square error.

Observe that the usual approximation (1.2) has order h^2 for one stage.

It is natural to attempt in (2.1) to add the terms w^3 and hw and attain greater precision. However it is not hard to see that in the general case it is impossible to obtain a more precise result in this way. But if we introduce the random variable $y = \int_0^h w\,dt$ and construct $\bar{X}(t_0+h)$ according to the formula

$$\bar{X}(t_0+h) = x_0 + c_1h + c_2w + c_3w^2 + c_4hw + c_5w^3 + c_6y, \tag{2.10}$$

then one can choose the coefficients $c_1, \cdots, c_6$ so that the square of the mean square error in one stage will be equal to $O(h^4)$.

To this end, consider the random process $\{X, w, y\}$ generated by the system

$$\begin{aligned} dX &= a\,dt + \sigma\,dw,\\ dw &= 1\cdot dw,\\ dy &= w\,dt. \end{aligned}$$

The infinitesimal operator L of this process is given by

$$Lf(h, x, w, y) = \frac{\partial f}{\partial h} + a\frac{\partial f}{\partial x} + w\frac{\partial f}{\partial y} + \frac{1}{2}\sigma^2\frac{\partial^2 f}{\partial x^2} + \frac{1}{2}\frac{\partial^2 f}{\partial w^2} + \sigma\frac{\partial^2 f}{\partial x\partial w}. \tag{2.11}$$

Formulas (2.4)–(2.7) with the new variable y remain valid for $\{X, w, y\}$. Here the function $\varphi(h, x, w, y)$ is equal to

$$\varphi = x - x_0 - c_1 h - c_2 w - c_3 w^2 - c_4 hw - c_5 w^3 - c_6 y.$$

Finding $L^3\varphi^2$ from (2.4) and (2.7) and setting $(L^3\varphi^2)_0$ equal to zero, we obtain along with (2.8)

$$(SL\varphi)_0 = 0, \qquad (LS\varphi)_0 = 0, \qquad (S^3\varphi)_0 = 0. \tag{2.12}$$

Formulas (2.8) and (2.12) yield (2.9) and the following expressions for the coefficients c_4, c_5 and c_6:

$$\begin{aligned} c_4 &= \left(\frac{\partial\sigma}{\partial t} + a\frac{\partial\sigma}{\partial x} - \frac{1}{2}\sigma\left(\frac{\partial\sigma}{\partial x}\right)^2\right)_0, \\ c_5 &= \left(\frac{1}{6}\frac{\partial}{\partial x}\left(\sigma\frac{\partial\sigma}{\partial x}\right)\right)_0, \\ c_6 &= \left(\sigma\frac{\partial a}{\partial x} - \frac{\partial\sigma}{\partial t} - a\frac{\partial\sigma}{\partial x} - \frac{1}{2}\sigma^2\frac{\partial^2\sigma}{\partial x^2}\right)_0. \end{aligned} \tag{2.13}$$

The approximation (2.10) with the coefficients (2.9) and (2.13) has order h^4 for the square of the mean square error in one stage.

Introducing the random variable $z = \int_0^h w^2\, dt$ and defining $\bar{X}(t_0+h)$ by the formula

$$\begin{aligned} \bar{X}(t_0+h) &= x_0 + c_1 h + c_2 w + c_3 w^2 + c_4 hw + c_5 w^3 \\ &\quad + c_6 y + c_7 h^2 + c_8 hw^2 + c_9 w^4 + c_{10} wy + c_{11} z \end{aligned} \tag{2.14}$$

one can choose coefficients $c_1, \cdots, c_{11}$, which are not hard to compute following the above arguments, such that the order of the square of the mean square error in one stage will be equal to h^5.

Thus introduction of the new random variables increases the accuracy. However numerical realization of the random variables (2.10) and (2.14) is significantly more complicated than realization of (2.1).

REMARK. For the sake of brevity we do not mention explicitly in this paper all the conditions under which the calculations are valid.

Proper justification can be attained without especial difficulty under appropriate conditions having to do with the smoothness and growth of the coefficients a and σ and their derivatives.

