London Mathematical Society ISSN 1461–1570

NUMERICAL ANALYSIS OF EXPLICIT ONE-STEP METHODS FOR STOCHASTIC DELAY DIFFERENTIAL EQUATIONS

CHRISTOPHER T. H. BAKER AND EVELYN BUCKWAR

Abstract

We consider the problem of strong approximations of the solution of stochastic differential equations of Itô form with a constant lag in the argument. We indicate the nature of the equations of interest, and give a convergence proof in full detail for explicit one-step methods. We provide some illustrative numerical examples, using the Euler–Maruyama scheme.

1. *Introduction*

We shall study the evolutionary problem for Itô stochastic delay differential equations or *SDDE*s of the form

$$dX(t) = f\big(t, X(t), X(t-\tau)\big)\,dt + g\big(t, X(t), X(t-\tau)\big)\,dW(t), \qquad t \in [0, T], \tag{1}$$

$$X(t) = \Psi(t), \qquad t \in [-\tau, 0], \tag{2}$$

(with the 'lag' $\tau > 0$). SDDEs are a generalization of both deterministic delay differential equations (DDEs) and stochastic ordinary differential equations (SODEs). In many areas of science (such as population problems, and the study of materials or systems with memory) there has been an increasing interest in the investigation of functional differential equations incorporating memory or 'after-effect'. These systems frequently provide *more realistic models* for phenomena that display time-lag or after-effect than do their instantaneous counterparts. Deterministic models require that the parameters involved be completely known, though in the original problem one often has insufficient information on parameter values. These may fluctuate due to some external or internal 'noise', which is random—or at least appears to be so. Thus we move from deterministic problems to *stochastic problems* (or, respectively, stochastic ordinary differential equations [SODEs], stochastic delay differential equations [SDDEs], and so forth). A range of basic stochastic concepts are considered in [**29**]. For the theoretical prerequisites on SODEs we refer to [**1**] or [**14**]; for the theory of SDDEs, see (for example) [**16, 19, 23**].

In general, there is no analytical closed-form solution of the problems considered here, and we usually require numerical techniques to investigate the models quantitatively. The analysis of numerical methods for SDDEs is based on the numerical analysis of DDEs and the numerical analysis of SODEs. We refer to [**3, 30**] for discussions of issues in the numerical treatment of DDEs. For an overview of applications and objectives of numerical methods for SODEs, see [**6**], [**7**], [**25**] or [**27**]; for more extensive treatments, see [**15, 21**]. There are few articles on numerical analysis of SDDEs to date (see [**17, 28**]); the relevant

---

The second author is supported by the TMR Grant ERBFMBICT983282.
Received 18 January 2000, revised 11 October 2000*; published 24 November 2000.*
2000 Mathematics Subject Classification 65C20

numerical analysis has received insufficient attention in the literature, and the present paper is intended to rectify this in some small measure.

In this paper we shall be interested in obtaining approximations to strong solutions of an SDDE. One reason to be interested in this kind of approximation is the wish to examine the dependence of the solution on the initial function, or on parameters that are contained in the definition of the SDDE.

This paper is organized as follows: in Section 2 we shall describe model problems, Section 3 contains background material of the theory of SDDEs, Section 4 is devoted to the mean-square analysis of general explicit one-step methods for SDDEs with constant lag, and in Section 5 we shall prove consistency of the Euler–Maruyama method. The main theoretical results are Theorems 3, 5 and 6. In Section 6 we present numerical illustrations, and in the last section we draw the attention of the reader to open problems in this area.

2. *Some model problems*

We shall use a brief discussion of some model problems to introduce SDDEs to the reader. A more extensive discussion can be found in [**24**].

Example 1 (Cell population growth). In a recent paper [**4**] several mathematical models for cell proliferation are discussed. The deterministic models presented there range from exponential growth to a neutral delay differential equation. The extension of the exponential growth model by the introduction of delay terms can be justified by assuming that, once activated, cell division is not instantaneous. Thus the use of delay differential equations greatly increases the range of qualitative behaviour that can be modelled.

Consider a large population $N(t)$ of cells at time t evolving with a proportionate rate $\rho_0 > 0$ of 'instantaneous' and a proportionate rate ρ_1 of 'delayed' cell growth. (The population is assumed to be *large* in order to justify continuous as opposed to discrete growth models. By 'instantaneous' cell growth, we mean that the rate of growth is dependent on the *current* cell population, and by 'delayed' cell growth, we mean that the rate of growth is dependent on some *previous* cell population.) The number $\tau > 0$ denotes the average cell-division time. A model is then

$$\begin{aligned} N'(t) &= \rho_0 N(t) + \rho_1 N(t-\tau), & t \geqslant 0, \\ N(t) &= \Psi(t), & -\tau \leqslant t < 0. \end{aligned} \tag{3}$$

This equation may also be used to model a single-sex population evolving with a constant birth rate $\rho_1 > 0$ and a constant death rate per capita ($\rho_0 < 0$). Then the occurrence of the delay in the birth term denotes the development (maturation) period.

Now assume that these biological systems operate in a noisy environment with an overall noise rate that is distributed like white noise, $\beta dW(t)$. Then we shall have a population $X(t)$, now a random process, with growth that is described by the SDDE

$$\begin{aligned} dX(t) &= \big(\rho_0 X(t) + \rho_1 X(t-\tau)\big)\,dt + \beta dW(t), & t \geqslant 0, \\ X(t) &= \Psi(t), & -\tau \leqslant t < 0. \end{aligned} \tag{4}$$

This is a linear autonomous equation with a constant lag and additive noise (and the delay is only in the drift term).

Example 2 (Population growth again). Assume now that in equation (4) we want to model noisy behaviour in the system itself; for example, the intrinsic variability of the

https://doi.org/10.1112/S1461157000000322 Published online by Cambridge University Press

cell proliferation, or other individual differences and interaction between individuals. This leads to a multiplicative noise term in equation (4).

$$\begin{array}{rcll} dX(t) &=& \big(\rho_0 X(t) + \rho_1 X(t-\tau)\big)\,dt + \beta X(t)\,dW(t), & t \geqslant 0, \\ X(t) &=& \Psi(t), & -\tau \leqslant t < 0. \end{array} \tag{5}$$

More examples. For additional examples we can refer to applications in neural control mechanisms: neurological diseases [**5**], human postural sway [**10**] and pupil light reflex [**18**].

3. *General formulation*

Let $(\Omega, \mathcal{A}, P)$ be a complete probability space with a filtration $(\mathcal{A}_t)$ satisfying the usual conditions; that is, the filtration $(\mathcal{A}_t)_{t\geqslant 0}$ is right-continuous, and each $\mathcal{A}_t$, where $t \geqslant 0$, contains all P-null sets in $\mathcal{A}$. For the general theory we refer to [**29**]. With $\mathcal{E}(X) = \int_\Omega X\,dP$ we say for $1 \leqslant p \leqslant \infty$ that $X \in \mathcal{L}^p = \mathcal{L}^p(\Omega, \mathcal{A}, P)$ if

$$\mathcal{E}\left(|X|^p\right) < \infty, \qquad \text{and we define} \qquad \|X\|_p = \left(\mathcal{E}\left(|X|^p\right)\right)^{\frac{1}{p}}.$$

Here, $\mathcal{E}$ denotes the expectation; for a useful summary of the properties of expectation and conditional expectation that will be used here, refer to the work of Mao [**19**, pp. 8–9].

In the literature of stochastic numerical analysis, convergence is usually considered either in the mean-square sense or in the absolute mean; that is, with $p = 2$ or $p = 1$, respectively, in the following definition.

Definition 1. Let $\{X_\nu\}_{\nu\geqslant 0}$ be a sequence of random variables defined on $\mathcal{L}^p(\Omega, \mathcal{A}, P)$. Then *convergence* as $\nu \to \nu_*$ of X_ν to a random variable X in $\mathcal{L}^p(\Omega, \mathcal{A}, P)$ in the pth mean takes place when

$$\mathcal{E}|X_\nu - X|^p \to 0 \quad \text{as} \quad \nu \to \nu_*, \tag{6}$$

or, equivalently, when

$$\|X_\nu - X\|_p \to 0 \quad \text{as} \quad \nu \to \nu_*. \tag{7}$$

Remark 1. Due to Jensen's inequality (which states that $g(E(Z)) \leqslant E(g(Z))$ for any convex function $g : \mathbb{R} \to \mathbb{R}$) we have

$$\mathcal{E}\left(|Z|^q\right) \leqslant \left(\mathcal{E}(|Z|^p)\right)^{q/p} \qquad \text{for all } 0 < q \leqslant p,\ Z \in \mathcal{L}^p(\Omega, \mathcal{A}, P), \tag{8}$$

so if $X \in \mathcal{L}^p(\Omega, \mathcal{A}, P)$ and $p > q$, then $X \in \mathcal{L}^q(\Omega, \mathcal{A}, P)$. Also, on taking $Z = X_\nu - X$, we see that if condition (7) holds for $p > q$, it is also valid with p replaced by q; thus, in particular, *convergence in the mean* ($q = 1$) *is implied by convergence in the mean-square* ($p = 2$).

Let $W(t)$ be a 1-dimensional Wiener process given on the filtered probability space $(\Omega, \mathcal{A}, P)$. We consider the scalar stochastic delay differential equation ($0 = t_0 < T < \infty$):

$$\left.\begin{array}{rcll} dX(t) &=& f\big(t, X(t), X(t-\tau)\big)\,dt + g\big(t, X(t), X(t-\tau)\big)\,dW(t), & t \in [0, T], \\ X(t) &=& \Psi(t), & t \in [-\tau, 0], \end{array}\right\} \tag{9}$$

with one fixed lag, where $\Psi(t)$ is an $\mathcal{A}_{t_0}$-measurable $C([-\tau, 0], \mathbb{R})$-valued random variable such that $\mathcal{E}\|\Psi\|^2 < \infty$. (By $C([-\tau, 0], \mathbb{R})$ we mean the Banach space of all continuous paths from $[-\tau, 0] \to \mathbb{R}$ equipped with the supremum norm $\|\eta\| := \sup_{s\in[-\tau,0]} |\eta(s)|$, where $\eta \in C$.) The first term on the right-hand side is called the *drift function*, characterizing the local trend, and the second term denotes the *diffusion function*, which influences the average size of the fluctuations of X.

