Error and Convergence of Two Numerical Schemes for Stochastic Differential Equations

Brian D. Ewald
Gardner-Webb University, Boiling Springs, North Carolina 28017

Received 5 January 2005; accepted 6 March 2006
Published online 29 May 2006 in Wiley InterScience (www.interscience.wiley.com).
DOI 10.1002/num.20165

We examine the convergence and error rate of two stochastic numerical schemes using the method of proof used by G. N. Mil'shtein (Theory Prob Appl 19 (1974), 557–562). © 2006 Wiley Periodicals, Inc. Numer Methods Partial Differential Eq 22: 1247–1253, 2006

Keywords: stochastic numerical schemes

1. INTRODUCTION

In G. N. Mil'shtein's paper [1], first published in 1974, the author introduced the scheme now named for him:

$$\bar{x}_{n+1} = \bar{x}_n + \left(a(t_n, \bar{x}_n) - \frac{1}{2}\,\sigma(t_n, \bar{x}_n)\,\frac{\partial \sigma}{\partial x}(t_n, \bar{x}_n) \right)\Delta t + \sigma(t_n, \bar{x}_n)\Delta W_n \tag{1}$$

$$+ \frac{1}{2}\,\sigma(t_n, \bar{x}_n)\,\frac{\partial \sigma}{\partial x}(t_n, \bar{x}_n)(\Delta W_n)^2,$$

which can be used to approximate the solution to the one-dimensional Itô stochastic differential equation

$$dx_t = a(t, x_t)dt + \sigma(t, x_t)dW_t. \tag{2}$$

He then proved that his scheme converges in the root-mean-square sense with order $O(\Delta t)$.

Much work has since been done in this area, and many different schemes have been proposed (see [2] and [3] for a detailed analysis of many such schemes). However, these schemes use

Correspondence to: Brian D. Ewald, P.O. Box 7324, Gardner-Webb University, Boiling Springs, NC 28017 (e-mail: bewald@gardner-webb.edu)

methods largely different from that originally used by Mil'shtein. In this article, we see that Mil'shtein's original method can also be used to prove the convergence of other schemes.

In this article, we study two schemes. The first is a backward Mil'shtein scheme. That is, we alter Mil'shtein's scheme by making it implicit in the deterministic term. The second scheme is a finite difference scheme. In it, we alter Mil'shtein's scheme by replacing the derivative of σ by a finite difference. This is sometimes useful, for example, if the calculation of the derivative of σ is numerically intensive, or if σ itself is only known empirically (e.g., from tables).

2. A BACKWARD MIL'SHTEIN SCHEME

We consider the stochastic differential equation

$$dx_t = [a(x_t) + \lambda\sigma(x_t)\sigma'(x_t)]dt + \sigma(x_t)dw_t, \qquad x(t_0) = x_0, \tag{3}$$

where the stochastic integral is taken in the sense of Itô. We note that, for instance, if $\lambda = \frac{1}{2}$, this is equivalent to the stochastic differential equation

$$dx_t = a(x_t)dt + \sigma(x_t) \circ dw_t, \qquad x(t_0) = x_0, \tag{4}$$

where now the stochastic integral is taken in the sense of Stratanovich.

In this section, we would like to consider an implicit scheme where $\bar{x}(t_0) = x_0$ and the $\bar{x}(t_{k+1})$ are recursively given as the solutions to

$$\begin{aligned} \bar{x}(t_{k+1}) = \bar{x}(t_k) + a(\bar{x}(t_{k+1}))\Delta t + \sigma(\bar{x}(t_k))\Delta w_k \\ + \sigma(\bar{x}(t_k))\sigma'(\bar{x}(t_k))[\mu_1\Delta w_k^2 - \mu_2\Delta t]; \end{aligned} \tag{5}$$

here μ_1 and μ_2 are constants to be determined later. For simplicity in notation, in what follows, we will write $\bar{x}_k$ for $\bar{x}(t_k)$ and $\bar{a}_k$ for $a(\bar{x}_k)$, etc.