For example, conditions of boundedness and continuity of the coefficients a and σ and their partial derivatives in t and x up to fourth order inclusive are sufficient for the validity of the formula

$$\mathbf{E}(X(t_0+h) - \bar{X}(t_0+h))^2 = O(h^3). \tag{2.15}$$

To prove this we adduce the formula

$$\begin{aligned} &\mathbf{E}_{t_0,x_0,w_0} f(t_0+h, X(t_0+h), w(t_0+h)) = f(t_0, x_0, w_0) \\ &\quad + Lf(t_0, x_0, w_0)h + \frac{1}{2}L^2 f(t_0, x_0, w_0)h^2 \\ &\quad + \frac{1}{2}\int_{t_0}^{t_0+h} (t_0+h-\xi)^2 \mathbf{E}_{t_0,x_0,w_0} L^3 f(\xi, X(\xi), w(\xi))\, d\xi. \end{aligned} \tag{2.16}$$

Formula (2.16) is valid for functions f which are bounded together with the functions Lf, $L^2 f$ and $L^3 f$. It is obtained from the corresponding formula from the theory of semigroups (see [4], p. 311) if one passes, as usual, from the non-homogeneous Markov process $\{X, w\}$ to the homogeneous one $\{X, w, t\}$ by introducing the new variable τ and the equation $dt/d\tau = 1$.

Under the above conditions on the coefficients a and σ, formula (2.16) is valid for $f=\varphi^2$.

To prove this it is necessary to consider the sequence of functions φ_N^2:

$$\varphi_N^2=\begin{cases}\varphi^2, & \varphi^2\leqq N,\\ N+\frac{1}{2}, & \varphi^2\geqq N+1,\\ L_N(\varphi^2), & N<\varphi^2<N+1,\end{cases}$$

where $L_N(z)$ is the Lagrange interpolation polynomial such that $L_N(N)=N$, $L'_N(N)=N$, $L'_N(N)=1$, $L''_N(N)=\cdots=L_N^{(6)}(N)=0$, $L_N(N+1)=N+\frac{1}{2}$, $L'_N(N+1)=\cdots=L_N^{(6)}(N+1)=0$. Formula (2.16) is valid for each function φ_N^2. Passing to the limit as $N\to\infty$, we obtain formula (2.16) for the function φ^2; whence (2.15) follows immediately.

3. Approximation on a Finite Interval

Consider the sequence

$$\bar{X}(t_0)=x_0,$$
$$\bar{X}(t_0+(k+1)h)=\bar{X}(t_0+kh)+\bar{\sigma}w_{k+1}+\left(\bar{a}-\frac{1}{2}\bar{\sigma}\overline{\frac{\partial\sigma}{\partial x}}\right)h+\frac{1}{2}\bar{\sigma}\overline{\frac{\partial\sigma}{\partial x}}w_{k+1}^2, \tag{3.1}$$

where $h=T/m$, $k=0,1,\cdots,m-1$, $w_1,\cdots,w_m$ are independent random variables with normal distribution, zero expectation and variance h; the bar signifies that the function is evaluated at the point $(t_0+kh,\bar{X}(t_0+kh))$. Set

$$\varepsilon_k^2=\mathbf{E}_{t_0,x_0}(X(t_0+kh)-\bar{X}(t_0+kh))^2. \tag{3.2}$$

Let us estimate ε_{k+1}^2 in terms of its dependence on ε_k^2. We have

$$\begin{aligned}\varepsilon_{k+1}^2&=\mathbf{E}_{t_0,x_0}(X(t_0+(k+1)h)-\bar{X}(t_0+(k+1)h))^2\\&=\mathbf{E}_{t_0,x_0}(\mathbf{E}_{t_0+kh,X(t_0+kh),\bar{X}(t_0+kh)}\varphi_{k+1}^2(h,X(t_0+(k+1)h),w_{k+1}(h);\bar{X}(t_0+kh))),\end{aligned} \tag{3.3}$$

where

$$\begin{aligned}\varphi_{k+1}(h,x,w_{k+1};\bar{X}(t_0+kh))&=x-\bar{X}(t_0+kh)\\&\quad-\left(\bar{a}-\frac{1}{2}\bar{\sigma}\overline{\frac{\partial\sigma}{\partial x}}\right)h-\sigma w_{k+1}-\frac{1}{2}\sigma\overline{\frac{\partial\sigma}{\partial x}}w_{k+1}^2.\end{aligned}$$