If the functions f and g do not explicitly depend on t the equation is called *autonomous*, and we consider this case for simplicity. Equation (9) can then be formulated equivalently as

$$X(t) = X(0) + \int_0^t f\big(X(s), X(s-\tau)\big)\,ds + \int_0^t g\big(X(s), X(s-\tau)\big)\,dW(s), \tag{10}$$

for $t \in [0, T]$ and with $X(t) = \Psi(t)$, for $t \in [-\tau, 0]$. The second integral in equation (10) is a stochastic integral which is to be interpreted in the Itô sense.

3.1. *Assumptions on the functions f, g, and Ψ*

We have $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $g : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ and $\Psi : [-\tau, 0] \to \mathbb{R}$, and we shall, at various points, assume subsets of the following set of conditions.

A1. The functions f and g are continuous.

A2. (a) The functions f and g satisfy a uniform Lipschitz condition; that is, there exist positive constants L_1, L_2, L_3 and L_4 such that for all $\phi_1, \phi_2, \psi_1, \psi_2 \in \mathbb{R}$ and $t \in [0, T]$

$$|f(\phi_1, \psi_1) - f(\phi_2, \psi_2)| \leqslant L_1 |\phi_1 - \phi_2| + L_2 |\psi_1 - \psi_2|, \tag{11}$$

and

$$|g(\phi_1, \psi_1) - g(\phi_2, \psi_2)| \leqslant L_3|\phi_1 - \phi_2| + L_4|\psi_1 - \psi_2|. \tag{12}$$

(b) The function Ψ is Hölder-continuous with exponent γ; that is, there exists a positive constant L_5 such that for $t, s \in [-\tau, 0]$

$$\mathcal{E}\left(|\Psi(t) - \Psi(s)|^p\right) \leqslant L_5|t - s|^{p\gamma}, \qquad p = 1, 2. \tag{13}$$

A3. The functions f and g satisfy a linear growth condition; that is, there exist positive constants K_1 and K_2 such that for all $\phi, \phi_1, \psi, \psi_1 \in \mathbb{R}$ and $t \in [0, T]$,

$$|f(\phi, \phi_1)|^2 \leqslant K_1\left(1 + |\phi|^2 + |\phi_1|^2\right) \tag{14}$$

and

$$|g(\psi, \psi_1)|^2 \leqslant K_2\left(1 + |\psi|^2 + |\psi_1|^2\right). \tag{15}$$

A4. The partial derivatives of $f(\phi, \psi)$,

$$\frac{\partial f}{\partial \phi}, \quad \frac{\partial f}{\partial \psi}, \quad \frac{\partial^2 f}{\partial \phi^2}, \quad \frac{\partial^2 f}{\partial \psi^2}, \quad \text{and} \quad \frac{\partial^2 f}{\partial \phi\, \partial \psi},$$

exist and are uniformly bounded.

A5. (a) The function g does not depend on X.
(b) The function $f(\phi, \psi)$ is decomposable as $f_1(\phi) + f_2(\psi)$.

Concerning Assumption A2(b), if $\gamma > 1/2$, its value does not impinge upon the theory given in Section 5 unless Assumption A5 holds, in which case its value is irrelevant when $\gamma > 1$. Assumption A4 is the natural extension of an assumption made by Milstein [**21**, p. 20] in his discussion of Euler's method for SODEs. If assumption A5(a) is valid, the SDDE has *additive noise*; otherwise, the equation has *multiplicative noise*.

https://doi.org/10.1112/S1461157000000322 Published online by Cambridge University Press

Definition 2. An $\mathbb{R}$-valued stochastic process $X(t) : [-\tau, T] \times \Omega \to \mathbb{R}$ is called a *strong solution* of equation (9), if it is a measurable, sample-continuous process such that $X|[0, T]$ is $(\mathcal{A}_t)_{0 \leqslant t \leqslant T}$-adapted, f and g are continuous functions and X satisfies equation (9) or, equivalently equation (10), almost surely, and satisfies the initial condition $X(t) = \Psi(t)$, where $t \in [-\tau, 0]$. A solution $X(t)$ is said to be *path-wise unique* if any other solution $\widehat{X}(t)$ is indistinguishable from it; that is,

$$P\Big(X(t) = \widehat{X}(t) \text{ for all } -\tau \leqslant t \leqslant T\Big) = 1.$$

Theorem 1. *Assume that the functions f and g satisfy the assumptions* A1 *to* A3 *above. Then there exists a unique strong solution to equation* (9).

Proof. Proofs of Theorem 1 can be found in [**19**], [**22**] and [**23**]. □

We cite a theorem from Mao ([**19**, Lemma 5.5.2]), which we shall use in our analysis. It was originally stated and proved for more general equations.

Theorem 2. *Let inequalities* (14) *and* (15) *hold. Then the solution of equation* (9) *has the property*

$$\mathcal{E}\left(\sup_{-\tau \leqslant t \leqslant T} |X(t)|^2\right) \leqslant C_1, \tag{16}$$

with

$$C_1 := \left(\frac{1}{2} + 4\mathcal{E}\|\Psi\|^2\right) e^{6KT(T+4)}, \qquad K := \max(K_1, K_2). \tag{17}$$

Moreover, for any $0 \leqslant s < t \leqslant T$ *with* $t - s < 1$,

$$\mathcal{E}|X(t) - X(s)|^2 \quad \leqslant \quad C_2(t - s), \tag{18}$$

where $C_2 = 4K(1 + 2C_1)$.

It is well known in the theory of deterministic DDEs that a scalar DDE with a single fixed lag may be interpreted on each interval of length τ as a system of ODEs. Denote $\gamma_0(t) = t$, $\gamma_1(t) = t - \tau$, and $\gamma_i(t) = \gamma_1(\gamma_{i-1}(t))$, where $i \geqslant 2$; also, $X(t) = Y_m(t)$, where $t \in [m\tau, (m+1)\tau]$, $Y_{-1}(t) = \Psi(t)$ and $dW_m(t) = dW(\gamma_{m-r}(t))$. Then equation (9) becomes

$$\begin{aligned} dY_r(t) = {} & \gamma'_{m-r}(t) f\big(\gamma_{m-r}(t), Y_r(t), Y_{r-1}(t)\big)\, dt \\ & + \gamma'_{m-r}(t) g\big(\gamma_{m-r}(t), Y_r(t), Y_{r-1}(t)\big)\, dW_m(t), \\ & \text{for } t \in [m\tau, (m+1)\tau] \text{ and } r = 0, \ldots, m. \end{aligned} \tag{19}$$

With this approach, the problem of solving an SDDE is reduced to one of solving a sequence of systems of SODEs of increasing dimension on successive intervals $[m\tau, (m+1)\tau]$. If one wishes to solve the SDDE on an unbounded interval, the dimensionality of the system of SODEs, obtained by the above procedure, is also unbounded. This approach has been followed in [**17**].

4. *Numerical analysis for an autonomous SDDE*

For simplicity we shall in the sequel consider equation (9) in the autonomous form; that is, we shall work with

$$\left.\begin{array}{rcll} dX(t) & = & f\big(X(t), X(t-\tau)\big)dt + g\big(X(t), X(t-\tau)\big)dW(t), & t \in [0, T], \\ X(t) & = & \Psi(t), & t \in [-\tau, 0]. \end{array}\right\} \tag{20}$$

We define a mesh with a uniform step on the interval $[0, T]$, $h = T/N$, $t_n = n \cdot h$, where $n = 0, \dots N$, and where we assume that for the given h there is a corresponding integer N_τ such that the lag can be expressed in terms of the stepsize as $\tau = N_\tau \cdot h$.

We consider strong approximations $\widetilde{X}_n$ of the solution to equation (20), using a stochastic explicit one-step method of the form

$$\widetilde{X}_{n+1} = \widetilde{X}_n + \phi\left(h, \widetilde{X}_n, \widetilde{X}_{n-N_\tau}, I_\phi\right), \qquad n = 0, \dots, N-1, \tag{21}$$

where the initial values are given by $\widetilde{X}_{n-N_\tau} := \Psi(t_n - \tau)$ for $n - N_\tau \leqslant 0$. The increment function $\phi(h, \cdot, \cdot, I_\phi) : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ incorporates a finite number of multiple Itô-integrals (see [**15**] or [**21**]) of the form

$$I_{(j_1,\dots,j_l),h} = \int_t^{t+h} \int_t^{s_l} \dots \int_t^{s_2} dW^{j_1}(s_1) \dots dW^{j_{l-1}}(s_{l-1}) dW^{j_l}(s_l),$$

where $j_i \in \{0, 1\}$ and $dW^0(t) = dt$, and with $t = t_n$ in the case (21). We denote by I_ϕ the collection of Itô-integrals required to compute the increment function ϕ. The increment function ϕ is assumed to generate approximations $\widetilde{X}_n$ which are $\mathcal{A}_{t_n}$-measurable.

