ANNALES DE L'I. H. P., SECTION B

FABIENNE CASTELL
JESSICA GAINES

# The ordinary differential equation approach to asymptotically efficient schemes for solution of stochastic differential equations

*Annales de l'I. H. P., section B*, tome 32, nº 2 (1996), p. 231-250
<http://www.numdam.org/item?id=AIHPB_1996__32_2_231_0>

*Ann. Inst. Henri Poincaré,*

Vol. 32, n° 2, 1996, p. 231-250 *Probabilités et Statistiques*

# The ordinary differential equation approach to asymptotically efficient schemes for solution of stochastic differential equations

by

**Fabienne CASTELL**

Université Paris-Sud, Bât 425 Mathématiques, 91405 Orsay Cedex.

and

**Jessica GAINES**

Department of Mathematics and Statistics,
University of Edinburgh, Mayfield Road, Edinburgh EH9 3JZ.

ABSTRACT. – We consider the numerical approximation to strong solutions of stochastic differential equations (SDE's) using a fixed time step and given only the increments of the Brownian path over each time step. Using the approach generalised by Ben Arous, Castell and Hu, of approximating the solution to an SDE over small time by the solution to a time inhomogeneous ordinary differential equation (ODE), we obtain ODE's which, as the number of time steps increases, yield an asymptotically efficient sequence of approximations to the solution of an SDE, where the concept of asymptotic efficiency is that of Clark and Newton. We distinguish between the two cases of an SDE driven by a one-dimensional Brownian path or satisfying the commutativity condition on the one hand and an SDE driven by a multi-dimensional Brownian path and with a non-commutative Lie algebra on the other hand. When the ODE's presented are solved numerically, the property of asymptotic efficiency is preserved as long as the solution is accurate enough. The methods of this paper represent an alternative and easily generalisable way of looking at the approximation of strong solutions to SDE's.

*Mathematics Subject Classification:* 60 H 10, 65 F 05.

RÉSUMÉ. – Nous nous intéressons aux approximations numériques des solutions fortes d'une équation différentielle stochastique (EDS), utilisant un pas de temps fixe, et les incréments de la trajectoire Brownienne. Nous utilisons l'approche développée par Ben Arous, Castell, et Hu, qui permet d'approcher en temps petit la solution d'une EDS, par la solution d'une équation différentielle ordinaire (EDO) inhomogène en temps. Nous obtenons ainsi des EDO's, qui lorsque le pas de temps diminue, fournissent une suite d'approximations de la solution de l'EDS asymptotiquement efficace, au sens de Clark et Newton. Nous distinguons d'une part le cas d'une EDS conduite par un Brownien de dimension 1, ou satisfaisant la condition de commutativité; d'autre part le cas d'une EDS conduite par un Brownien multi-dimensionnel, et ne satisfaisant pas la condition de commutativité. Lorsque les EDO's obtenues sont résolues numériquement de façon suffisamment précise, la propriété d'efficacité asymptotique est préservée. Les méthodes exposées dans cet article sont des méthodes alternatives et facilement généralisables d'approximation des solutions fortes d'une EDS.

## 1. INTRODUCTION

A modern theoretical approach to understanding the solution of stochastic differential equations (SDE's) is to approximate the solution for small time by the solution of a time inhomogeneous ordinary differential equation (ODE). It is natural to ask whether this approach is useful in the context of the numerical solution of SDE's and whether, in particular, it could lead to asymptotically efficient schemes. In this article we present ODE's, the solutions to which are shown to be asymptotically efficient approximations to the solution of an SDE. Asymptotically efficient numerical approximations can then be obtained by numerical solution of the ODE's using a suitable method.

The problem considered in this article is the numerical approximation of strong solutions to the SDE:

$$\begin{cases} d\xi_t = X_0(\xi_t)\ dt + \sum_{j=1}^{r} X_j(\xi_t)\ \circ dB_t^j \\ \xi_0 = x_0 \end{cases} \tag{1}$$

where $\xi_t$ is an element of $\mathbb{R}^d$, and $B_t = (B_t^1, \cdots, B_t^r)$ is a Brownian motion of dimension $r$. The equation (1) is written in the Stratonovich sense, the corresponding equation expressed in the Itô sense being

$$\begin{cases} d\xi_t = \tilde{X}_0(\xi_t)\ dt + \sum_{j=1}^{r} X_j(\xi_t)\ dB_t^j \\ \xi_0 = x_0 \end{cases}$$

where $\tilde{X}_0 = X_0 + \frac{1}{2}\sum_{j=1}^{r}\sum_{i=1}^{d} \frac{\partial X_j}{\partial x_i} X_j^i$.

Approximations we consider in this article are evaluated at the points of a partition of the interval $[0, T]$, separated by a time step $h = \frac{T}{N}$. So they are given by $(\hat{\xi_0} = x_0, \hat{\xi_1}, \ldots, \hat{\xi_N})$, where $\hat{\xi_k}$ is expected to be close to $\xi_{kh}$. Furthermore we will concentrate on approximations to (1) that depend only on samples of $B_t$ taken at the points of the partition. Under this assumption, it has been shown (*see* [6]) that in general no numerical method can guarantee accuracy along the trajectory, in the mean-square sense, of a higher order than $O(h)$ for a Brownian of dimension one, and than $O(\sqrt{h})$ in dimension greater than one, when the commutativity condition is not satisfied. From the asymptotic expansion of the solution to (1), it is clear that in order to achieve a higher order of convergence, one has to introduce the iterated integrals $\int_{kh}^{(k+1)h} \int_{kh}^{t} dB_s^i\ dB_t^j$, $(k = 0, \cdots, N-1)$, when the Brownian motion is of dimension higher than one. It is only when the vector fields $X_j$, $(j = 1, \cdots, r)$ all commute that these integrals are not needed. Similarly, when the Brownian motion is of dimension one, the integrals $\int_{kh}^{(k+1)h} \int_{kh}^{t} dB_s\ dt$ $(k = 0, \cdots, N-1)$ are needed to obtain a higher order of accuracy than $O(h)$, unless $X_0$ and $X_1$ commute.

A variety of numerical schemes exists for approximating strong solutions to SDE's. These can be found in survey articles and books such as [14], [16] and [11]. Many of the discretisation schemes used are obtained by suitable truncation of the stochastic Taylor series expansion of the solution (*see* [10]).

Among $\mathcal{P}_N$-measurable approximations (where $\mathcal{P}_N$ is the $\sigma$-field generated by $(B_{ih}^j)$ for $i = 1, \ldots, N$, and $j = 1, \ldots, r$), there is a class of numerical schemes with optimal order of strong convergence that have the property of being asymptotically efficient in the $L_2$-sense, as defined by Clark in [5]. These schemes, presented by Newton in [12] and [13], could be considered to be the best available schemes for simulating strong solutions under the constraints laid out above.