2.1. Numerical Realization of the Backward Scheme

We see that we will need a method to solve the equation

$$\hat{x}_k = \bar{x}_k + a(\hat{x}_k)t + \bar{\sigma}_k w + \bar{\sigma}_k\bar{\sigma}'_k[\mu_1 w^2 - \mu_2 t], \tag{6}$$

for $\hat{x}_k = \hat{x}_k(t, w)$. For this purpose, we see that

$$\begin{aligned} a(\hat{x}_k) &= a(\bar{x}_k + a(\hat{x}_k)t + \bar{\sigma}_k w + \bar{\sigma}_k\bar{\sigma}'_k[\mu_1 w^2 - \mu_2 t]) \\ &= a(\bar{x}_k) + a'(\bar{x}_k)[a(\hat{x}_k)t + \bar{\sigma}_k w + \bar{\sigma}_k\bar{\sigma}'_k[\mu_1 w^2 - \mu_2 t]] \\ &\quad + \tfrac{1}{2}a''(\bar{x}_k)[a(\hat{x}_k)t + \bar{\sigma}_k w + \bar{\sigma}_k\bar{\sigma}'_k[\mu_1 w^2 - \mu_2 t]]^2 \\ &\quad + O(t^3, t^2 w, tw^2, w^3) \end{aligned} \tag{7}$$

$$= \bar{a}_k + \bar{a}'_k a(\hat{x}_k)t + \bar{a}'_k \bar{\sigma}_k w + \bar{a}'_k \bar{\sigma}_k \bar{\sigma}'_k (\mu_1 w^2 - \mu_2 t)$$

$$+ \tfrac{1}{2} \bar{a}''_k \bar{\sigma}_k^2 w^2 + O(t^2, tw, w^3)$$

$$= \bar{a}_k + \bar{a}'_k \bar{a}_k t + \bar{a}'_k \bar{\sigma}_k w + \bar{a}'_k \bar{\sigma}_k \bar{\sigma}'_k (\mu_1 w^2 - \mu_2 t)$$

$$+ \tfrac{1}{2} \bar{a}''_k \bar{\sigma}_k^2 w^2 + O(t^2, tw, w^3).$$

Therefore, instead of the scheme (5) above, we will consider the following scheme, which agrees with (5) at order $O(t^2, tw, w^3)$ (namely, its order of accuracy):

$$\bar{x}_{k+1} = \bar{x}_k + \bar{a}_k \Delta t + \bar{a}'_k \bar{a}_k \Delta t^2 + \bar{a}'_k \bar{\sigma}_k \Delta w \Delta t \tag{8}$$

$$+ \bar{a}_k \bar{\sigma}_k \bar{\sigma}'_k (\mu_1 \Delta w^2 - \mu_2 \Delta t)\Delta t + \tfrac{1}{2} \bar{a}''_k \bar{\sigma}_k^2 \Delta w^2 \Delta t$$

$$+ \bar{\sigma}_k \Delta w + \bar{\sigma}_k \bar{\sigma}'_k (\mu_1 \Delta w^2 - \mu_2 \Delta t).$$

2.2. One-step Approximation

We note, as in [1], that the random process $\{x, w\}$, where x is the solution to the Itô equation (3), has the infinitesimal generator L given by

$$Lf(t, x, w) = \frac{\partial f}{\partial t} + [a(t, x) + \lambda \sigma(t, x)\sigma'(t, x)] \frac{\partial f}{\partial x} \tag{9}$$

$$+ \frac{1}{2} \sigma^2(t, x) \frac{\partial^2 f}{\partial x^2} + \sigma(t, x) \frac{\partial^2 f}{\partial x \partial w} + \frac{1}{2} \frac{\partial^2 f}{\partial w^2}.$$

We then have the following Taylor expansion for the semigroup

$$\mathbf{E}_{t_0, x_0, w_0} f(t_0 + t, x(t_0 + t), w(t_0 + t)) \tag{10}$$

$$= f(t_0, x_0, w_0) + Lf(t_0, x_0, w_0)t + \tfrac{1}{2} L^2 f(t_0, x_0, w_0)t^2 + O(t^3),$$

when f, a, and σ satisfy suitable conditions.

We note also the formula

$$L(f \cdot g) = Lf \cdot g + f \cdot Lg + Sf \cdot Sg, \tag{11}$$

where

$$Sf(t, x, w) = \sigma(t, x) \frac{\partial f}{\partial x} + \frac{\partial f}{\partial w}. \tag{12}$$

Therefore we have that

$$L\phi^2 = 2\phi \cdot L\phi + (S\phi)^2, \tag{13}$$

$$L^2\phi^2 = 2(L\phi)^2 + 2\phi \cdot L^2\phi + 2S\phi \cdot (LS\phi + SL\phi) + (S^2\phi)^2. \tag{14}$$