4.1. *Assumptions on the increment function ϕ*

We suppose that there exist positive constants C_1, C_2, C_3 such that for all $\xi, \xi', \eta, \eta' \in \mathbb{R}$

$$\left|\mathcal{E}\left(\phi(h, \xi, \eta, I_\phi) - \phi(h, \xi', \eta', I_\phi)\right)\right| \leqslant C_1 h \left(|\xi - \xi'| + |\eta - \eta'|\right), \tag{22}$$

$$\mathcal{E}\left(|\phi(h, \xi, \eta, I_\phi) - \phi(h, \xi', \eta', I_\phi)|^2\right) \leqslant C_2 h \left(|\xi - \xi'|^2 + |\eta - \eta'|^2\right), \tag{23}$$

and

$$\mathcal{E}\left(|\phi(h, \xi, \eta, I_\phi)|^2\right) \leqslant C_3 h \left(1 + |\xi|^2 + |\eta|^2\right). \tag{24}$$

Lemma 1. *If the increment function ϕ in equation* (21) *satisfies condition* (24), *then $\mathcal{E}|\widetilde{X}_n|^2 < \infty$ for all $n \leqslant N$.*

Proof. We have

$$\begin{aligned} \mathcal{E}\left(\left|\widetilde{X}_n\right|^2 |\mathcal{A}_{t_0}\right) &= \mathcal{E}\left(\left|\widetilde{X}_{n-1} + \phi(h, \widetilde{X}_{n-1}, \widetilde{X}_{n-1-N_\tau}, I_\phi)\right|^2 |\mathcal{A}_{t_0}\right) \\ &\leqslant 2\mathcal{E}\left(\left|\widetilde{X}_{n-1}\right|^2 |\mathcal{A}_{t_0}\right) + 2\mathcal{E}\left(\left|\phi(h, \widetilde{X}_{n-1}, \widetilde{X}_{n-1-N_\tau}, I_\phi)\right|^2 |\mathcal{A}_{t_0}\right) \\ &\leqslant 2\mathcal{E}\left(\left|\widetilde{X}_{n-1}\right|^2 |\mathcal{A}_{t_0}\right) + 2C_3 h \mathcal{E}\left(1 + \left|\widetilde{X}_{n-1}\right|^2 + \left|\widetilde{X}_{n-1-N_\tau}\right|^2 |\mathcal{A}_{t_0}\right) \\ &= 2(1 + C_3 h)\,\mathcal{E}\left(\left|\widetilde{X}_{n-1}\right|^2 |\mathcal{A}_{t_0}\right) + 2C_3 h \mathcal{E}\left(\left|\widetilde{X}_{n-1-N_\tau}\right|^2 |\mathcal{A}_{t_0}\right) + 2C_3 h. \end{aligned}$$

The lemma follows from this result. To display the argument in detail, we define

$$\begin{aligned} \rho_0 &= \max_{-N_\tau\leqslant r\leqslant 0} \mathcal{E}\left(\left|\widetilde{X}_r\right|^2 \;|\mathcal{A}_{t_0}\right), & \rho_n &= \max_{0<r\leqslant n} \mathcal{E}\left(\left|\widetilde{X}_r\right|^2 \;|\mathcal{A}_{t_0}\right), \\ \widehat{\rho}_0 &= \rho_0, & \widehat{\rho}_n &= \max_{-N_\tau\leqslant r\leqslant n} \mathcal{E}\left(\left|\widetilde{X}_r\right|^2 \;|\mathcal{A}_{t_0}\right) = \max(\rho_0, \rho_n). \end{aligned}$$

Note that the sequences $\{\rho_n\}_{n\geqslant 1}$ and therefore $\{\widehat{\rho}_n\}_{n\geqslant 0}$ are monotonically non-decreasing. Thus, we obtain

$$\widehat{\rho}_n \leqslant \begin{cases} 2(1+C_3h)\,\widehat{\rho}_{n-1} + 2C_3h\widehat{\rho}_0 + 2C_3h, & \text{for } 0 < n \leqslant N_\tau, \\ 2(1+2C_3h)\,\widehat{\rho}_{n-1} + 2C_3h, & \text{for } n > N_\tau; \end{cases}$$

whence

$$\widehat{\rho}_n \leqslant 2(1+C_3h)\,\widehat{\rho}_{n-1} + 2C_3h\widehat{\rho}_0 + 2C_3h, \qquad \text{for } n > 0.$$

By induction, when $\widehat{\rho}_n \leqslant \alpha\widehat{\rho}_{n-1} + \zeta$ for $n > 0$, we find that

$$\widehat{\rho}_n \leqslant \alpha^n\zeta + (1+\alpha+\cdots+\alpha^n)\widehat{\rho}_0;$$

setting $\alpha = 2(1+C_3h)$ and $\zeta = 2\,C_3h\widehat{\rho}_0 + 2C_3h$, and using the assumptions on the initial function Ψ to bound $\widehat{\rho}_0$, we deduce the desired result. □

Notation 1. We denote by $X(t_{n+1})$ the value of the exact solution of equation (20) at the meshpoint t_{n+1}, by $\widetilde{X}_{n+1}$ the value of the approximate solution using equation (21), and by $\widetilde{X}(t_{n+1})$ the value obtained after just one step of equation (21); that is,

$$\widetilde{X}(t_{n+1}) = X(t_n) + \phi\left(h, X(t_n), X(t_n-\tau), I_\phi\right).$$

Using the above notation we can give the following definitions, employing terminology used for SODEs by Artemiev and Averina [**2**, pp. 89–91].

Definition 3. The *error* of the above approximation $\{\widetilde{X}_n\}$ on the mesh-points is the sequence of random variables

$$\epsilon_n := X(t_n) - \widetilde{X}_n, \quad n = 1, \ldots, N. \tag{25}$$

Note that ϵ_n is $\mathcal{A}_{t_n}$-measurable since both $X(t_n)$ and $\widetilde{X}_n$ are $\mathcal{A}_{t_n}$-measurable random variables, and that $\left(\mathcal{E}|\epsilon_n|^2\right)^{1/2}$ is the $\mathcal{L}^2$-norm of (25).

Definition 4. Let

$$\delta_{n+1} = X(t_{n+1}) - \widetilde{X}(t_{n+1}), \qquad n = 0, \ldots, N-1. \tag{26}$$

The method (21) is said to be *consistent* with order p_1 in the mean and with order p_2 in the mean-square sense if, with

$$p_2 \geqslant \frac{1}{2} \quad \text{and} \quad p_1 \geqslant p_2 + \frac{1}{2}, \tag{27}$$

the estimates

$$\max_{0\leqslant n\leqslant N-1} |\mathcal{E}(\delta_{n+1})| \leqslant Ch^{p_1} \quad \text{as } h\to 0, \tag{28}$$

and

$$\max_{0\leqslant n\leqslant N-1} \left(\mathcal{E}|\delta_{n+1}|^2\right)^{\frac{1}{2}} \leqslant Ch^{p_2} \quad \text{as } h\to 0, \tag{29}$$

hold, where the (generic) constant C does not depend on h, but may depend on T, and on the initial data.

https://doi.org/10.1112/S1461157000000322 Published online by Cambridge University Press

We now state the main theorem of this paper, which is the analogue in the case of delay equations of a theorem by Milstein for SODEs (see [**21**], in particular for a discussion of the necessity to employ consistency in the mean *and* in the mean-square, as well as the application of conditional versions of the inequalities (28) and (29)).

Theorem 3. *We assume that the conditions of Theorem 1 are fulfilled. Suppose that the method defined by equation (21) is consistent with order p_1 in the mean and order p_2 in the mean-square sense, with $p1$, p_2 satisfying inequality (27), and that the increment function ϕ in equation (21) satisfies the estimates (22) and (23). Then the approximation (21) for equation (20) is convergent in $\mathcal{L}^2$ (as $h \to 0$ with $\tau/h \in \mathbb{N}$) with order $p = p_2 - 1/2$. That is, convergence is in the mean-square sense, and*

$$\max_{1\leqslant n\leqslant N} \left(\mathcal{E}|\epsilon_n|^2\right)^{\frac{1}{2}} \leqslant Ch^p \quad \text{as } h\to 0. \tag{30}$$

Proof. Using Notation 1, adding and subtracting $X(t_n)$ and $\phi(h, X(t_n), X(t_n-\tau), I_\phi)$, and rearranging, we obtain

$$\begin{aligned}
\epsilon_{n+1} &= X(t_{n+1}) - \widetilde{X}_{n+1} \\
&= \underbrace{X(t_n) - \widetilde{X}_n}_{\epsilon_n} + \underbrace{X(t_{n+1}) - X(t_n) - \phi\left(h, X(t_n), X(t_n-\tau), I_\phi\right)}_{\delta_{n+1}} \\
&\quad + \underbrace{\phi\left(h, X(t_n), X(t_n-\tau), I_\phi\right) - \phi\left(h, \widetilde{X}_n, \widetilde{X}_{n-N_\tau}, I_\phi\right)}_{u_n} \\
&= \epsilon_n + \delta_{n+1} + u_n,
\end{aligned}$$

where u_n is defined as

$$u_n := \phi\left(h, X(t_n), X(t_n-\tau), I_\phi\right) - \phi\left(h, \widetilde{X}_n, \widetilde{X}_{n-N_\tau}, I_\phi\right). \tag{31}$$

Thus, squaring, employing the conditional mean with respect to the σ-algebra $\mathcal{A}_{t_0}$, and taking the modulus, we obtain

$$\begin{aligned}
\mathcal{E}\left(|\epsilon_{n+1}|^2 \ |\mathcal{A}_{t_0}\right) \leqslant \mathcal{E}\left(|\epsilon_n|^2 \ |\mathcal{A}_{t_0}\right) &+ \underbrace{\mathcal{E}\left(|\delta_{n+1}|^2 \ |\mathcal{A}_{t_0}\right)}_{(1)} + \underbrace{\mathcal{E}\left(|u_n|^2 \ |\mathcal{A}_{t_0}\right)}_{(2)} \\
&+ \underbrace{2\left|\mathcal{E}\left(\delta_{n+1} \cdot \epsilon_n \mid \mathcal{A}_{t_0}\right)\right|}_{(3)} + \underbrace{2\left|\mathcal{E}\left(\delta_{n+1} \cdot u_n \ |\mathcal{A}_{t_0}\right)\right|}_{(4)} \\
&+ \underbrace{2\left|\mathcal{E}\left(\epsilon_n \cdot u_n \ |\mathcal{A}_{t_0}\right)\right|}_{(5)},
\end{aligned} \tag{32}$$

which holds almost surely.

We shall now estimate the separate terms in inequality (32) individually and in sequence; all the estimates hold almost surely. We shall frequently use the Hölder inequality, the inequality $2ab \leqslant a^2 + b^2$ and properties of conditional expectation, which can be found in [**29**]. In the sequel we shall use c to denote an unspecified constant, which depends only on the constants L_1, L_2, L_3, L_4, K_1, K_2, C_1 and C_2, and on T and the initial data.