It follows from the viewpoint of Ben Arous [3], Castell [4] and Hu [9] that for small random time the solution to an SDE can be approximated, as accurately as required, by the exponential of a Lie series, or in other words by the solution of an ordinary differential equation (ODE) taken at time 1. This ODE is defined by a time-dependent stochastic vector field, which is a linear combination of Lie brackets of the $X_i$, with iterated stochastic integrals as coefficients.

The question then arises as to whether numerical solution methods can be obtained from this theoretical approach, by truncation of the Lie series, and whether, in particular, asymptotically efficient schemes can be obtained. The answer is 'yes'. In this paper we show how asymptotically efficient schemes can be derived, both in the case of a one-dimensional Brownian path and in the multi-dimensional case, and prove that the schemes are indeed asymptotically efficient. Numerical schemes are obtained in two steps. First we obtain ODE's, the (true) solutions of which are asymptotically efficient approximations to the solution of the SDE (1). These ODE's have an appealing simplicity. Then we choose numerical methods to solve the ODE's that preserve the property of asymptotic efficiency.

For convenience of notation, the equation (1) will be written:

$$\begin{cases} d\xi_t = \sum_{j=0}^{r} X_j(\xi_t)\ dB_t^j \\ \xi_0 = x_0 \end{cases} \tag{2}$$

where $B_t^0 = t$. The vector fields $X_j$ are assumed to be smooth enough to ensure the existence of a unique strong solution to (2).

Using the exponential Lie series expansion with a one-dimensional Brownian path and replacing the iterated integrals with their conditional means given the Brownian path, yields the following ODE:

$$\begin{cases} \dfrac{du}{ds} = (B^1_{(n+1)h} - B^1_{nh})X_1(u(s)) + hX_0(u(s)) - \dfrac{h^2}{12}X^{(101)}(u(s)) \\ u(0) = \hat{\xi}_n \end{cases} \tag{3}$$

where $X^{(101)} = [X_1, [X_0, X_1]]$.

The solution to (3) at time 1, is an asymptotically efficient approximation to the solution of (1) with $r = 1$ at time $(n+1)h$. To obtain a numerical scheme it is necessary to choose a numerical method for solving the ODE (3).

The discretisation method used must yield an accurate enough solution to the ODE for the property of asymptotic efficiency to still hold. A practical choice is to use the well-known order 4 Runge-Kutta scheme (so avoiding the calculation of any further derivatives) and to take just one step of length 1 along the ODE.

Similarly, in the multi-dimensional case an asymptotically efficient method is to solve the $d$-dimensional ODE

$$\begin{cases} \dfrac{du}{ds} = \displaystyle\sum_{j=0}^{r}(B^j_{(n+1)h} - B^j_{nh})X_j(u(s)) \\ u(0) = \hat{\xi}_n \end{cases} \tag{4}$$

using a suitable numerical scheme, such as the Heun scheme.

The numerical schemes derived in this paper are clearly not better than those previously used, in that when solving the same SDE and using the same information about the Brownian path there can be no improvement in order of convergence or in efficiency. Yet we consider that these methods constitute a valuable addition to the numerical tool-box. Looking at an old problem from a new perspective is often helpful. The exponential Lie series approach used here throws light on the ODE approximation in [5], for example.

In addition, the method outlined in this paper can be generalised, to apply when using more information about the Brownian path, *i.e.* when iterated stochastic integrals are known or generated. It would be sufficient to replace iterated integrals in the Lie series expansion by their conditional means given the information available.

In section 2 of the paper the results are laid out in detail and the main theorem proved. Some numerical examples follow in section 3.

## 2. ASYMPTOTIC EFFICIENCY RESULTS FOR APPROXIMATIONS BY ODE

Before presenting asymptotic results for approximations (3) or (4), we would like to explain where they come from. First of all, we introduce some notations for later use.

Let $J = (j_1, \ldots, j_m) \in \{0, \ldots, r\}^m$ be a multi-index, and $\sigma \in \sigma_m$ be a permutation of order m.

Let us denote

- $|J| = m$ the cardinality of $J$;

• $||J||$ the order of $J$: $||J|| = |J|+$ number of 0's in $J$;
• $B^J_{s,t}$ the Stratonovich iterated integral:

$$B^J_{s,t} = \int \dots \int_{s<t_1<\dots<t_m<t} \circ dB^{j_1}_{t_1} \dots \circ dB^{j_m}_{t_m};$$

• for $n \in \{1, \dots, N\}$, $B^J_n = B^J_{(n-1)h,nh}$.
• $X^J$ the Lie bracket $[X_{j_1}[\dots[X_{j_{m-1}} X_{j_m}]\dots]$;
• $e(\sigma)$ the number of errors in ordering $\sigma(1), \dots, \sigma(m)$, that is the cardinality of the set $\{j \in \{1, \dots, m-1\} | \sigma(j) > \sigma(j+1)\}$;
• $J \circ \sigma = (j_{\sigma(1)}, \dots, j_{\sigma(m)})$.
• $\mathcal{P}_N = \sigma(B^j_{kh}, k = 1, \dots, N; j = 1 \dots r)$
• for $s \in \mathbb{R}^+$, $\mathcal{B}_s = \sigma(B^1_u, \dots, B^r_u, u \leq s)$
• when $X$ is a vector field on $\mathbb{R}^n$, and when $x \in \mathbb{R}^n$, $\exp(X)(x)$ denotes the solution at time 1 of the ordinary differential equation given by $X$, *i.e.* $\exp(X)(x) = u(1)$ where $u$ is solution to

$$\begin{cases} \dfrac{du}{ds} = X(u(s)) \\ u(0) = x. \end{cases}$$

Then, the result stated in [3] or [4] is the following. (*see* theorem 2.1 of [4]).