We now consider the function

$$\phi_k(t, x, w) = x + x_k - \bar{x}_k - \bar{a}_k t - \bar{a}_k'\bar{a}_k t^2 - \bar{a}_k'\bar{\sigma}_k wt \tag{15}$$

$$- \bar{a}_k\bar{\sigma}_k\bar{\sigma}_k'(\mu_1 w^2 - \mu_2 t)t - \tfrac{1}{2}\bar{a}_k''\bar{\sigma}_k^2 w^2 t$$

$$- \bar{\sigma}_k w - \bar{\sigma}_k\bar{\sigma}_k'(\mu_1 w^2 - \mu_2 t).$$

Then, we see that

$$L\phi_k = -\bar{a}_k - 2\bar{a}_k\bar{a}_k' t - \bar{a}_k'\bar{\sigma}_k w - \bar{a}_k\bar{\sigma}_k\bar{\sigma}_k'(\mu_1 w^2 - \mu_2 t) + \mu_2\bar{a}_k\bar{\sigma}_k\bar{\sigma}_k' t - \tfrac{1}{2}\bar{a}_k''\bar{\sigma}_k^2 w^2 + \mu_2\bar{\sigma}_k\bar{\sigma}_k'$$
$$+ (a + \lambda\sigma\sigma') - \bar{a}_k\bar{\sigma}_k\bar{\sigma}_k'\mu_1 t - \tfrac{1}{2}\bar{a}_k''\bar{\sigma}_k^2 t - \mu_1\bar{\sigma}_k\bar{\sigma}_k',$$

$$S\phi_k = \sigma - \bar{a}_k'\bar{\sigma}_k t - 2\mu_1\bar{a}_k\bar{\sigma}_k\bar{\sigma}_k' wt - \bar{a}_k''\bar{\sigma}_k^2 wt - \bar{\sigma}_k - 2\mu_1\bar{\sigma}_k\bar{\sigma}_k' w,$$

$$L^2\phi_k = -2\bar{a}_k\bar{a}_k' + 2\mu_2\bar{a}_k\bar{\sigma}_k\bar{\sigma}_k' - \mu_1\bar{a}_k\bar{\sigma}_k\bar{\sigma}_k' - \frac{1}{2}\bar{a}_k''\bar{\sigma}_k^2 + (a + \lambda\sigma\sigma')\frac{\partial}{\partial x}(a + \lambda\sigma\sigma')$$
$$+ \frac{1}{2}\sigma^2\frac{\partial^2}{\partial x^2}(a + \lambda\sigma\sigma') - \mu_1\bar{a}_k\bar{\sigma}_k\bar{\sigma}_k' - \frac{1}{2}\bar{a}_k''\bar{\sigma}_k^2,$$

$$SL\phi_k = \sigma\frac{\partial}{\partial x}(a + \lambda\sigma\sigma') - \bar{a}_k'\sigma_k - 2\mu_1\bar{a}_k\bar{\sigma}_k\bar{\sigma}_k' w - \bar{a}_k''\bar{\sigma}_k^2 w,$$

$$LS\phi_k = -\bar{a}_k'\bar{\sigma}_k - 2\mu_1\bar{a}_k\bar{\sigma}_k\bar{\sigma}_k' w - \bar{a}_k''\bar{\sigma}_k^2 w + (a + \lambda\sigma\sigma')\sigma' + \tfrac{1}{2}\sigma^2\sigma'',$$

$$S^2\phi_k = \sigma\sigma' - 2\mu_1\bar{a}_k\bar{\sigma}_k\bar{\sigma}_k' t - \bar{a}_k''\bar{\sigma}_k^2 t - 2\mu_1\bar{\sigma}_k\bar{\sigma}_k'.$$

Therefore, we will have

$$\phi_k^2(t_{k+1} - t_k, x_{k+1} - x_k, w_{k+1} - w_k) = (x_{k+1} - \bar{x}_{k+1})^2,$$

$$\phi_k^2(0, 0, 0) = (x_k - \bar{x}_k)^2,$$

$$L\phi_k^2(0, 0, 0) = 2(x_k - \bar{x}_k)[(a_k - \bar{a}_k) + (\lambda\sigma_k\sigma_k' - (\mu_1 - \mu_2)\bar{\sigma}_k\bar{\sigma}_k')] + (\sigma_k - \bar{\sigma}_k)^2,$$