• For the term labelled (1) in inequality (32), we have, due to the assumed consistency in the mean-square sense of the method,

$$\mathcal{E}\left(|\delta_{n+1}|^2 \ |\mathcal{A}_{t_0}\right) = \mathcal{E}\left(\mathcal{E}\left(|\delta_{n+1}|^2 \ |\mathcal{A}_{t_n}\right) |\mathcal{A}_{t_0}\right) \leqslant ch^{2p_2}.$$

• For the term labelled (2) in inequality (32), we have, due to property (23) of the increment function,

$$\mathcal{E}\left(|u_n|^2 \ |\mathcal{A}_{t_0}\right) \leqslant ch\mathcal{E}\left(|\epsilon_n|^2 \ |\mathcal{A}_{t_0}\right) + ch\mathcal{E}\left(|\epsilon_{n-N_\tau}|^2 \ |\mathcal{A}_{t_0}\right).$$

• For the term labelled (3) we have, due to the consistency condition,

$$\begin{aligned}
2\left|\mathcal{E}\left(\delta_{n+1} \cdot \epsilon_n \ |\mathcal{A}_{t_0}\right)\right| &\leqslant 2\left|\mathcal{E}\left(\mathcal{E}(\delta_{n+1}|\mathcal{A}_{t_n})\epsilon_n \ |\mathcal{A}_{t_0}\right)\right| \\
&\leqslant 2\left(\mathcal{E}\left|\mathcal{E}(\delta_{n+1} \ |\mathcal{A}_{t_n})\right|^2\right)^{\frac{1}{2}} \cdot \left(\mathcal{E}\left(|\epsilon_n|^2 \ |\mathcal{A}_{t_0}\right)\right)^{\frac{1}{2}} \\
&\leqslant 2\left(\mathcal{E}(ch^{p_1})^2\right)^{\frac{1}{2}} \cdot \left(\mathcal{E}\left(|\epsilon_n|^2 \ |\mathcal{A}_{t_0}\right)\right)^{\frac{1}{2}} \\
&= 2\left(\mathcal{E}\left(ch^{2p_1-1}\right)\right)^{\frac{1}{2}} \cdot \left(h\mathcal{E}\left(|\epsilon_n|^2 \ |\mathcal{A}_{t_0}\right)\right)^{\frac{1}{2}} \\
&\leqslant ch^{2p_1-1} + h\mathcal{E}\left(|\epsilon_n|^2 \ |\mathcal{A}_{t_0}\right).
\end{aligned}$$

• For the term labelled (4) in inequality (32), we obtain, by employing the consistency condition and property (23) of the increment function ϕ,

$$\begin{aligned}
2\left|\mathcal{E}(\delta_{n+1} \cdot u_n \ |\mathcal{A}_{t_0})\right| &\leqslant 2\left(\mathcal{E}\left(|\delta_{n+1}|^2 \ |\mathcal{A}_{t_0}\right)\right)^{\frac{1}{2}}\left(\mathcal{E}\left(|u_n|^2 \ |\mathcal{A}_{t_0}\right)\right)^{\frac{1}{2}} \\
&\leqslant \mathcal{E}\left(\mathcal{E}\left(|\delta_{n+1}|^2 \ |\mathcal{A}_{t_n}\right)|\mathcal{A}_{t_0}\right) + \mathcal{E}\left(|u_n|^2 \ |\mathcal{A}_{t_0}\right) \\
&\leqslant ch^{2p_2} + ch\mathcal{E}\left(\epsilon_n^2 \ |\mathcal{A}_{t_0}\right) + ch\mathcal{E}\left(\epsilon_{n-N_\tau}^2 \ |\mathcal{A}_{t_0}\right).
\end{aligned}$$

• For the term labelled (5) in inequality (32) we have, using definition (31) and property (22) of the increment function ϕ,

$$\begin{aligned}
2\left|\mathcal{E}\left(u_n \cdot \epsilon_n \ |\mathcal{A}_{t_0}\right)\right| &\leqslant 2\mathcal{E}\left(\left|\mathcal{E}(u_n \ |\mathcal{A}_{t_n})\right| \cdot |\epsilon_n| \ |\mathcal{A}_{t_0}\right) \\
&\leqslant ch\mathcal{E}\left(|\epsilon_n|^2 \ |\mathcal{A}_{t_0}\right) + 2ch\mathcal{E}\left(|\epsilon_n| \ |\epsilon_{n-N_\tau}| \ |\mathcal{A}_{t_0}\right) \\
&\leqslant ch\mathcal{E}\left(|\epsilon_n|^2 \ |\mathcal{A}_{t_0}\right) + ch\left\{2\left(\mathcal{E}(|\epsilon_n|^2 \ |\mathcal{A}_{t_0})\right)^{\frac{1}{2}} \cdot \left(\mathcal{E}(|\epsilon_{n-N_\tau}|^2 \ |\mathcal{A}_{t_0})\right)^{\frac{1}{2}}\right\} \\
&\leqslant ch\mathcal{E}\left(|\epsilon_n|^2 \ |\mathcal{A}_{t_0}\right) + ch\mathcal{E}\left(|\epsilon_n|^2 \ |\mathcal{A}_{t_0}\right) + ch\mathcal{E}\left(|\epsilon_{n-N_\tau}|^2 \ |\mathcal{A}_{t_0}\right), \\
&\leqslant ch\mathcal{E}\left(|\epsilon_n|^2 \ |\mathcal{A}_{t_0}\right) + ch\mathcal{E}\left(|\epsilon_{n-N_\tau}|^2 \ |\mathcal{A}_{t_0}\right).
\end{aligned}$$

Combining these results, we obtain, with $2p_2 \leqslant 2p_1 - 1$,

$$\mathcal{E}\left(\epsilon_{n+1}^2 \ |\mathcal{A}_{t_0}\right) \leqslant (1+ch)\mathcal{E}\left(\epsilon_n^2 \ |\mathcal{A}_{t_0}\right) + ch^{2p_2} + ch\mathcal{E}\left(|\epsilon_{n-N_\tau}|^2 \ |\mathcal{A}_{t_0}\right).$$

Now we shall prove the assertion by an induction argument over consecutive intervals of length τ up to the end of the interval $[0, T]$. Since we have exact initial values, we set $\epsilon_n = 0$ for $n = -N_\tau, \ldots, 0$.

Step 1. Suppose that $t_n \in [0, \tau]$; that is, $n = 1, \ldots, N_\tau$ and $\epsilon_{n-N_\tau} = 0$.

$$\begin{aligned}
\mathcal{E}\left(\epsilon_{n+1}^2 \ |\mathcal{A}_{t_0}\right) &\leqslant (1+ch)\mathcal{E}\left(\epsilon_n^2 \ |\mathcal{A}_{t_0}\right) + ch^{2p_2} \\
&\leqslant ch^{2p_2}\sum_{k=0}^{n}(1+ch)^k \\
&= ch^{2p_2}\frac{(1+ch)^{n+1}-1}{(1+ch)-1}
\end{aligned}$$

https://doi.org/10.1112/S1461157000000322 Published online by Cambridge University Press

$$\leqslant ch^{2p_2-1}\left(\left(e^{ch}\right)^{n+1}-1\right)$$
$$\leqslant ch^{2p_2-1}\left(e^{cT}-1\right).$$

Step 2. Suppose that $t_n \in [k\tau, (k+1)\tau]$, and make the assumption that

$$\mathcal{E}\left(|\epsilon_{n-N_\tau}|^2 \,|\mathcal{A}_{t_0}\right) \leqslant ch^{2p_2-1}.$$

Then

$$\begin{aligned}\mathcal{E}\left(\epsilon_{n+1}^2 \,|\mathcal{A}_{t_0}\right) &\leqslant (1+ch)\mathcal{E}\left(\epsilon_n^2 \,|\mathcal{A}_{t_0}\right) + ch^{2p_2} + ch\mathcal{E}\left(|\epsilon_{n-N_\tau}|^2 \,|\mathcal{A}_{t_0}\right)\\ &\leqslant (1+ch)\mathcal{E}\left(\epsilon_n^2 \,|\mathcal{A}_{t_0}\right) + ch^{2p_2} + hch^{2p_2-1}\\ &= (1+ch)\mathcal{E}\left(\epsilon_n^2 \,|\mathcal{A}_{t_0}\right) + ch^{2p_2}\\ &\leqslant ch^{2p_2-1}(e^{cT}-1),\end{aligned}$$

by the same arguments as above. This implies, almost surely, that

$$\left(\mathcal{E}\left(\epsilon_{n+1}^2 \,|\mathcal{A}_{t_0}\right)\right)^{\frac{1}{2}} \leqslant ch^{p_2-\frac{1}{2}},$$

which proves the theorem. □

The above theorem is an analogue of [**21**, Theorem 1.1], but our proof follows different lines. In the remainder of this section we shall discuss stochastic zero-stability. We adapt the definition given in [**12**].

Definition 5. The stochastic one-step method (21) is *zero-stable in the quadratic mean-square sense* if, given $\varepsilon > 0$, there is a $\delta = \delta(\varepsilon, h_0) > 0$ such that for all $0 < h < h_0$ and positive integers $n \leqslant T/h$,

$$\rho_0 \equiv \max_{-N_t \leqslant r \leqslant 0} \mathcal{E}\left|\widetilde{X}_r - \widetilde{X}_r^*\right|^2 \leqslant \delta \quad \Longrightarrow \quad \rho_n \equiv \mathcal{E}\left|\widetilde{X}_n - \widetilde{X}_n^*\right|^2 \leqslant \varepsilon$$

holds, where $\widetilde{X}_n^*$ denotes the sequence defined by the method (21) with the initial values $\widetilde{X}_r$ for $r = -N_t, \ldots, 0$ replaced by $\widetilde{X}_r^*$ for $r = -N_t, \ldots, 0$. If the method is stable and, further, if $\rho_n \to 0$ whenever ρ_0 is sufficiently small, the method is *asymptotically zero-stable in the quadratic mean-square sense*.