THEOREM 2.1. – *Let us assume that $X_0, \dots, X_r$ are $C^\infty$ vector fields, bounded with bounded derivatives. For all integer $p \geq 1$, and for all $s, t \in \mathbb{R}^+, s \leq t$, let us define the stochastic vector field*

$$\zeta^p_{s,t} = \sum_{m=1}^{p} \sum_{||J||=m} c^J_{s,t} X^J,$$

$$\textit{where } c^J_{s,t} = \sum_{\sigma \in \sigma_{|J|}} \frac{(-1)^{e(\sigma)}}{|J|^2 \binom{|J|-1}{e(\sigma)}} B^{J\circ\sigma^{-1}}_{s,t};$$

*and let $R_{p+1}(s,t)$ be the process defined on $\mathbb{R}^d$ by*

$$\xi_t = \exp(\zeta^p_{s,t})(\xi_s) + (t-s)^{\frac{p+1}{2}} R_{p+1}(s,t). \tag{5}$$

*Then, $R_{p+1}$ is bounded in probability when t tends to s. More precisely, a.s, $\exists \alpha, c > 0$ such that $\forall R > c$,*

$$\lim_{t \to s} P\left[ \sup_{s \leq u \leq t} (u-s)^{\frac{p+1}{2}} ||R_{p+1}(s,u)|| \geq R(t-s)^{\frac{p+1}{2}} \Big| \mathcal{B}_s \right] \leq \exp\left(-\frac{R^\alpha}{c}\right).$$

*Remark.* – In [4], the result is given between 0 and $t$. However using the very same proof as in [4], it is easy to check that theorem 2.1 is valid.

One can expect that if the remainder term is dropped, expression (5) will provide a good approximation to the true solution. However, if the integer $p$ appearing in the expansion is chosen greater than 1, the corresponding approximation will not be $\mathcal{P}_N$-measurable, since it involves iterated integrals. Therefore, it seems quite natural to replace the iterated integrals by their $\mathcal{P}_N$-conditional expectation.

Another question which arises when trying to obtain an approximation from (5) is that of the "best" order p which has to be used in order to obtain some asymptotic efficiency results (in the $L_2$-sense). This question is not so easy to answer, since the optimal $p$ depends on the rate of convergence of the $L_2$ optimal error. In the "general" case, $p = 2$ is sufficient to obtain asymptotic efficiency, yielding the approximation

$$\begin{cases} \hat{\xi}_0 = x_0 \\ \hat{\xi}_{n+1} = \exp\left(\sum_{j=0}^{r}(B^j_{(n+1)h} - B^j_{nh})X_j\right)(\hat{\xi}_n). \end{cases} \tag{6}$$

We now state asymptotic results for (6), clarifying what we mean by asymptotic efficiency.

THEOREM 2.2. – *If $X_0$ has continuous bounded derivatives of orders up to 2, and if the $X_i$ $(i = 1, \ldots, r)$ have continuous bounded derivatives of orders up to 3, then $(\hat{\xi}_N)_N$ defined by (6) satisfies for all $c \in \mathbb{R}^d$,*

$$\frac{E\left[\left\langle c, \frac{1}{h^{1/2}}(\xi_T - \hat{\xi}_N)\right\rangle^2 \middle| \mathcal{P}_N\right] + 1}{E\left[\left\langle c, \frac{1}{h^{1/2}}(\xi_T - E[\xi_T|\mathcal{P}_N])\right\rangle^2 \middle| \mathcal{P}_N\right] + 1} \overset{a.s}{\rightarrow} 1. \tag{7}$$

Before proving theorem 2.2, let us make some comments. In general cases, that is when the rate of convergence of $\hat{\xi}_N$ to $\xi_T$ is of order $h^{1/2}$, theorem 2.2 says that $(\hat{\xi}_N)_N$ is a first-order asymptotically efficient sequence of approximations to $\xi_T$ as defined by Clark [5], or Newton [12]. However, in some particular cases (for instance for a Brownian of dimension 1, or when the vector fields $X_i$ commute), the rate of convergence of the $L_2$-optimal error $\xi_T - E[\xi_T|\mathcal{P}_N]$ is greater than $h^{1/2}$. In these cases, theorem 2.2 does not give any information about asymptotic efficiency. It does not even say that $\hat{\xi}_N$ has the best rate of convergence. In order to obtain asymptotic efficiency, we have to go further in the expansion (5).

For instance, for a Brownian of dimension 1, asymptotic efficiency is reached with $p = 4$. Denoting $B^J_{nh,(n+1)h}$ by $B^J_{n+1}$, the approximating

ODE is then given by

$$\begin{cases} \hat{\xi}_0 = x_0 \\ \hat{\xi}_{n+1} = \exp[B^1_{n+1}X_1 + hX_0 + E(B^{(10)}_{n+1} - \frac{1}{2}hB^1_{n+1}|\mathcal{P}_N)X^{(10)} \dots \\ \qquad \dots + \frac{1}{6}E(3B^{(101)}_{n+1} - \frac{1}{2}h(B^1_{n+1})^2|\mathcal{P}_N)X^{(101)}](\hat{\xi}_n). \end{cases}$$

Using $E(B^{(10)}_{n+1}|\mathcal{P}_N) = \frac{hB^1_{n+1}}{2}$, and $E(B^{(101)}_{n+1}|\mathcal{P}_N) = \frac{h}{6}[(B^1_{n+1})^2 - h]$, we obtain

$$\begin{cases} \hat{\xi}_0 = x_0 \\ \hat{\xi}_{n+1} = \exp[(B^1_{(n+1)h} - B^1_{nh})X_1 + hX_0 - \frac{h^2}{12}X^{(101)}](\hat{\xi}_n). \end{cases} \tag{8}$$

THEOREM 2.3. – *If $r = 1$, if $X_0$ has continuous bounded derivatives of orders up to 3, and if $X_1$ has continuous bounded derivatives of orders up to order 4, then $(\hat{\xi}_N)_N$ defined by (8) is a first-order asymptotically efficient sequence of approximations to $\xi_T$, that is*

$$\frac{E\Big[\Big\langle c, \frac{1}{h}(\xi_T - \hat{\xi}_N)\Big\rangle^2 \Big|\mathcal{P}_N\Big] + 1}{E\Big[\Big\langle c, \frac{1}{h}(\xi_T - E[\xi_T|\mathcal{P}_N])\Big\rangle^2 \Big|\mathcal{P}_N\Big] + 1} \xrightarrow{a.s} 1. \tag{9}$$

When the Lie algebra generated by the $X_i$ is Abelian ($[X_i, X_j] = 0$) $\hat{\xi}_{n+1} = \xi_{(n+1)h}$ whenever $p \geq 2$ (*see* [3] or [4]). Therefore, approximation (6) is in this case asymptotically efficient of any order.