$$L^2\phi_k^2(0, 0, 0) = 2[(a_k - \bar{a}_k) + (\lambda\sigma_k\sigma_k' - (\mu_1 - \mu_2)\bar{\sigma}_k\bar{\sigma}_k')]^2 + 2(x_k - \bar{x}_k)\Bigg[-2\bar{a}_k\bar{a}_k' + 2(\mu_1 - \mu_2)\bar{a}_k\bar{\sigma}_k\bar{\sigma}_k' - \bar{a}_k''\bar{\sigma}_k^2 + (a_k + \lambda\sigma_k\sigma_k')\frac{\partial}{\partial x}(a + \lambda\sigma\sigma')_k + \frac{1}{2}\sigma_k^2\frac{\partial^2}{\partial x^2}(a + \lambda\sigma\sigma')_k\Bigg] + 2(\sigma_k - \bar{\sigma}_k) \times \left[\sigma_k\frac{\partial}{\partial x}(a + \lambda\sigma\sigma')_k - 2\bar{a}_k'\bar{\sigma}_k + (a_k + \lambda\sigma_k\sigma_k')\sigma_k' + \frac{1}{2}\sigma_k^2\sigma_k''\right] + (\sigma_k\sigma_k' - 2\mu_1\bar{\sigma}_k\bar{\sigma}_k')^2.$$

We now note that we will have equations such as

$$a_k - \bar{a}_k = a'(\xi)(x_k - \bar{x}_k),$$

for some ξ between x_k and $\bar{x}_k$. Therefore, if we set $\mu_1 = \lambda + \mu_2 = \frac{1}{2}$ and define the mean-square errors $\epsilon_k^2 = \mathbf{E}(x_k - \bar{x}_k)^2$, we can take the expectation of (10) for $t = t_{k+1} - t_k$, $x = x_{k+1} - x_k$, and $w = w_{k+1} - w_k$ to arrive at

$$\epsilon_{k+1}^2 \leq \epsilon_k^2 + C_1\epsilon_k^2\Delta t + C_2\epsilon_k\Delta t^2 + C_3\Delta t^3. \tag{16}$$

Remark 1. *The constants C_k above depend on uniform bounds on a', σ', $(\sigma\sigma')'$, etc., so we must assume that a and σ have the necessary bounds on these derivatives. More specifically, in this case it suffices to have $C_k = C_k(|a|_{C^4}, |\sigma|_{C^4})$. It is also necessary that ϕ_k^2 satisfy* (10). *For this, we note that the formula*

$$\begin{aligned} &\mathbf{E}_{t_0, x_0, w_0} f(t_0 + t, x(t_0 + t), w(t_0 + t)) \\ &= f(t_0, x_0, w_0) + Lf(t_0, x_0, w_0)t + \tfrac{1}{2}L^2 f(t_0, x_0, w_0)t^2 \\ &+ \frac{1}{2}\int_{t_0}^{t_0+t}(t_0 + t - h)^2\mathbf{E}_{t_0, x_0, w_0}L^3 f(h, x(h), w(h))dh \end{aligned} \tag{17}$$

holds whenever f, Lf, L^2f, and L^3f are uniformly bounded. To see that this is true for $f = \phi_k^2$, we can apply a smooth cutoff function. That is, define $\psi_n = \theta_n(\phi_k^2)\cdot\phi_k^2$, where θ_n is a smooth real function such that

$$\begin{cases} \theta_n(s) = 1, & \text{for } |s| \leq n - 1, \\ \theta_n(s) = 0, & \text{for } |s| \geq n, \\ \theta_n(s) \in [0, 1], & \text{for } n - 1 \leq |s| \leq n. \end{cases} \tag{18}$$

Then, each ψ_n satisfies (33), *and we achieve the desired result by allowing $n \to \infty$.*

Remark 2. *We note that if we only consider the equation* (10) *up to order Δt, we will obtain the similar equation*

$$\epsilon_{k+1}^2 \leq \epsilon_k^2 + C_1\epsilon_k^2\Delta t + C_2\Delta t^2. \tag{19}$$

However, this requires only that $\mu_1 - \mu_2 = \lambda$ with no requirement that $\mu_1 = \frac{1}{2}$. That is, we can be more flexible in our choice of μ_1, but, as we will see in Remark 3, *we will pay a price in order of convergence.*

2.3. Approximation on the Entire Interval

We shall now see what estimate we obtain for the mean-square error for the entire interval $[t_0, t_0 + T]$, using the previous one-step estimates. This is the analogue of the argument of Mil'shtein.