Theorem 4. *If the increment function ϕ of the approximation method* (21) *satisfies the estimates* (22) *and* (23), *then the one-step method* (21) *is zero-stable in the quadratic mean-square sense.*

Proof. We have, for $0 < n \leqslant N = T/h$,

$$\begin{aligned}\left(\widetilde{X}_n - \widetilde{X}_n^*\right)^2 \leqslant{}& \left(\widetilde{X}_{n-1} - \widetilde{X}_{n-1}^*\right)^2\\ &+2\left(\widetilde{X}_{n-1} - \widetilde{X}_{n-1}^*\right)\cdot\\ &\qquad\left(\phi\left(h, \widetilde{X}_{n-1}, \widetilde{X}_{n-1-N_\tau}, I_\phi\right) - \phi\left(h, \widetilde{X}_{n-1}^*, \widetilde{X}_{n-1-N_\tau}^*, I_\phi\right)\right)\\ &+\left(\phi\left(h, \widetilde{X}_{n-1}, \widetilde{X}_{n-1-N_\tau}, I_\phi\right) - \phi\left(h, \widetilde{X}_{n-1}^*, \widetilde{X}_{n-1-N_\tau}^*, I_\phi\right)\right)^2.\end{aligned}$$

Now we take expected values conditioned on the σ-algebra $(\mathcal{A}_{t_0})$, take the modulus and use properties of conditional expectation and the estimates (22) and (23), and proceed with the manipulations in the same way as we did for the terms labelled (2) and (5) in inequality (32). We obtain

$$
\begin{aligned}
&\mathcal{E}\left(\left|\widetilde{X}_n - \widetilde{X}_n^*\right|^2 \,|\mathcal{A}_{t_0}\right) \\
\leqslant\ & \mathcal{E}\left(\left|\widetilde{X}_{n-1} - \widetilde{X}_{n-1}^*\right|^2 \,|\mathcal{A}_{t_0}\right) \\
& +2\Big|\mathcal{E}\Big(\,(\widetilde{X}_{n-1} - \widetilde{X}_{n-1}^*)\cdot \\
&\qquad \Big(\phi\left(h, \widetilde{X}_{n-1}, \widetilde{X}_{n-1-N_\tau}, I_\phi\right) - \phi\big(h, \widetilde{X}_{n-1}^*, \widetilde{X}_{n-1-N_\tau}^*, I_\phi\big)\Big) \,|\mathcal{A}_{t_0}\Big)\Big| \\
& +\mathcal{E}\left(\left|\phi(h, \widetilde{X}_{n-1}, \widetilde{X}_{n-1-N_\tau}, I_\phi) - \phi(h, \widetilde{X}_{n-1}^*, \widetilde{X}_{n-1-N_\tau}^*, I_\phi)\right|^2 |\mathcal{A}_{t_0}\right) \\
\leqslant\ & \mathcal{E}\left(\left|\widetilde{X}_{n-1} - \widetilde{X}_{n-1}^*\right|^2 |\mathcal{A}_{t_0}\right) \\
& +ch\mathcal{E}\left(\left|\widetilde{X}_{n-1} - \widetilde{X}_{n-1}^*\right|^2 |\mathcal{A}_{t_0}\right) + ch\mathcal{E}\left(\left|\widetilde{X}_{n-1-N_\tau} - \widetilde{X}_{n-1-N_\tau}^*\right|^2 |\mathcal{A}_{t_0}\right) \\
& +h\mathcal{E}\left(\left|\widetilde{X}_{n-1} - \widetilde{X}_{n-1}^*\right|^2 |\mathcal{A}_{t_0}\right) + ch\mathcal{E}\left(\left|\widetilde{X}_{n-1-N_\tau} - \widetilde{X}_{n-1-N_\tau}^*\right|^2 \,|\mathcal{A}_{t_0}\right) \\
=\ & (1+ch)\mathcal{E}\left(\left|\widetilde{X}_{n-1} - \widetilde{X}_{n-1}^*\right|^2 |\mathcal{A}_{t_0}\right) + ch\mathcal{E}\left(\left|\widetilde{X}_{n-1-N_\tau} - \widetilde{X}_{n-1-N_\tau}^*\right|^2 \,|\mathcal{A}_{t_0}\right),
\end{aligned}
$$

where c denotes a generic positive constant.

The proof now follows similar lines to the proof of Lemma 1. We define the quantities

$$
R_0 = \max_{-N_\tau \leqslant r \leqslant 0} \mathcal{E}\left(\left|\widetilde{X}_r - \widetilde{X}_r^*\right|^2 \,|\mathcal{A}_{t_0}\right) \qquad \text{and} \qquad R_n = \max_{0 < r \leqslant n} \mathcal{E}\left(\left|\widetilde{X}_r - \widetilde{X}_r^*\right|^2 \,|\mathcal{A}_{t_0}\right);
$$

$$
\widehat{R}_0 = R_0, \qquad \widehat{R}_n = \max_{-N_\tau \leqslant r \leqslant n} \mathcal{E}\left(\left|\widetilde{X}_r - \widetilde{X}_r^*\right|^2 \,|\mathcal{A}_{t_0}\right) = \max\left(R_0, R_n\right),
$$

and we note that $\{R_n\}_{n>0}$ and $\{\widehat{R}_n\}_{n\geqslant 0}$ are monotonically non-decreasing. We now obtain, for $0 < n \leqslant N_\tau$, the result $\widehat{R}_n \leqslant (1+ch)\widehat{R}_{n-1} + ch\widehat{R}_0$, whilst for $n > N_\tau$ we have $\widehat{R}_n \leqslant (1+ch)\widehat{R}_{n-1} + ch\widehat{R}_{j(n)}$ for some $j(n) < n$. Thus

$$
\widehat{R}_n \leqslant (1+2ch)\widehat{R}_{n-1} \text{ for } n > 0.
$$

It follows (by induction, and using the property that $1 + 2ch < \exp(2ch)$) that

$$
\widehat{R}_n \leqslant \exp(2cT)\widehat{R}_0.
$$

We deduce that, given $\varepsilon > 0$, we have

$$
\widehat{R}_n \leqslant \varepsilon \quad \text{if} \quad R_0 \leqslant \delta \equiv \varepsilon \exp(-2cT), \quad \text{when } n \leqslant N,
$$

which proves the theorem. □

Conjecture. We conjecture that if the method (21) is consistent in both the mean and the mean-square senses, and is asymptotically zero-stable in the mean-square, then the method is convergent in the mean-square sense. We have not located the corresponding discussion for SODEs in the literature.

5. *The Euler–Maruyama scheme*

The most widely used approximation method for stochastic differential equations is the Euler–Maruyama scheme, which we shall use to provide some numerical illustrations. In this section we shall prove that it satisfies the consistency conditions (28) and (29), as well as conditions (22) and (23).

https://doi.org/10.1112/S1461157000000322 Published online by Cambridge University Press

Recall that we consider strong approximations with a fixed step-size on the interval $[0, T]$; that is, $h = T/N$, $t_n = n \cdot h$, $n = 0, \ldots, N$, and that we assume the existence of an integer $N_\tau = N/(m \cdot r)$, such that the lag can be expressed in terms of the step-size as $\tau = N_\tau \cdot h$.

The Euler–Maruyama method has the following form for equation (20):

$$\begin{aligned} \widetilde{X}_{n-N_\tau} &= \Psi(t_n - \tau), && n - N_\tau \leqslant 0 \\ \widetilde{X}_{n+1} &= \widetilde{X}_n + hf\left(\widetilde{X}_n, \widetilde{X}_{n-N_\tau}\right) + g\left(\widetilde{X}_n, \widetilde{X}_{n-N_\tau}\right)\Delta W_{n+1}, && 1 \leqslant n \leqslant N-1 \end{aligned} \tag{33}$$

with $\Delta W_{n+1} := W_{(n+1)h} - W_{nh}$, denoting independent $N(0, h)$-distributed Gaussian random variables. We denote the increment function of the Euler–Maruyama scheme (33) by ϕ_{EM}. It contains only the most basic multiple Itô-integrals, namely $I_{(0),h} = h$ and $I_{(1),h} = \Delta W_{n+1}$.

Theorem 5. (1) *If the functions f, g and Ψ in equation* (20) *satisfy the conditions of Theorem* 1 (*that is, assumptions* A1 *to* A3 *and, in addition, assumption* A4), *then the Euler–Maruyama approximation is consistent* (*a*) *with order $p_1 = \min(1 + \gamma, 3/2)$ in the mean, and* (*b*) *with order $p_2 = \min(1/2 + \gamma, 1)$ in the mean-square, where γ is the exponent of Hölder-continuity of Ψ in assumption* A2.

(2) *For equations* (20) *with additive noise and a decomposable drift function f* (*that is, assumption* A5 *holds*), *the Euler–Maruyama approximation is consistent with order $p_1 = \min(1 + \gamma, 2)$ in the mean, and with order $p_2 = \min(1/2 + \gamma, 3/2)$ in the mean-square.*

Proof. We concentrate first on part (1). We shall frequently make use of the fact (see [**1**, Remark 6.1.7]) that for all $0 \leqslant u \leqslant t \leqslant T$ the equation

$$X(t) - X(0) = \int_0^t f\big(X(s), X(s-\tau)\big)\, ds + \int_0^t g\big(X(s), X(s-\tau)\big)\, dW(s)$$

is equivalent to

$$X(t) - X(u) = \int_u^t f\big(X(s), X(s-\tau)\big)\, ds + \int_u^t g\big(X(s), X(s-\tau)\big)\, dW(s). \tag{34}$$

First we prove consistency in the mean with order $p_1 = \min(1 + \gamma, 3/2)$. We thank Dr. Tretyakov for pointing out, in a private communication, a gap in an earlier version of the proof. We have

$$\begin{aligned} \delta_{n+1} &= X(t_{n+1}) - X(t_n) - \phi_{EM}\big(h, X(t_n), X(t_n - \tau), I_{\phi_{EM}}\big) \\ &= \int_{t_n}^{t_{n+1}} f\big(X(s), X(s-\tau)\big)\, ds + \int_{t_n}^{t_{n+1}} g\big(X(s), X(s-\tau)\big)\, dW(s) \\ &\quad -hf\big(X(t_n), X(t_n-\tau)\big) - g\big(X(t_n), X(t_n-\tau)\big)\Delta W_{n+1} \\ &= \int_{t_n}^{t_{n+1}} f\big(X(s), X(s-\tau)\big) - f\big(X(t_n), X(t_n-\tau)\big)\, ds \\ &\quad + \int_{t_n}^{t_{n+1}} g\big(X(s), X(s-\tau)\big) - g\big(X(t_n), X(t_n-\tau)\big)\, dW(s); \end{aligned}$$

https://doi.org/10.1112/S1461157000000322 Published online by Cambridge University Press

$$\text{hence } |\mathcal{E}(\delta_{n+1})| = \left|\mathcal{E}\int_{t_n}^{t_{n+1}} f\big(X(s), X(s-\tau)\big) - f\big(X(t_n), X(t_n-\tau)\big)\, ds\right|$$
$$\leqslant \underbrace{\left|\mathcal{E}\int_{t_n}^{t_{n+1}} \left\{\frac{\partial f}{\partial x_1}\big(X(t_n), X(t_n-\tau)\big)\,\big(X(s)-X(t_n)\big)\right\} ds\right|}_{(1)}$$
$$+ \underbrace{\left|\mathcal{E}\int_{t_n}^{t_{n+1}} \left\{\frac{\partial f}{\partial x_2}\big(X(t_n), X(t_n-\tau)\big)\,\big(X(s-\tau)-X(t_n-\tau)\big)\right\} ds\right|}_{(2)}$$
$$+ \underbrace{\left|\mathcal{E}\int_{t_n}^{t_{n+1}} \rho(s)\, ds\right|}_{(3)}, \tag{35}$$

using Taylor's theorem for f and denoting by $\partial f/\partial x_i$ the derivative of f with respect to the ith argument.