*Proof of Theorems* 2.2 *and* 2.3. – We only give the proof of theorem 2.2, following the proof of asymptotic efficiency in [13]; theorem 2.3 can be obtained in a very similar way. First of all, let us remark that

$$\frac{\Big\langle c, \frac{1}{h^{1/2}}(\hat{\xi}_N - E(\xi_T|\mathcal{P}_N))\Big\rangle^2}{E\Big[\Big\langle c, \frac{1}{h^{1/2}}(\xi_T - E(\xi_T|\mathcal{P}_N))\Big\rangle^2 \Big|\mathcal{P}_N\Big] + 1} + 1$$
$$= \frac{E\Big[\Big\langle c, \frac{1}{h^{1/2}}(\xi_T - \hat{\xi}_N)\Big\rangle^2 \Big|\mathcal{P}_N\Big] + 1}{E\Big[\Big\langle c, \frac{1}{h^{1/2}}(\xi_T - E(\xi_T|\mathcal{P}_N))\Big\rangle^2 \Big|\mathcal{P}_N\Big] + 1}.$$

Therefore, since $\forall p \geq 1$, $\sup_{n,N} E\left[\left|\left|\frac{1}{h}(\xi_T - E(\xi_T|\mathcal{P}_N)\right|\right|^p\right] < +\infty$, it is sufficient to show that $\frac{1}{h^{1/2}}\left(E(\xi_T|\mathcal{P}_N) - \hat{\xi}_N\right) \stackrel{a.s}{\rightarrow} 0$. It is this latter condition that will be used to check for first-order asymptotic efficiency. During the proof, we will use the following notation.

• When $X$ is a vector-field on $\mathbb{R}^d$, $L_X$ is the Lie derivative associated with $X$

$$L_X := \sum_{j=1}^{d} X^j \frac{\partial}{\partial x_j}$$

where $X^j$ is the $j$-th coordinate of $X$;

• If $x = (x_n, n = 1, \ldots, N)_N$ is a sequence of a vector-valued (or matrix-valued) processes, $x$ satisfies $P_k$ (or $x$ is at least of order k) iff

$$\sup_{n,N} E[||N^{\frac{k}{2}} x_n||^p] < \infty \quad \forall p \geq 1.$$

It is clear that

– If $x$ and $y$ satisfy $P_k$, so does $x + y$.

– If $k < l$, and if $x$ satisfies $P_l$, $x$ satisfies $P_k$.

– If $x$ satisfies $P_k$, and if $y_n = E[x_n|\mathcal{F}_n]$ (for some $\sigma$-fields $(\mathcal{F}_n, n = 1, \ldots, N)_N$), then $y = (y_n, n = 1, \ldots, N)_N$ satisfies $P_k$.

– If $x_n = B_n^J$, $x$ satisfies $P_{||J||}$.

• We will denote by $\bar{X}$ the $\mathcal{P}_N$-conditional expectation of $X$. If in addition, $X$ is $\mathcal{B}_n$-measurable, $\bar{X}$ will mean a $\mathcal{B}_n$-measurable version of $E[X|\mathcal{P}_N]$. Such a version exists since

$$E[E[\bullet|\mathcal{B}_n]|\mathcal{P}_N] = E[E[\bullet|\mathcal{P}_N]|\mathcal{B}_n] = E[\bullet|B_h, \ldots, B_{nh}]. \tag{10}$$

Finally, we denote $\tilde{X} = X - \bar{X}$.

The stochastic Taylor formula gives the following expansion for $\xi_{(n+1)h}$ (*see* for instance [1], or [10])

$$\xi_{(n+1)h} = \xi_{nh} + \sum_{||J|| \leq 3} p_J(\xi_{nh}) B_{n+1}^J + R_4(n) \tag{11}$$

where

• under the assumptions of theorem 2.2, $(R_4(n))$ satisfies $P_4$;

• for $J = (j_1, \ldots, j_m)$, $p_J(x) = L_{X_{j_1}} \ldots L_{X_{j_{m-1}}} X_{j_m}(x)$.

Using the Taylor formula, we now expand $p_J$ around $\bar{\xi}_{nh}$. If $Dp_J$ denotes the Jacobian matrix of $p_J$, we obtain,

$$\xi_{(n+1)h} = \xi_{nh} + \sum_{||J||=1} \left[ p_J(\bar{\xi}_{nh}) + (Dp_J(\bar{\xi}_{nh}) + \delta_n^J)\tilde{\xi}_{nh} \right] B_{n+1}^J + \sum_{2 \leq ||J|| \leq 3} \left[ p_J(\bar{\xi}_{nh}) + \delta_n^J \right] B_{n+1}^J + R_4(n).$$

where the $\delta_n^J$ are of the same order as $\tilde{\xi}_{nh}$, that is $\sqrt{h}$, and are $\mathcal{B}_n$-measurable. Taking the $\mathcal{P}_N$-conditional expectation yields

$$\bar{\xi}_{(n+1)h} = \bar{\xi}_{nh} + \sum_{||J||=1} \left( p_J(\bar{\xi}_{nh}) B_{n+1}^J + \overline{\delta_n^J \tilde{\xi}_{nh} B_{n+1}^J} \right) + \sum_{2 \leq ||J|| \leq 3} \left( p_J(\bar{\xi}_{nh}) \overline{B_{n+1}^J} + \overline{\delta_n^J B_{n+1}^J} \right) + \overline{R_4(n)}. \quad (12)$$

We now give an expansion for $\hat{\xi}_n$. For each multi-index $J$, let us introduce the vector field $\beta_J$ defined by:

$$\beta_J = \sum_{\sigma \in \sigma_{|J|}} \frac{(-1)^{e(\sigma)}}{|J|^2 \binom{|J|-1}{e(\sigma)}} X^{J \circ \sigma}.$$

It is easy to check that a.s

$$\forall x, \ \forall p, \quad \exp\left( \sum_{||J|| \leq p} \overline{c_{n+1}^J} X^J \right)(x) = \exp\left( \sum_{||J|| \leq p} \beta_J \overline{B_{n+1}^J} \right)(x)$$

where $c_{n+1}^J = c_{nh,(n+1)h}^J$. Therefore,

$$\hat{\xi}_{n+1} = \exp\left( \sum_{||J|| \leq 2} \overline{c_{n+1}^J} X^J \right)(\hat{\xi}_n) = \exp\left( \sum_{||J|| \leq 2} \overline{B_{n+1}^J} \, \beta_J \right)(\hat{\xi}_n).$$