If we define the additional sequence η_k by $\eta_0 = \epsilon_0 = 0$ and

$$\eta_{k+1}^2 = \eta_k^2 + C_1\eta_k^2\Delta t + C_2\eta_k\Delta_t^2 + C_3\Delta t^3, \tag{20}$$

we see that the η-sequence is monotone increasing and supplies an upper bound for the ϵ-sequence. Also, since $\eta_0 = 0$, either $\eta_K \leq \Delta t^2$, where $t_K = t_0 + T$, or there is some k_0 such that $\eta_{k_0} \leq \Delta t^2$, but $\eta_{k_0+1} > \Delta t^2$. In this case, we see that, for $k > k_0$

$$\eta_{k+1}^2 \leq \eta_k^2(1 + (C_1 + C_2 + C_3)\Delta t), \tag{21}$$

so that

$$\begin{aligned}\eta_K^2 &\leq \eta_{k_0}^2(1 + (C_1 + C_2 + C_3)\Delta t)^{(K-k_0)} \\ &\leq \Delta t^2(1 + (C_1 + C_2 + C_3)\Delta t)^K \\ &\leq C\Delta t^2.\end{aligned} \tag{22}$$

In all cases we see that the mean-square error over the entire interval is $\epsilon_K^2 = O(\Delta t^2)$.

Remark 3. *If we again have only that $\mu_1 - \mu_2 = \lambda$, but not that $\mu_1 = \frac{1}{2}$, as in Remark 2, a similar argument will give us that $\epsilon_K^2 = O(\Delta t)$. For instance, we can choose $\mu_1 = \mu_2 = 0$, and $\lambda = 0$, and although we have a rather slow convergence rate, we have the advantage that the scheme no longer uses any derivative of σ; if these are particularly difficult to calculate numerically, this may result in overall computational savings.*

3. A FINITE DIFFERENCE SCHEME

As noted in Remark 3, it sometimes happens that the derivative of σ is difficult to calculate (e.g., when the function σ is given by its values from tables). Here we propose to circumvent this difficulty by replacing the derivative $\partial\sigma/\partial x$ in Mil'shtein's scheme by a finite difference, and we see how this affects the convergence and order of the scheme. For the sake of simplicity, we set $\lambda = 0$.

We use the scheme with time-step

$$\bar{x}_{k+1} = \bar{x}_k + \bar{\sigma}_k\Delta w + \bar{a}_k\Delta t + \frac{1}{4\alpha\Delta t}\bar{\sigma}_k(\sigma(\bar{x}_k + \alpha\Delta t) - \sigma(\bar{x}_k - \alpha\Delta t))(\Delta w^2 - \Delta t). \tag{23}$$

This leads us to define the functions

$$\phi_k(t, x, w) = x + x_k - \bar{x}_k - \bar{\sigma}_k w - \bar{a}_k t - \frac{1}{4\alpha t} \bar{\sigma}_k(\sigma(\bar{x}_k + \alpha t) - \sigma(\bar{x}_k - \alpha t))(w^2 - t), \quad (24)$$

and we find, after some calculation,

$$L\phi_k^2(0, 0, 0) = 2(x_k - \bar{x}_k)(a_k - \bar{a}_k) + (\sigma_k - \bar{\sigma}_k)^2$$

and

$$L^2\phi_k^2(0, 0, 0) = 2(a_k - \bar{a}_k)^2 + 2(x_k - \bar{x}_k)(a_k a_k' + \tfrac{1}{2} a_k''\sigma_k^2) + 2(\sigma_k - \bar{\sigma}_k)(a_k'\sigma_k + a_k\sigma_k' + \tfrac{1}{2}\sigma_k^2\sigma_k'' - \tfrac{1}{2}\bar{\sigma}_k\bar{\sigma}_k'') + (\sigma_k\sigma_k' - \bar{\sigma}_k\bar{\sigma}_k')^2.$$

Therefore, in a manner analogous to the previous scheme, we have that $L\phi_k^2(0, 0, 0) \leq C_1\epsilon_k^2$ and $L^2\phi_k^2(0, 0, 0) \leq C_2\epsilon_k$, where $C_k = C_k(|a|_{C^4}, |\sigma|_{C^4})$. Hence, we again have the inequality

$$\epsilon_{k+1}^2 \leq \epsilon_k^2 + C_1\epsilon_k^2\Delta t + C_2\epsilon_k\Delta t^2 + C_3\Delta t^3, \quad (25)$$

and so, as in section 2.3, we arrive at $\epsilon_K^2 = O(\Delta t^2)$.

References

1. G. N. Mil'shtein, Approximate integration of stochastic differential equations, Theory Prob Appl 19 (1974), 557–562.
2. Peter E. Kloeden and Eckhard Platen, Numerical solution of stochastic differential equations, Application of Mathematics, Vol. 23, Springer-Verlag, Berlin, 1992.
3. Peter E. Kloeden, Eckhard Platen, and Henri Schurz, Numerical solution of SDE through computer experiments, Universitext, Springer-Verlag, Berlin, 1994.