We have two cases to consider for the integrands in equation (35): (i) $s-\tau \leqslant 0$ for $s \in [t_n, t_{n+1}]$ (so we have $X(s-\tau) = \Psi(s-\tau)$), and (ii) $t_n - \tau > 0$.

- For the term (1) in equation (35) we obtain in both cases, by invoking equation (34) and assumption A4 on f,

$$\left|\mathcal{E}\int_{t_n}^{t_{n+1}} \left\{\frac{\partial f}{\partial x_1}\big(X(t_n), X(t_n-\tau)\big)\,\big(X(s)-X(t_n)\big)\right\} ds\right|$$
$$\leqslant c\left|\mathcal{E}\int_{t_n}^{t_{n+1}}\int_{t_n}^{s} f\big(X(u), X(u-\tau)\big)\, du\, ds\right|$$
$$\leqslant c\mathcal{E}\left(\sqrt{K(1+2\sup_{-\tau\leqslant r\leqslant T}|X(r)|^2)}\right)\cdot\left(\int_{t_n}^{t_{n+1}}\int_{t_n}^{s} du\, ds\right)$$
$$\leqslant c\left(\sqrt{K(1+2\mathcal{E}\sup_{-\tau\leqslant r\leqslant T}|X(r)|^2)}\right) h^2$$
$$\leqslant c\left(\sqrt{K(1+2\,C_1(\|\Psi\|, T)}\right) h^2$$

with $C_1(\|\Psi\|, T) = (1/2 + 4\mathcal{E}\|\Psi\|^2)\, e^{6K(T+4)T}$, due to inequality (16).

- For the term (2) in equation (35) assumption A4 on f and assumption A2 on Ψ yield in case (i):

$$\left|\mathcal{E}\int_{t_n}^{t_{n+1}} \left\{\frac{\partial f}{\partial x_2}\big(X(t_n), \Psi(t_n-\tau)\big)\,\big(\Psi(s-\tau)-\Psi(t_n-\tau)\big)\right\} ds\right|$$
$$\leqslant \mathcal{E}\int_{t_n}^{t_{n+1}} \left|\frac{\partial f}{\partial x_2}\big(X(t_n), \Psi(t_n-\tau)\big)\right|\;|\Psi(s-\tau)-\Psi(t_n-\tau)|\, ds$$
$$\leqslant c\int_{t_n}^{t_{n+1}} L_5\,|s-t_n|^{\gamma}\, ds$$
$$\leqslant \quad ch^{1+\gamma}.$$

https://doi.org/10.1112/S1461157000000322 Published online by Cambridge University Press

In case (ii) we obtain:

$$\begin{aligned}
&\left|\mathcal{E}\int_{t_n}^{t_{n+1}}\left\{\frac{\partial f}{\partial x_2}\big(X(t_n),X(t_n-\tau)\big)\big(X(s-\tau)-X(t_n-\tau)\big)\right\}ds\right| \\
&\leqslant\sqrt{\mathcal{E}\left(\int_{t_n}^{t_{n+1}}\frac{\partial f}{\partial x_2}\big(X(t_n),X(t_n-\tau)\big)\big(X(s-\tau)-X(t_n-\tau)\big)\,ds\right)^2} \\
&\leqslant\sqrt{h\mathcal{E}\int_{t_n}^{t_{n+1}}\left(\frac{\partial f}{\partial x_2}\big(X(t_n),X(t_n-\tau)\big)\right)^2\big(X(s-\tau)-X(t_n-\tau)\big)^2\,ds} \\
&\leqslant\sqrt{hc\int_{t_n}^{t_{n+1}}s-t_n\,ds} \\
&\leqslant ch^{\frac{3}{2}}.
\end{aligned}$$

- For the term (3) in equation (35), the remainder $\rho(s)$ has the form

$$\begin{aligned}
\rho(s)=&\ \frac{1}{2}\,\frac{\partial^2 f}{\partial x_1^2}(\phi,\varphi)\big(X(s)-X(t_n)\big)^2 \\
&+\frac{\partial^2 f}{\partial x_1\partial x_2}(\phi,\varphi)\big(X(s)-X(t_n)\big)\big(X(s-\tau)-X(t_n-\tau)\big) \\
&+\frac{1}{2}\,\frac{\partial^2 f}{\partial x_2^2}(\phi,\varphi)\big(X(s-\tau)-X(t_n-\tau)\big)^2,
\end{aligned}$$

where the derivatives of f are evaluated at appropriate intermediate values $X(t_n)\leqslant\phi\leqslant X(s)$ and $X(t_n-\tau)\leqslant\varphi\leqslant X(s-\tau)$. We can then calculate, using $2ab\leqslant a^2+b^2$ and assumption A4 on f,

$$|\rho(s)|\leqslant c\left(|X(s)-X(t_n)|^2+|X(s-\tau)-X(t_n-\tau)|^2\right). \tag{36}$$

We have, by inequality (36), assumption A2 on Ψ and inequality (18), in case (i):

$$\begin{aligned}
&\left|\mathcal{E}\int_{t_n}^{t_{n+1}}\rho(s)\,ds\right| \\
&\leqslant c\mathcal{E}\int_{t_n}^{t_{n+1}}|X(s)-X(t_n)|^2+|\Psi(s-\tau)-\Psi(t_n-\tau)|^2\,ds \\
&\leqslant c\int_{t_n}^{t_{n+1}}(s-t_n)\,ds+c\int_{t_n}^{t_{n+1}}(s-t_n)^{2\gamma}\,ds\quad\leqslant\quad ch^2+ch^{1+2\gamma}.
\end{aligned}$$

In case (ii), we obtain

$$\begin{aligned}
&\left|\mathcal{E}\int_{t_n}^{t_{n+1}}\rho(s)\,ds\right| \\
&\leqslant c\mathcal{E}\int_{t_n}^{t_{n+1}}|X(s)-X(t_n)|^2+|X(s-\tau)-X(t_n-\tau)|^2\,ds \\
&\leqslant c\int_{t_n}^{t_{n+1}}(s-t_n)\,ds\quad\leqslant\quad ch^2.
\end{aligned}$$

In summary, we obtain

$$|\mathcal{E}(\delta_{n+1})| \leqslant ch^{\min(1+\gamma,\frac{3}{2})},$$

so part 1(a) of the theorem follows. We have used properties of multiple Itô-integrals, which may be found in [19] and [21]. We have also employed the following estimate of the drift term:

$$\left|f\big(X(u), X(u-\tau)\big)\right| \leqslant \sqrt{K\left(1+2\sup_{-\tau\leqslant r\leqslant T}|X(r)|^2\right)},$$

which is an immediate consequence of the linear growth bound (14).

Now we prove part 1(b)—in other words, consistency in the mean-square, with order $p_2 = \min(1/2+\gamma, 1)$. We use the Hölder inequality, the Schwarz inequality for integrals, $2ab \leqslant a^2+b^2$, $(a+b)^2 \leqslant 2(a^2+b^2)$ and property (17). We have:

$$\begin{aligned}
\mathcal{E}\,|\delta_{n+1}|^2 \leqslant\ & \mathcal{E}\left(\int_{t_n}^{t_{n+1}} \left|f\big(X(s), X(s-\tau)\big) - f\big(X(t_n), X(t_n-\tau)\big)\right|\, ds\right)^2 \\
& +2\mathcal{E}\left\{\begin{array}{l} \left(\int_{t_n}^{t_{n+1}} \left|f\big(X(s), X(s-\tau)\big) - f\big(X(t_n), X(t_n-\tau)\big)\right|\, ds\right) \\ \times\left(\int_{t_n}^{t_{n+1}} \left|g\big(X(s), X(s-\tau)\big) - g\big(X(t_n), X(t_n-\tau)\big)\right|\, dW(s)\right)\end{array}\right\} \\
& +\mathcal{E}\left(\int_{t_n}^{t_{n+1}} \left|g\big(X(s), X(s-\tau)\big) - g\big(X(t_n), X(t_n-\tau)\big)\right|\, dW(s)\right)^2 \\
\leqslant\ & \mathcal{E}\left(\int_{t_n}^{t_{n+1}} \left|f\big(X(s), X(s-\tau)\big) - f\big(X(t_n), X(t_n-\tau)\big)\right|\, ds\right)^2 \\
& +2\left\{\begin{array}{l} \left(\mathcal{E}\left(\int_{t_n}^{t_{n+1}} \left|f\big(X(s), X(s-\tau)\big) - f\big(X(t_n), X(t_n-\tau)\big)\right|\, ds\right)^2\right)^{1/2} \\ \times\left(\int_{t_n}^{t_{n+1}} \mathcal{E}\left(\left|g\big(X(s), X(s-\tau)\big) - g\big(X(t_n), X(t_n-\tau)\big)\right|\right)^2 ds\right)^{1/2}\end{array}\right\} \\
& +\int_{t_n}^{t_{n+1}} \mathcal{E}\left(\left|g\big(X(s), X(s-\tau)\big) - g\big(X(t_n), X(t_n-\tau)\big)\right|\right)^2\, ds \\
\leqslant\ & 2\mathcal{E}\left(\int_{t_n}^{t_{n+1}} \left|f\big(X(s), X(s-\tau)\big) - f\big(X(t_n), X(t_n-\tau)\big)\right|\, ds\right)^2 \\
& +2\int_{t_n}^{t_{n+1}} \mathcal{E}\left(\left|g\big(X(s), X(s-\tau)\big) - g\big(X(t_n), X(t_n-\tau)\big)\right|\right)^2\, ds \\
\leqslant\ & 2\mathcal{E}\left(\int_{t_n}^{t_{n+1}} L_1\,|X(s)-X(t_n)| + L_2\,|X(s-\tau)-X(t_n-\tau)|\, ds\right)^2 \\
& +2\mathcal{E}\int_{t_n}^{t_{n+1}} \big(L_3|X(s)-X(t_n)| + L_4|X(s-\tau)-X(t_n-\tau)|\big)^2\, ds \\
\leqslant\ & 2h\mathcal{E}\left(\int_{t_n}^{t_{n+1}} \big(L_1|X(s)-X(t_n)| + L_2|X(s-\tau)-X(t_n-\tau)|\big)^2\, ds\right) \\
& +2\int_{t_n}^{t_{n+1}} \mathcal{E}\big(L_3|X(s)-X(t_n)| + L_4|X(s-\tau)-X(t_n-\tau)|\big)^2\, ds
\end{aligned}$$

https://doi.org/10.1112/S1461157000000322 Published online by Cambridge University Press

$$\leqslant 4h \int_{t_n}^{t_{n+1}} L_1^2\, \mathcal{E}\big(|X(s)-X(t_n)|^2\big) + L_2^2\, \mathcal{E}\big(|X(s-\tau)-X(t_n-\tau)|^2\big)\, ds$$
$$+4 \int_{t_n}^{t_{n+1}} L_3^2\, \mathcal{E}\big(|X(s)-X(t_n)|^2\big) + L_4^2\, \mathcal{E}\big(|X(s-\tau)-X(t_n-\tau)|^2\big)\, ds. \tag{37}$$