Now, $\forall f \in C_b^1(\mathbb{R}^d)$, for all $s, t$, $s \leq t$, and for every vector-field $X$ on $\mathbb{R}^d$,

$$f(\exp(sX)(x)) = f(x) + \int_s^t L_X f(\exp(uX)(x)) du.$$

Using this formula recursively with $\sum_{||J||\leq 2} \overline{B^J_{n+1}}\beta_J$ in place of $X$, we obtain $\forall n \leq N-1$,

$$\hat{\xi}_{n+1} = \hat{\xi}_n + \sum_{p=1}^{3} \sum_{\substack{k, J_1, \ldots, J_k \\ ||J_1||+\ldots ||J_k||=p \\ ||J_i||\leq 2}} \frac{1}{k!} \times \overline{B^{J_1}_{n+1}} \ldots \overline{B^{J_k}_{n+1}}\, L_{\beta_{J_1}} \ldots L_{\beta_{J_{k-1}}} \beta_{J_k}(\hat{\xi}_n) + S_4(n) \tag{13}$$

where $S_4(n)$ satisfies $P_4$. But, in all the cases where $||J_1||+\ldots+||J_R|| \leq 3$, $||J_i|| \leq 2$, it is easy to check that

$$\overline{B^{J_1}_{n+1}} \ldots \overline{B^{J_k}_{n+1}} = \overline{B^{J_1}_{n+1} \ldots B^{J_k}_{n+1}}. \tag{14}$$

Let us now define, for all multi-indices $J_1, \ldots, J_k$, and for every multi-index $K$ a "shuffle" of $J_1, \ldots, J_k$, the coefficients $d_{J_1,\ldots,J_k}(K)$ by:

$$B^{J_1}_t \ldots B^{J_k}_t = \sum_{\substack{K, \text{shuffle} \\ of\, J_1 \ldots J_k}} d_{J_1,\ldots,J_k}(K) B^K_t. \tag{15}$$

We refer the reader to [7] for the definition of the shuffles and the existence of the coefficients $d_{J_1,\ldots,J_k}(K)$.
It follows from (13), (14) and (15) that

$$\hat{\xi}_{n+1} = \hat{\xi}_n + \sum_{||K||\leq 3} \overline{B^K_{n+1}} \sum_{\substack{k, J_1, \ldots, J_k \text{ such that} \\ K \text{ shuffle of } J_1, \ldots J_k}} \frac{1}{k!} d_{J_1,\ldots,J_k}(K) \times L_{\beta_{J_1}} \ldots L_{\beta_{J_{k-1}}} \beta_{J_k}(\hat{\xi}_n) - \sum_{||K||=3} \overline{B^K_{n+1}}\, \beta_K(\hat{\xi}_n) + S_4(n). \tag{16}$$

We claim that

LEMMA 2.1.

$$\forall K, \quad \forall x, \quad p_K(x) = \sum_{\substack{k, J_1, \ldots, J_k \text{ such that} \\ K \text{ shuffle of } J_1, \ldots J_k}} \frac{1}{k!} d_{J_1,\ldots,J_k}(K) L_{\beta_{J_1}} \ldots L_{\beta_{J_{k-1}}} \beta_{J_k}(x).$$

The reader is referred to the end of the section for the proof of lemma 2.1.
From lemma 2.1, it results that

$$\hat{\xi}_{n+1} = \hat{\xi}_n + \sum_{||K||\leq 3} p_K(\hat{\xi}_n)\overline{B^K_{n+1}} - \sum_{||K||=3} \overline{B^K_{n+1}}\beta_K(\hat{\xi}_n) + S_4(n). \tag{17}$$

Expressions (12) and (17) yield the following difference equation for the error $e_n = \bar{\xi}_{nh} - \hat{\xi}_n$.

$$e_{n+1} = (I + Y_{n+1})e_n + d_{n+1}$$

where

$$Y_{n+1} = \sum_{1\le||J||\le 2} A_J(\bar{\xi}_{nh}, \hat{\xi}_n)\overline{B^J_{n+1}}$$

and

$$A_J(x,y) = \int_0^1 Dp_J((1-u)x + uy)du;$$

$$d_{n+1} = \sum_{||J||=3} \Big(p_J(\bar{\xi}_{nh}) - p_J(\hat{\xi}_{nh}) + \beta_J(\hat{\xi}_n)\Big)\overline{B^J_{n+1}} + \sum_{||J||=1} \overline{\delta^J_n \tilde{\xi}_{nh} B^J_{n+1}} \dots$$

$$\dots + \sum_{2\le||J||\le 3} \overline{\delta^J_n B^J_{n+1}} + \overline{R_4(n)} - S_4(n).$$

We now use the following lemma proved by Newton in [12].

LEMMA 2.2 (theorem 1 of [12]). – *Let $(x_n; n = 0,\dots,N)_N$ be a sequence of vector-valued processes defined by the following difference equations:*

$$x_0 = 0$$
$$x_{n+1} = (I + Y_{n+1})x_n + d_{n+1}$$

*where $(Y_n; n = 0,\dots,N)_N$ and $(d_n; n = 0,\dots,N)_N$ are sequences of matrix-valued and vector-valued processes. Furthermore, let $(\mathcal{F}_n; n = 0,\dots,N)_N$ be a sequence of filtrations such that $Y_n$ and $d_n$ are $\mathcal{F}_n$-adapted. If for some $p \ge 2$, and some $K < \infty$, the following conditions are fulfilled:*

1. $\sup_{n,N} ||NE(Y_n|_{\mathcal{F}_{n-1}})||^p \le K$ a.s,
2. $\sup_{n,N} E(||N^{1/2}Y_n||^p|_{\mathcal{F}_{n-1}}) \le K$ a.s,
3. $E(d_n|_{\mathcal{F}_{n-1}})$ *fulfills* $P_2$,
4. $d_n$ *fulfills* $P_1$,

*then $(x_n)$ has the following property:* $\sup_{n,N} E(||x_n||^p) < \infty$.

In our case, $Y_{n+1}$ satisfies properties 1 and 2 of lemma 2.2 with $\mathcal{F}_n := \mathcal{B}_{nh}$. $d_{n+1}$ fulfills $P_3$. Moreover,

- for $||J|| = 3$, $E\Big[\Big(p_J(\bar{\xi}_{nh}) - p_J(\hat{\xi}_n) + \beta_J(\hat{\xi}_n)\Big)\overline{B^J_{n+1}}|\mathcal{B}_{nh}\Big]$

$$= (p_J(\bar{\xi}_{nh}) - p_J(\hat{\xi}_n) + \beta_J(\hat{\xi}_n))E\Big[\overline{B^J_{n+1}}|\mathcal{B}_{nh}\Big] = 0;$$

- $||J|| = 1$, $E\left[\overline{\delta_n^J \tilde{\xi}_{nh} B_{n+1}^J} | \mathcal{B}_{nh}\right] = \overline{\delta_n^J \tilde{\xi}_{nh}} E\left[B_{n+1}^J | \mathcal{B}_{nh}\right] = 0$;
- for $2 \leq ||J|| \leq 3$, $E\left[\overline{\delta_n^J B_{n+1}^J} | \mathcal{B}_{nh}\right] = E\left[\delta_n^J B_{n+1}^J | B_h, \dots B_{nh}\right]$

$$= E\left[\delta_n^J | B_h, \dots B_{nh}\right] E(B_{n+1}^J) = 0.$$

Therefore, $E[d_{n+1} | \mathcal{B}_{nh}]$ is at least of order 4. We deduce that $(e_n)$ verifies $P_2$, which allows us to end the proof of theorem 2.2. ■

It follows from the proof that for a numerical scheme to be asymptotically efficient, it is sufficient that its expansion up to order 2 is the same as (12), and that the terms of order 3 have null $\mathcal{B}_{nh}$-conditional expectation. Therefore, the approximate solution to the ODE (4) will have to fulfill this condition to remain asymptotically efficient.