Using (2.3) we find that

$$\begin{aligned}\mathbf{E}_{t_0+kh,X(t_0+kh),\bar{X}(t_0+kh)}\varphi_{k+1}^2&=(X(t_0+kh)-\bar{X}(t_0+kh))^2\\&\quad+(L\varphi_{k+1}^2)_0h+\tfrac{1}{2}(L^2\varphi_{k+1}^2)_0h^2+O(h^3),\end{aligned} \tag{3.4}$$

where

$$(L\varphi_{k+1}^2)_0=2(x_k-\bar{x}_k)(a-\bar{a})+(\sigma-\bar{\sigma})^2, \tag{3.5}$$

$$\begin{aligned}(L^2\varphi_{k+1}^2)_0&=2(a-\bar{a})^2+2(x_k-\bar{x}_k)\left(\frac{\partial a}{\partial t}+a\frac{\partial a}{\partial x}+\frac{1}{2}\sigma^2\frac{\partial^2a}{\partial x^2}\right)\\&\quad+2(\sigma-\bar{\sigma})\left(\sigma\frac{\partial a}{\partial x}+\frac{\partial\sigma}{\partial t}+a\frac{\partial\sigma}{\partial x}+\frac{1}{2}\sigma^2\frac{\partial^2\sigma}{\partial x^2}\right)+\left(\sigma\frac{\partial\sigma}{\partial x}-\bar{\sigma}\overline{\frac{\partial\sigma}{\partial x}}\right).\end{aligned} \tag{3.6}$$

In (3.5) and (3.6), $x_k = X(t_0+kh)$, $\bar{x}_k = \bar{X}(t_0+kh)$; the unbarred functions are evaluated at (t_0+kh, x_k), the barred functions at $(t_0+kh, \bar{x}_k)$.

Assuming the boundedness of the corresponding derivatives of the coefficients a and σ, we obtain from (3.2)–(3.6) the following inequality:

$$\varepsilon_{k+1}^2 \leqq \varepsilon_k^2 + A\varepsilon_k^2 h + B\varepsilon_k h^2 + Ch^3, \tag{3.7}$$

where A, B and C are positive constants.

Consider the equality

$$\eta_{k+1}^2 = \eta_k^2 + A\eta_k^2 h + B\eta_k h^2 + Ch^3, \quad \eta_0 = 0. \tag{3.8}$$

It is clear that $\varepsilon_k^2 \leqq \eta_k^2$ and therefore if $\eta_m^2 = O(h^2)$, then $\varepsilon_m^2 = O(h^2)$.

Let us prove that $\eta_m^2 = O(h^2)$. Note that η_k^2 is monotonically increasing and that for small k the values of η_k^2 are smaller than h^2. If this holds for all $k = 0, 1, \cdots, m-1$, then $\eta_m^2 = O(h^2)$. Let $\eta_{k_0}^2 \leqq h^2$ and, for $k > k_0$, $\eta_k^2 > h^2$. For $k > k_0$ we obtain from (3.8) that

$$\eta_{k+1}^2 \leqq \eta_k^2(1+(A+B+C)h).$$

From this we find that

$$\eta_m^2 \leqq \eta_{k_0}^2(1+(A+B+C)h)^m \leqq \text{const.} \cdot \eta_{k_0}^2 \leqq \text{const.} \cdot h^2.$$

Thus it is proved that the order of the square of the mean square error on $[t_0, t_0+T]$ in using the one-stage approximation (2.1), (2.9) is equal to h^2.

Similar computations show that using the usual approximation (1.2) this order is equal to h.

4. Formulas (2.1) and (2.9) may turn out to be inconvenient due to the necessity of computing $\partial\sigma/\partial x$ at each stage.

One can use the idea of the method of Runge–Kutte and seek a one-stage approximation in the form

$$\bar{X}(t_0+h) = x_0 + a_0 h + p_1\sigma_0 w + p_2\sigma\left(t_0, x_0 + \alpha_1\sigma_0 w + \alpha_2\sigma_0\frac{h}{w}\right)w, \tag{4.1}$$

where $a_0 = a(t_0, x_0)$, $\sigma_0 = \sigma(t_0, x_0)$.

If $\sigma(t_0, x_0 + \alpha_1\sigma_0 w + \alpha_2\sigma_0(h/w))w$ is replaced by the expression $\sigma_0 w + (\partial\sigma/\partial x)(t_0, x_0) \times(\alpha_1\sigma_0 w^2 + \alpha_2\sigma_0 h)$ and (4.1) is compared with (2.1) and (2.9), then we find the following relations for the coefficients p_1, p_2, α_1 and α_2:

$$p_1 + p_2 = 1, \qquad p_2\alpha_1 = -p_2\alpha_2 = \tfrac{1}{2}. \tag{4.2}$$

The order of the square of the mean square error in one stage of the approximation (4.1) and (4.2) is h^3.