Again, there are the two cases to consider for the delayed arguments. In case (i) we obtain, by using assumption A2 on Ψ, that the value of inequality (37) is

$$\leqslant 4h \int_{t_n}^{t_{n+1}} L_1^2\, C_2(\|\Psi\|, T)(s-t_n) + L_2^2\, L_5(s-t_n)^{2\gamma}\, ds$$
$$+4 \int_{t_n}^{t_{n+1}} L_3^2\, C_2(\|\Psi\|, T)(s-t_n) + L_4^2\, L_5(s-t_n)^{2\gamma}\, ds$$
$$\leqslant cC_2(\|\Psi\|, T)h^2 + ch^{1+2\gamma}.$$

In case (ii) we calculate that inequality (37) is

$$\leqslant 4h \int_{t_n}^{t_{n+1}} L_1^2\, C_2(\|\Psi\|, T)(s-t_n) + L_2^2\, C_2(\|\Psi\|, T)(s-t_n)\, ds$$
$$+4 \int_{t_n}^{t_{n+1}} L_3^2\, C_2(\|\Psi\|, T)(s-t_n) + L_4^2\, C_2(\|\Psi\|, T)(s-t_n)\, ds$$
$$\leqslant cC_2\,(\|\Psi\|, T)h^2.$$

This implies that

$$\Big(\mathcal{E}|\delta_{n+1}|^2\Big)^{\frac{1}{2}} \leqslant ch^{\min\left(\frac{1}{2}+\gamma, 1\right)}.$$

Now consider part (2), and note that

$$\delta_{n+1} = \int_{t_n}^{t_{n+1}} f\big(X(s), X(s-\tau)\big) - f\big(X(t_n), X(t_n-\tau)\big)\, ds$$

for equations (20) with additive noise. In the case that assumption A5 holds (that is, the equation (20) has additive noise and a decomposable drift function f), we obtain for the term (2) in inequality (35) and case (ii)

$$\left|\mathcal{E} \int_{t_n}^{t_{n+1}} \left\{ \frac{\partial f}{\partial x_2}(X(t_n-\tau))\, (X(s-\tau)-X(t_n-\tau)) \right\} ds\right|$$
$$\leqslant c \left|\mathcal{E} \int_{t_n}^{t_{n+1}} \int_{t_n-\tau}^{s-\tau} f(X(u), X(u-\tau))\, du\, ds\right|$$
$$\leqslant c\mathcal{E}\left(\sqrt{K(1+2 \sup_{-\tau\leqslant r\leqslant T} |X(r)|^2)}\right) \cdot \left(\int_{t_n}^{t_{n+1}} \int_{t_n-\tau}^{s-\tau} du\, ds\right)$$
$$\leqslant c\left(\sqrt{K(1+2\mathcal{E} \sup_{-\tau\leqslant r\leqslant T} |X(r)|^2)}\right) h^2$$
$$\leqslant c\left(\sqrt{K(1+2C_1(\|\Psi\|, T)}\right) h^2$$

with $C_1(\|\Psi\|, T) = (1/2 + 4\mathcal{E}\|\Psi\|^2)\, e^{6K(T+4)T}$, due to inequality (16). Using this bound to modify the proof of part (1), the result in part (2) follows. □

https://doi.org/10.1112/S1461157000000322 Published online by Cambridge University Press

Remark 2. A modification of the proof gives, with $p_1 = 2$ and $p_2 = 1$, the corresponding result for the case of an SODE; see [**21**].

Lemma 2. *If the functions f and g in equation* (9) *satisfy the conditions of Theorem* 1*, then the increment function ϕ_{EM} of the Euler–Maruyama scheme* (*given by equation* (33)) *satisfies the estimates* (22) *and* (23) *for all $\xi, \xi', \eta, \eta' \in \mathbb{R}$.*

Proof. We use the Lipschitz-continuity of the drift and diffusion function and properties of multiple Itô-integrals, which may be found in [**19**] and [**21**].

$$\begin{aligned}
&\left|\mathcal{E}\left(\phi_{EM}(h,\xi,\eta,\Delta W_{n+1}) - \phi_{EM}(h,\xi',\eta',\Delta W_{n+1})\right)\right| \\
&\quad = \left|\mathcal{E}\left(hf(\xi,\eta) + g(\xi,\eta)\Delta W_{n+1} - h\,f(\xi',\eta') - g(\xi',\eta')\Delta W_{n+1}\right)\right| \\
&\quad \leqslant h\left|f(\xi,\eta) - f(\xi',\eta')\right| + \left|g(\xi,\eta) - g(\xi',\eta')\right|\ \left|\mathcal{E}(\Delta W_{n+1})\right| \\
&\quad \leqslant h\left(L_1\,|\xi-\xi'| + L_2\,|\eta-\eta'|\right)
\end{aligned}$$

$$\begin{aligned}
&\mathcal{E}\left(\left|\phi_{EM}(h,\xi,\eta,\Delta W_{n+1}) - \phi_{EM}(h,\xi',\eta',\Delta W_{n+1})\right|^2\right) \\
&\quad = \mathcal{E}\left(\left|hf(\xi,\eta) + g(\xi,\eta)\Delta W_{n+1} - hf(\xi',\eta') - g(\xi',\eta')\Delta W_{n+1}\right|^2\right) \\
&\quad \leqslant 2h^2\left|f(\xi,\eta) - f(\xi',\eta')\right|^2 + 2\left|g(\xi,\eta) - g(\xi',\eta')\right|^2 \mathcal{E}\,|\Delta W_{n+1}|^2 \\
&\quad \leqslant 4h^2\big(L_1^2\,|\xi-\xi'|^2 + L_2^2\,|\eta-\eta'|^2\big) + 4h\big(L_3^2\,|\xi-\xi'|^2 + L_4^2\,|\eta-\eta'|^2\big),
\end{aligned}$$

from which the estimates follow. □

Lemma 3. *If the functions f and g in equation* (9) *satisfy the conditions of Theorem* 1*, then the increment function ϕ_{EM} of the Euler–Maruyama scheme* (*given by equation* (33)) *satisfies the estimate* (24) *for all $\xi, \eta \in \mathbb{R}$.*

Proof. We use the linear growth bounds of the drift and diffusion function and properties of multiple Itô-integrals, which may be found in [**19**] and [**21**].

$$\begin{aligned}
&\mathcal{E}\left(|\phi_{EM}(h,\xi,\eta,\Delta W_{n+1})|^2\right) \\
&\quad = \mathcal{E}\left(|hf(\xi,\eta) + g(\xi,\eta)\Delta W_{n+1}|^2\right) \\
&\quad \leqslant 2h^2\,|f(\xi,\eta)|^2 + 2|g(\xi,\eta)|^2\mathcal{E}|\Delta W_{n+1}|^2 \\
&\quad \leqslant 2h^2K_1\left(1+|\xi|^2+|\eta|^2\right) + 2K_2\left(1+|\xi|^2+|\eta|^2\right)h,
\end{aligned}$$

from which the estimates follow. □

The next theorem follows from our previous results in Theorem 5 and Lemmas 2 and 3.

Theorem 6. (1) *Theorem* 3 *is valid, for the Euler–Maruyama method applied to equations* (20)*, under conditions* A1 – A4*, with order of convergence $p = \min(\gamma, 1/2)$ in the mean-square sense.*

(2) *With the additional assumption* A5 (*that is, for equations with additive noise and decomposable f*) *Theorem* 3 *is even valid for the Euler–Maruyama method with order of convergence $p = \min(\gamma, 1)$ in the mean-square sense.*

https://doi.org/10.1112/S1461157000000322 Published online by Cambridge University Press

6. Numerical experiments

The theoretical discussion of numerical processes is intended to provide an insight into the performance of numerical methods in practice. We have used the equation

$$dX(t) = \big(aX(t) + bX(t-1)\big)\,dt + \big(\beta_1 + \beta_2 X(t) + \beta_3 X(t-1)\big)\,dW(t)$$

as a test equation for our Euler–Maruyama method; we shall use this section to report on some numerical results for this equation, and to relate them (to a limited extent) to the theory presented above.

Concerning 'exact solutions', in the case of additive noise ($\beta_2 = \beta_3 = 0$) we have calculated an explicit solution on the first interval $[0, \tau]$ by the method of steps (see, for example, [9]), using $\Psi(t) = 1 + t$ for $t \in [-1, 0]$ as an initial function. The solution on $t \in [0, 1]$ is given by

$$X(t) = e^{at}\left(1 + \frac{b}{a^2}\right) - \frac{b}{a}t - \frac{b}{a^2} + \beta e^{at}\int_0^t e^{-as}\,dW(s).$$

We have then used this solution as an initial function to compute an 'explicit solution' on the second interval [1, 2] with a standard SODE-method and a small step-size. In the case of multiplicative noise we have computed an 'explicit solution' on a very fine grid (usually 2048 steps) with the Euler–Maruyama scheme.