*Proof of Lemma* 2.1. – Let $\xi_t^\varepsilon$ be solution to the Stratonovich stochastic differential equation

$$\begin{cases} d\xi_t^\varepsilon = \sum_{i=1}^{r} \varepsilon X_i(\xi_t^\varepsilon) o dB_t^i + \varepsilon^2 X_0(\xi_t^\varepsilon) dt \\ \xi_0^\varepsilon = x. \end{cases} \tag{18}$$

It has been proved in [4] or [3] that a.s, for all $l \geq 0$,

$$\xi_t^\varepsilon = \exp\left(\sum_{||J|| \leq l} \varepsilon^{||J||} B_t^J \beta_J\right)(x) + \varepsilon^{l+1} R_{l+1}(\varepsilon, t)$$

where $R_{l+1}$ satisfies for some $\alpha$ and $c > 0$, and for all $R > c$,

$$\lim_{\varepsilon \to 0} P\left[\sup_{0 \leq s \leq T} ||R_{l+1}(\varepsilon, t)|| \geq R\right] \leq \exp\left(\frac{-R^\alpha}{cT}\right).$$

Therefore, the Taylor expansion of $\xi_t^\varepsilon$ with respect to $\varepsilon$ is given by

$$\xi_t^\varepsilon = x + \sum_{||K|| \leq l} \varepsilon^{||K||} B_t^K \left(\sum_{\substack{J_1, \dots, J_k \text{such that} \\ K \text{shuffle of } J_1, \dots, J_k}} \frac{1}{k!} \right.$$
$$\left. \times d_{J_1, \dots, J_k}(K) L_{\beta_{J_1}} \dots L_{\beta_{J_{k-1}}} \beta_{J_k}(x)\right) + \varepsilon^{l+1} Q_{l+1}(\varepsilon, t)$$

where $Q_{l+1}$ satisfies the same property as $R_{l+1}$.

But using (18), this Taylor expansion is nothing but

$$\xi_t^\varepsilon = x + \sum_{||K|| \le l} \varepsilon^{||K||} p_K(x) B_t^K + \varepsilon^{l+1} S_{l+1}(\varepsilon, t)$$

where $S_{l+1}$ satisfies the same property as $R_{l+1}$. Identifying these two expansions, it results that a.s., for all $x \in \mathbb{R}^d$, and for all $p \le l$,

$$\sum_{||K||=p} p_K(x) B_t^K$$

$$= \sum_{||K||=p} \left( \sum_{\substack{J_1,\ldots,J_k \text{ such that} \\ K \text{shuffle of } J_1,\ldots,J_k}} \frac{1}{k!} d_{J_1,\ldots,J_k}(K) L_{\beta_{J_1}} \ldots L_{\beta_{J_{k-1}}} \beta_{J_k}(x) \right) B_t^K.$$

Now using proposition 2.1 in [2], which states the linear independence of the iterated integrals $B_t^K$ for $||K||$ fixed, we obtain the assertion of lemma 2.1. ■

## 3. NUMERICAL EXAMPLES

In [13] Newton gave an extensive set of examples to illustrate the advantage of using an asymptotically efficient method when solving SDE's numerically. The saving in computation time was demonstrated by taking many different Brownian paths and calculating various statistics. Our aim here is not to repeat such a demonstration, but simply to show that the theoretical approach of approximating the solution to an SDE by the solution to an ODE gives rise to practical numerical methods.

We have simulated solutions to two SDE's as illustrations of the asymptotically efficient methods developed in this paper. The first example has one-dimensional noise, illustrating use of the ODE (3), while the second example has a Brownian path of dimension two and we have used the ODE (4).

In each case we call our method, consisting of numerical approximation to solutions of the ODE (3) or (4), the AE-ODE method. We refer to the two Stratonovich Runge-Kutta methods of Newton [13] by the names FRKS and ERKS used in that paper. The FRKS discretisation method is a simple first order Runge-Kutta scheme (involving the minimum number of function evaluations required by such a scheme), while the ERKS method is an asymptotically efficient first order Runge-Kutta scheme. For our first example, we generated numerical solutions using each of the AE-ODE, the FRKS and the ERKS methods.

The equation is a random linear oscillator (*see* [15]) satisfying the SDE

$$\begin{cases} du_t = v_t \, dt \\ dv_t = (-2c\omega v_t - \omega^2 u_t) \, dt + \sigma u_t \, dB_t \end{cases} \tag{19}$$

where $u$ is the position and $v$ the velocity of a particle under the simple harmonic restoring force $-\omega^2 u$, the damping force $-2c\omega$ and a random white noise force proportional to $u$. The chosen parameter values were $c = 1/2$, $\omega = 1$ and $\sigma = 2$ and the initial values $u(0) = 1$, $v(0) = 0$. The final time was $T = 10$ and the selected step size $h = 2^{-6}$.

To obtain a numerical solution to the ODE (3) over one time step along the SDE (19), we used the well-known fourth order Runge-Kutta method and took just one step of length 1. The Runge-Kutta scheme can be written

$$\hat{\xi}_{n+1} = \hat{\xi}_n + \tilde{h}(\frac{1}{6} f_0 + \frac{1}{3} f_1 + \frac{1}{3} f_2 + \frac{1}{6} f_3)(\hat{\xi}_n)$$

where

$$\begin{aligned} f_0(\xi) &= f(\xi) \\ f_1(\xi) &= f(\xi + \frac{1}{2}\tilde{h} f_0(\xi)) \\ f_2(\xi) &= f(\xi + \frac{1}{2}\tilde{h} f_1(\xi)) \\ f_3(\xi) &= f(\xi + \tilde{h} f_2(\xi)) \end{aligned}$$

with

$$f(\hat{\xi}_n) = \left[ B^1_{n+1} X_1 + h X_0 - \frac{h^2}{12} X^{(101)} \right] (\hat{\xi}_n). \tag{20}$$

As in the previous section, we have written $B^1_{n+1}$ for $B^1_{(n+1)h} - B^1_{nh}$ and the time step along the ODE is denoted by $\tilde{h}$. On expanding the approximate solution produced by this method as a Taylor series over one step we get

$$\hat{\xi}_{n+1} = \hat{\xi}_n + \left( \tilde{h} f + \frac{\tilde{h}^2}{2} L_f f + \frac{\tilde{h}^3}{6} L_f L_f f + \frac{\tilde{h}^4}{24} L_f L_f L_f f \right) (\hat{\xi}_n) + R \tag{21}$$

where the remainder $R$ contains only terms that are products of at least 5 partial derivatives of $f$ (including $f$ itself as a derivative of order 0).