Let us adduce analogues of formulas (2.1), (2.9), and (4.1), (4.2) for equation (1.1) interpreted in the sense of Stratonovich [3], [5]:

$$dX = a(t, X)\,dt + \sigma(t, X)\,d^*w. \tag{4.3}$$

In this case,

$$\bar{X}(t_0+h) = x_0 + a_0 h + \sigma_0 w + \frac{1}{2}\sigma_0\left(\frac{\partial\sigma}{\partial x}\right)_0 w^2. \tag{4.4}$$

The formulas which are analogous to (4.1) and (4.2) are more simple here:

$$\bar{X}(t_0+h) = x_0 + a_0 h + p_1\sigma_0 w + p_2\sigma(t_0, x_0 + \alpha\sigma_0 w)w, \tag{4.5}$$

$$p_1 + p_2 = 1, \qquad p_2\alpha = \tfrac{1}{2}. \tag{4.6}$$

5. Generalization to Systems of Equations

Consider the system of stochastic equations

$$dX = a(t, X)\, dt + \sigma(t, X)\, dw. \tag{5.1}$$

In (5.1), $X(t)$, $a(t, X)$ and $\sigma(t, X)$ are n-dimensional vectors and $w(t)$ is a standard Wiener process.

All the above results are easily carried over to systems of the form (5.1). The recurrent approximation formulas (the analogue of (3.1)) have the form

$$\begin{aligned} &\bar{X}(t_0) = X_0, \\ &\bar{X}(t_0+(k+1)h) = \bar{X}(t_0+kh) + \bar{\sigma} w_{k+1} + \left(\bar{a} - \frac{1}{2}\overline{\left(\sigma, \frac{\partial}{\partial x}\right)\sigma}\right) h + \frac{1}{2}\overline{\left(\sigma, \frac{\partial}{\partial x}\right)\sigma} w_{k+1}^2. \end{aligned} \tag{5.2}$$

In (5.2), the bar signifies that the function is evaluated at the point $(t_0+kh, \bar{X}(t_0+kh))$; the i-th coordinate of the vector $(\sigma, \partial/\partial x)\sigma$ is equal to $\sum_{j=1}^n \sigma_j \partial\sigma_i/\partial x_j$.

Consider the more general system of stochastic equations

$$dX(t) = a(t, X)\, dt + \sum_{r=1}^{q} \sigma_r(t, X)\, dw_r(t). \tag{5.3}$$

Here X, a and σ_n are n-dimensional vectors, and $w_n(t)$, $r = 1, \cdots, q$, are independent Wiener processes. For the system (5.3) it is natural to use an approximation of the form

$$\bar{X}(t_0+h) = X_0 + c_0 h + \sum_{r=1}^{q} c_r w_r + \sum_{i,j=1}^{q} c_{ij} w_i w_j. \tag{5.4}$$

It is of interest to note that in this case one is not successful in choosing vectors c_0, c_n and c_{ij} so that the order of the square of the mean square error in one stage is h^3. This is connected with the fact that the operators S_r, which are the analogues of the operator S, are not commutative.

However if we introduce the random variables $z_{ij} = \int_0^h w_i\, dw_j$, then by constructing a random vector of the form

$$\bar{X}(t_0+h) = X_0 + c_0 h + \sum_{r=1}^{q} c_r w_r + \sum_{i,j=1}^{q} c_{ij} z_{ij}, \tag{5.5}$$

one can now determine such vectors c_0, c_n and c_{ij}:

$$c_0 = a, \qquad c_r = \sigma_r, \qquad c_{ij} = \left(\sigma_i, \frac{\partial}{\partial x}\right)\sigma_j. \tag{5.6}$$

All the functions in (5.6) are evaluated at $(t_0, X(t_0))$.

Received by the editors September 23, 1973

REFERENCES

[1] I. I. GIKHMAN and A. V. SKOROKHOD, *Introduction to the Theory of Random Processes*, Moscow, Izd-vo "Nauka", 1965. (In Russian.)

[2] I. I. GIKHMAN and A. V. SKOROKHOD, *Stochastic Differential Equations*, Springer, New York, 1972.

[3] R. Z. KHAS'MINSKII, *Stability of Systems of Differential Equations under Random Perturbations of their Parameters*, Moscow, Izd-vo "Nauka", 1969. (In Russian.)

[4] E. HILLE and R. PHILLIPS, *Functional Analysis and Semigroups*, American Mathematical Society, Providence, R.I., 1957.

[5] R. L. STRATONOVICH, *Conditional Markov Processes and their Application to the Theory of Optimal Control*, American Elsevier, New York, 1968.