Our tests concerned the illustration of the theoretical order of convergence. If we square both sides of inequality (30) in Theorem 6, conditions for which are satisfied in the examples, we see that the mean-square error $\mathcal{E}|X(T) - \widetilde{X}_N|^2$ should be bounded by Ch^{2p} for some C:

$$\mathcal{E}\left|X(T) - \widetilde{X}_N\right|^2 \leqslant Ch^{2p}. \tag{38}$$

In our experiments, the mean-square error at the final time $T = 2$ was estimated in the following way. A set of 20 blocks, each containing 100 outcomes ($\omega_{i,j}$; $1 \leqslant i \leqslant 20$, $1 \leqslant j \leqslant 100$), were simulated, and for each block the estimator

$$\epsilon_i = \frac{1}{100}\sum_{j=1}^{100}\left|X(T, \omega_{i,j}) - \widetilde{X}_N(\omega_{i,j})\right|^2$$

was formed. In Table 1, $\epsilon \equiv \epsilon(h)$ denotes the mean of this estimator, which was itself estimated in the usual way. Thus we have

$$\epsilon(h) := \frac{1}{20}\sum_{i=1}^{20}\epsilon_i \qquad \text{and} \qquad \epsilon(h) \approx \mathcal{E}\left|X(T) - \widetilde{X}_N\right|^2. \tag{39}$$

We therefore ask whether the numerical results suggest the existence of a constant C such that

$$\epsilon(h) \leqslant Ch^{2p}. \tag{40}$$

Using the set of coefficients

$$\begin{aligned}
\text{I:}\quad & a = -2, b = 0.1, \beta_1 = 1, \beta_2 = \beta_3 = 0,\\
\text{II:}\quad & a = -2, b = 0.1, \beta_2 = 1, \beta_1 = \beta_3 = 0,\\
\text{III:}\quad & a = -2, b = 0.1, \beta_3 = 1, \beta_1 = \beta_2 = 0,
\end{aligned}$$

we obtained the results (corresponding to $h_0 = 1/4$, $h_1 = h_0/2$, $h_2 = h_1/2$, $h_3 = h_2/2$) shown in Table 1.

Table 1: Estimated errors ϵ for the Euler–Maruyama method

h	I, ϵ	I, ratio	II, ϵ	II, ratio	III, ϵ	III, ratio
0.25	0.0184	*	0.1089	*	0.02011	*
0.125	0.00404	0.22	0.04913	0.45	0.00987	0.5
0.0625	0.000973	0.24	0.02437	0.5	0.004823	0.5
0.03125	0.000244	0.25	0.012135	0.5	0.0025	0.5
Suggested values of p		1		1/2		1/2

It is, of course, impossible to prove a result such as that in inequality (40) by numerical experimentation. In fact, however, the computed ratio of terms $\epsilon(h/2)/\epsilon(h)$ approximates $\{1/2\}^{2p}$ for an appropriate p (as suggested in the table), which is suggestive of the stronger result $\epsilon(h) = \mu_p h^{2p} + \mathcal{O}(h^{2p+1})$, at least for a restricted class of problems. Observe that for very small h, rounding error effects can obscure the behaviour predicted by such a result. To the best of our knowledge, the existence of an expansion of the error in the case of stochastic differential equations is established only for weak approximations.

In summary, the 'ratio' of errors, given in Table 1 for the approximations to the test equation, are consistent (in the sense indicated above) with the property

$$\left(\mathcal{E}|X(T) - \widetilde{X}_N|^2\right)^{\frac{1}{2}} = \mu_p(T)h^p + \mathcal{O}(h^{p+1})$$

and hence with the theoretical order of convergence as stated in Theorems 3 and 6.

7. *Further directions*

This paper provides an introduction to the numerical analysis of stochastic delay differential equations. We concentrated here on autonomous SDDEs; for an indication of the way in which the theory extends to non-autonomous equations refer to the comparable extension for SODEs [**21**]. When one seeks to advance the study further, one observes a number of open questions, involving (for example):

(a) classification of the terms involving time-lag (for example, a bounded or a fading memory);

(b) the design of numerical methods for more general problems;

(c) weak approximation methods;

(d) the stability and dynamic properties of the numerical methods;

(e) variable time-step algorithms.

For stochastic *ordinary* differential equations, the issues (d) and (e) have only recently attracted attention (see, for example, [**26**] and the relevant articles in [**8**], and [**11, 13, 20**], respectively). We hope to address such issues in sequels to this paper.

8. *Acknowledgements*

We thank the editor, the referees, and Dr. M. Tretyakov, for their interest and advice.

References

1. L. ARNOLD, *Stochastic differential equations: theory and applications* (Wiley-Interscience, New York, 1974). 315, 326
2. T. A. AVERINA and S. S. ARTEMIEV, *Numerical analysis of systems of ordinary and stochastic differential equations* (VSP, Utrecht, 1997). 321
3. C. T. H. BAKER, 'Retarded differential equations', *J. Comput. Appl. Math.*, in press. 315
4. C. T. H. BAKER, G. A. BOCHAROV, C. A. H. PAUL and F. A. RIHAN, 'Modelling and analysis of time-lags in some basic patterns of cell proliferation', *J. Math. Biol.* 37 (1998) 341–371. 316
5. A. BEUTER and J. BÉLAIR, 'Feedback and delays in neurological diseases: a modelling study using dynamical systems', *Bull. Math. Biol.* 55 (1993) 525–541. 317
6. K. BURRAGE and E. PLATEN, 'Runge–Kutta methods for stochastic differential equations', *Ann. Numer. Math.* 1 (1994) 63–78. 315
7. J. M. C. CLARK, 'The discretization of stochastic differential equations: a primer', *Road Vehicle Systems and Related Mathematics*, Proc. 2nd DMV-GAMM Workshop, Torino, 1987 (ed. H. Neunzert, Teubner, Stuttgart, and Kluwer Academic Publishers, Amsterdam, 1988) 163–179. 315
8. H. CRAUEL and M. GUNDLACH (eds), *Stochastic dynamics*, Papers from the Conference on Random Dynamical Systems, Bremen, April 28–May 2 1997 (Springer, New York, 1999). 333
9. R. D. DRIVER, *Ordinary and delay differential equations*, Appl. Math. Sci. 20 (Springer, New York, 1977). 332
10. C. W. EURICH and J. G. MILTON, 'Noise-induced transitions in human postural sway', *Phys. Rev. E* 54 (1996) 6681–6684. 317
11. J. G. GAINES and T. L. LYONS, 'Variable step size control in the numerical solution of stochastic differential equations', *SIAM J. Appl. Math.* 57 (1997) 1455–1484. 333
12. T. C. GARD, *Introduction to stochastic differential equations* (Marcel–Dekker, New York, Basel, 1988). 324
13. N. HOFMANN, T. MÜLLER-GRONBACH and K. RITTER, 'Optimal approximation of stochastic differential equations by adaptive step-size control', *Math. Comp.* 69 (2000) 1017–1034. 333
14. I. KARATZAS and S. E. SHREVE, *Brownian motion and stochastic calculus* (Springer, New York, 1991). 315
15. P. E. KLOEDEN and E. PLATEN, *Numerical solution of stochastic differential equations* (Springer, Berlin, 1992). 315, 320
16. V. KOLMANOVSKIĬ and A. MYSHKIS, *Applied theory of functional-differential equations* (Kluwer Academic Publishers Group, Dordrecht, 1992). 315
17. U. KÜCHLER and E. PLATEN, 'Strong discrete time approximation of stochastic differential equations with time delay', Discussion paper, Interdisciplinary Research Project 373 (ISSN 1436-1086) 25, Humboldt University, Berlin (1999). 315, 319
18. M. C. MACKEY, A. LONGTIN, J. G. MILTON and J. E. BOS, 'Noise and critical behaviour of the pupil light reflex at oscillation onset', *Phys. Rev. A* 41 (1990) 6992–7005. 317

19. X. Mao, *Stochastic differential equations and their applications* (Horwood Publishing, Chichester, 1997). 315, 317, 319, 319, 329, 331, 331

20. S. Mauthner, 'Step size control in the numerical solution of stochastic differential equations', *J. Comput. Appl. Math.* 100 (1998) 93–109. 333

21. G. N. Milstein, *Numerical integration of stochastic differential equations* (translated and revised from the 1988 Russian original) (Kluwer, Dordrecht, 1995). 315, 318, 320, 322, 324, 329, 331, 331, 331, 333

22. V. J. Mizel and V. Trutzer, 'Stochastic hereditary equations: existence and asymptotic stability', *J. Integral Equations* 7 (1984) 1–72. 319

23. S. E. A. Mohammed, *Stochastic functional differential equations* (Pitman Advanced Publishing Program, Boston, MA, 1984). 315, 319

24. S. E. A. Mohammed, 'Stochastic differential systems with memory. Theory, examples and applications', *Stochastic analysis and related topics VI*, The Geilo Workshop, 1996, Progress in Probability (ed. L. Decreusefond, J. Gjerde, B. Oksendal and A. S. Ustunel, Birkhäuser, 1998) 1–77. 316

25. E. Platen, 'An introduction to numerical methods for stochastic differential equations', *Acta Numerica* 8 (1999) 195–244. 315

26. Y. Saito and T. Mitsui, 'Stability analysis of numerical schemes for stochastic differential equations', *SIAM J. Numer. Anal.* 33 (1996) 2254–2267. 333

27. D. Talay, 'Simulation of stochastic differential systems', *Probabilistic methods in applied physics*, Lecture Notes in Phys. 451 (ed. P. Krée and W. Wedig, Springer, Berlin, 1995) 54–96. 315

28. M. Tudor, 'Approximation schemes for stochastic equations with hereditary argument' (in Romanian), *Stud. Cerc. Mat.* 44 (1992) 73–85. 315

29. D. Williams, *Probability with martingales* (Cambridge University Press, Cambridge, 1991). 315, 317, 322

30. M. Zennaro, 'Delay differential equations: theory and numerics', *Theory and numerics of ordinary and partial differential equations*, Leicester, 1994 (ed. M. Ainsworth, J. Levesley, W. A. Light and M. Marletta, Oxford University Press, New York, 1995) 291–333. 315

Christopher T. H. Baker cthbaker@maths.man.ac.uk
Evelyn Buckwar ebuckwar@maths.man.ac.uk

Department of Mathematics
The Victoria University of Manchester
Manchester M13 9PL