Taking $\tilde{h} = 1$ and substituting for $f(\hat{\xi}_n)$ from (20), we then obtain

$$\hat{\xi}_{n+1} = \hat{\xi}_n + \Big( B^1_{n+1} X_1 + hX_0 + \frac{1}{2}(B^1_{n+1})^2 L_{X_1} X_1 - \frac{1}{12} h^2 X_{101}$$
$$+ \frac{1}{2} h B^1_{n+1} (L_{X_0} X_1 + L_{X_1} X_0) + \frac{1}{6} L_{X_1} L_{X_1} X_1$$
$$+ \frac{1}{6} h (B^1_{n+1})^2 (L_{X_0} L_{X_1} X_1 + L_{X_1} L_{X_0} X_1 + L_{X_1} L_{X_1} X_0)$$
$$+ \frac{1}{24} L_{X_1} L_{X_1} L_{X_1} X_1 \Big)(\hat{\xi}_n) + \tilde{R}. \tag{22}$$

The remainder $\tilde{R}$ contains terms of order $h^{5/2}$ or higher. It is clear that the expansion in (22) is, up to and including terms of order $h^2$, identical to the expansion of the true solution of the ODE (3), and that therefore the property of asymptotic efficiency holds for this approximate solution. We could have chosen any other numerical solution method that yields an expansion with terms of all orders up to and including $h^{3/2}$ identical to those in the expansion of the true solution of the ODE and with terms of order $h^2$ such that the mean error in these terms is zero over each step.

The solution to (19) cannot be calculated analytically, so an accurate approximation to the true solution was obtained by using both a very small step size and a higher order discretisation scheme. By generating both the increments along the Brownian path, $B^1_n$, and the integrals $B^{(10)}_n$, $n = 1 \dots N$, for $T = Nh$, we can obtain an approximation of order $3/2$. (This higher order approximation was also obtained by local solution of an ODE.)

We have plotted in Figure 1 the differences from the accurate solution of the simulations obtained using the FRKS, ERKS and AE-ODE schemes. Since the trajectories obtained using ERKS and AE-ODE are extremely close to each other, it is hard to distinguish them on the graphs. This is because the two schemes only differ in fourth order terms.

The second example chosen is the bilinear Stratonovich SDE

$$d\xi_t = A\xi \; dt + B\xi \circ dB^1_t + C\xi \circ dB^2_t \tag{23}$$

where $\xi$ is a vector of dimension two. The numerical values chosen for the matrices $A$,$B$,$C$ are as follows: $A(1,1) = A(2,2) = 1$, $A(1,2) = A(2,1) = -0.75$, $B(1,1) = C(2,1) = 0.5$, $B(2,2) = C(1,2) = -0.2$ and all other elements are zero. With these values the commutativity condition is not satisfied.

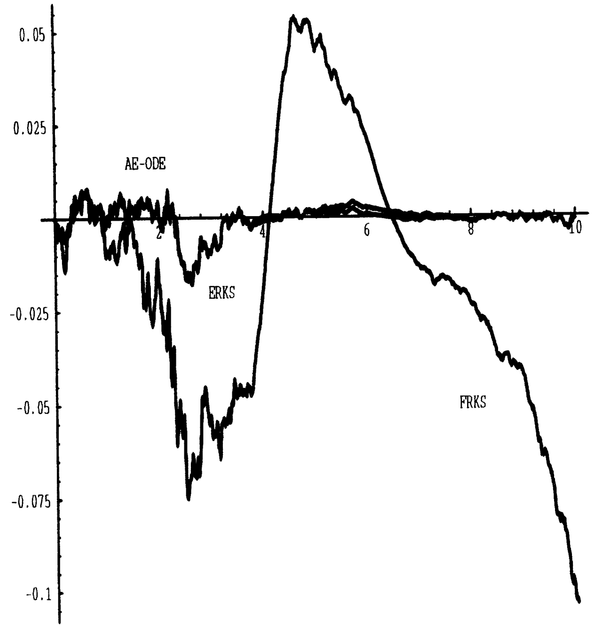


Fig. 1. – Difference from accurate solution.

This time we obtain two different approximate solutions: that obtained by numerical solution of the ODE (4), and an approximation using the well-known Euler discretisation scheme. We have plotted in Figure 2 the differences from an accurate solution of simulations obtained using these two schemes with a step-size of $h = 2^{-8}$. The accurate solution was obtained using a step-size of $h = 2^{-15}$ and an order $O(h)$ scheme using both the increments along the Brownian path, $B_n^1$, $B_n^2$ and the integrals $B_n^{(12)} - B_n^{(21)}$, $n = 1 \dots N$, for $T = Nh$. (*See* [8] for a description of how these integrals were generated.)

The numerical scheme chosen for solution of the ODE is the Heun scheme:

$$\hat{\xi}_{n+1} = \hat{\xi}_n + \frac{\tilde{h}}{2}(g_0 + g_1)(\hat{\xi}_n)$$

where

$$g_0(\xi) = g(\xi)$$
$$g_1(\xi) = g(\xi + \tilde{h} g_0(\xi))$$

with

$$g(\hat{\xi}_n) = \sum_{j=0}^{r} B_{n+1}^j X_j(\hat{\xi}_n). \tag{24}$$

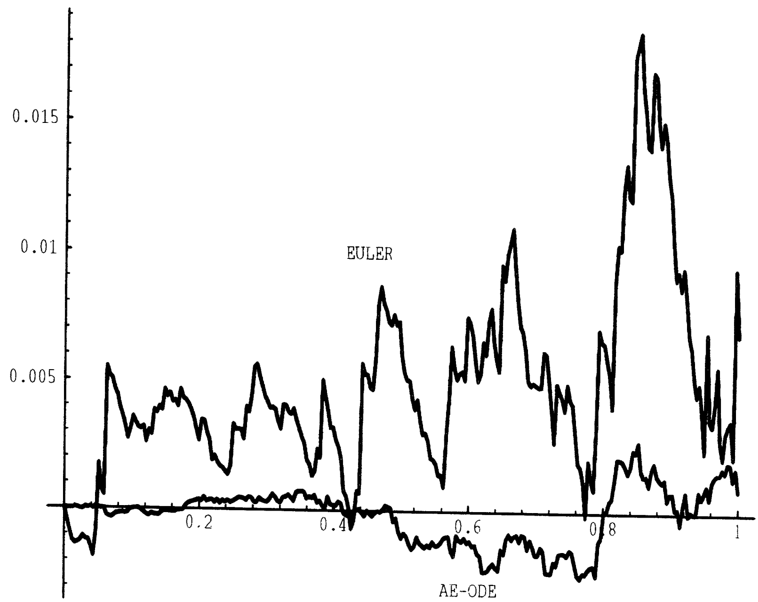


Fig. 2. – Differences from accurate solution.

We will see that the discretisation scheme obtained is identical to that obtained when solving the SDE (1) directly using the Heun method for SDE's.

Again expanding (24) as a Taylor series over one step, we get

$$\hat{\xi}_{n+1} = \hat{\xi}_n + \left( \tilde{h} g + \frac{\tilde{h}^2}{2} L_g g \right) (\hat{\xi}_n) + R \tag{25}$$

where the remainder $R$ contains only terms that are products of at least 3 partial derivatives of $g$ (including $g$ itself as a derivative of order 0).

Taking $\tilde{h} = 1$ and substituting for $g(\hat{\xi}_n)$ from (24), we then obtain

$$\hat{\xi}_{n+1} = \hat{\xi}_n + \left( \sum_{j=0}^{r} B_{n+1}^j X_j + \frac{1}{2} \sum_{j,k=1}^{r} B_{n+1}^j B_{n+1}^k L_{X_j} X_k \right) (\hat{\xi}_n) + \tilde{R}$$

where the remainder $\tilde{R}$ contains terms of order at least $h^{3/2}$. This expansion matches the expansion of the true solution of the ODE (4) up to and including terms of order $h$. This is a necessary and sufficient condition for an approximate solution to the ODE (4) to remain asymptotically efficient. Note that whereas one might suppose that the terms of order $h^{3/2}$ were important for obtaining asymptotic efficiency, in fact they are not, due to all such terms having $\mathcal{B}_n$-conditional mean zero.

The generalisation of the Heun scheme used commonly for solving the SDE (1) is

$$\hat{\xi}_{n+1} = \hat{\xi}_n + \frac{1}{2}\sum_{j=0}^{r} B_{n+1}^j (X_j + \tilde{X}_j)(\hat{\xi}_n)$$

where

$$\tilde{X}_j(\xi) = X_j\left(\xi + \sum_{j=0}^{r} B_{n+1}^j X_j(\xi)\right).$$

It is known that this is an asymptotically efficient scheme in the general case. It is easy to see that this scheme is in fact identical to that obtained by solving the ODE (4) using the Heun scheme for ODE's.

## REFERENCES


[1] R. AZENCOTT, *Formule de Taylor stochastique et développement asymptotique d'intégrales de Feynmann*, in Séminaire de Probabilités XVI, Supplément: Géométrie différentielle stochastique, Springer-Verlag, 1980/81, pp. 237-284.

[2] V. BALLY, On the connection between the Malliavin covariance matrix and Hörmander's condition, *Journal of Functional Analysis*, Vol. **96**, 1991, pp. 219-255.

[3] G. BEN AROUS, Flots et séries de Taylor stochastiques, *Probab. Theory Related Fields*, Vol. **81**, 1989, pp. 29-77.

[4] F. CASTELL, Asymptotic expansion of stochastic flows, *Probab. Theory Related Fields*, Vol. **96**, 1993, pp. 225-239.

[5] J. M. C. CLARK, *An efficient approximation for a class of stochastic differential equations*, in Advances in filtering and optimal stochastic control, Proceedings of IFIP-WG7/1 Working Conference, Cocoyoc, Mexico, 1982, W. H. Fleming and L. G. Gorostiza, eds., no. 42 in *Lecture Notes in Control and Information Sciences*, Springer-Verlag, Berlin, 1982.

[6] J. M. C. CLARK and R. J. CAMERON, *The maximum rate of convergence of discrete approximations for stochastic differential equations*, in Stochastic Differential Systems, B. Grigelionis, ed., no. 25 in *Lecture Notes in Control and Information Sciences*, Springer-Verlag, Berlin, 1980.

[7] J. G. GAINES, *The algebra of iterated stochastic integrals*. To appear in Stochastics and Stochastic Reports.

[8] J. G. GAINES and T. J. LYONS, *Random generation of stochastic area integrals*. To appear in SIAM *J. of Applied Math.*

[9] Y. Z. HU, Série de Taylor stochastique et formule de Campbell-Haussdorff, d'après Ben Arous, in Séminaire de Probabilités XXV, J. Azema, P. A. Meyer, and M. Yor, eds., no. 1485 in *Lecture Notes in Mathematics*, Springer-Verlag, 1991/92, pp. 579-586.

[10] P. E. KLOEDEN and E. PLATEN, Stratonovich and Itô stochastic Taylor expansions, *Math. Nachr.*, Vol. **151**, 1991, pp. 33-50.

[11] P. E. KLOEDEN and E. PLATEN, Numerical Solution of Stochastic Differential Equations, Vol. **23** of *Applications of Mathematics*, Springer-Verlag, 1992.

[12] N. J. NEWTON, An asymptotically efficient difference formula for solving stochastic differential equations, *Stochastics*, Vol. **19**, 1986, pp. 175-206.

[13] N. J. NEWTON, Asymptotically efficient Runge-Kutta methods for a class of Itô and Stratonovich equations, SIAM *J. of Applied Mathematics*, Vol. **51**, 1991, pp. 542-567.
[14] E. PARDOUX and D. TALAY, Discretization and simulation of stochastic differential equations, *Acta Appl. Math.*, Vol. **3**, 1985, pp. 23-47.
[15] L. C. G. ROGERS and D. WILLIAMS, *Diffusions, Markov Processes and Martingales 2, Itô Calculus*, John Wiley and Sons, 1987.
[16] D. TALAY, *Simulation and numerical analysis of stochastic differential systems: A review*, Tech. Report 1313, INRIA, 1990.

*(Manuscript received November 16, 1992; revised version received January 31, 1994.)*