Electronic Transactions on Numerical Analysis.
Volume 11, pp. 131-151, 2000.

ISSN 1068-9613.

ETNA
Kent State University
etna@mcs.kent.edu

CONTINUOUS Θ-METHODS FOR THE STOCHASTIC PANTOGRAPH EQUATION*

CHRISTOPHER T.H. BAKER† AND EVELYN BUCKWAR‡

Abstract. We consider a stochastic version of the pantograph equation:

$$\begin{aligned} dX(t) &= \{aX(t) + bX(qt)\}\,dt + \{\sigma_1 + \sigma_2 X(t) + \sigma_3 X(qt)\}\,dW(t),\\ X(0) &= X_0, \end{aligned}$$

for $t \in [0,T]$, a given Wiener process W and $0 < q < 1$. This is an example of an Itô stochastic delay differential equation with unbounded memory. We give the necessary analytical theory for existence and uniqueness of a strong solution of the above equation, and of strong approximations to the solution obtained by a continuous extension of the Θ-Euler scheme ($\Theta \in [0,1]$). We establish $\mathcal{O}(\sqrt{h})$ mean-square convergence of approximations obtained using a bounded mesh of uniform step h, rising in the case of additive noise to $\mathcal{O}(h)$. Illustrative numerical examples are provided.

Key words. stochastic delay differential equation, continuous Θ-method, mean-square convergence.

AMS subject classifications. 65C30, 65Q05.

1. Introduction. The paper is organised as follows: In this section, we set the scene and motivate the later discussion. Section 2 provides the necessary theoretical background concerning a stochastic version of the pantograph equation; in §3 we detail the construction of an approximate solution on a uniform mesh of width h and give necessary definitions. We state the main theoretical results in Theorem 3.3 and its corollary. In §4 we present some numerical experiments; finally, §5 contains the proofs of Theorems 3.2 and 3.3.

1.1. Background on the deterministic and stochastic equations. To orientate the reader, we shall relate the problem of interest

$$dX(t) = \Big\{aX(t) + bX(qt)\Big\}dt + \Big\{\sigma_1 + \sigma_2 X(t) + \sigma_3 X(qt)\Big\}dW(t), \tag{1.1a}$$

$$X(0) = X_0, \tag{1.1b}$$

to the corresponding deterministic *pantograph equation*

$$y'(t) \;=\; ay(t) \;+\; by(qt), \qquad q \in (0,1), \tag{1.2a}$$

in which it is conventional to take $y'(t)$ to denote the right-hand derivative. If (1.2a) holds for $t \geq 0$ then a solution is uniquely determined by an intial value:

$$y(0) \;=\; y_0. \tag{1.2b}$$

If (1.2a) holds for $t \geq t_0 > 0$ then a solution is uniquely determined by requiring

$$y(t) \;=\; \psi(t) \qquad (qt_0 \leq t \leq t_0), \tag{1.2c}$$

where $\psi(\cdot)$ is a given intial function. In the latter case, even if $\psi(\cdot)$ is smooth on $[qt_0, t_0]$, successively higher derivatives of the solution display 'jumps' at $\frac{1}{q}t_0$, $\frac{1}{q^2}t_0$, $\frac{1}{q^3}t_0$, ... in the

*Received May 12, 2000. Accepted for publication December 4, 2000. Recommended by L. Reichel.
†Department of Mathematics, University of Manchester, Manchester M13 9PL, UK.
‡Department of Mathematics, University of Manchester, Manchester M13 9PL, UK. This research was supported by Marie Curie Individual Fellowship HPMF-CT-1999-00090.

ETNA
Kent State University
etna@mcs.kent.edu

case that $\psi'(t_0) \neq a\psi(t_0) + b\psi(qt_0)$. Some of the difficulties involved in the numerical simulation of long-time behaviour of the solution of (1.2a–b) are discussed in [17], where earlier work, including [12], is also cited.

Since $qt < t$ when $t \geq 0$, (1.2a) is (for $t \geq 0$) a *differential equation with time lag*. The quantity $t - qt$ will be called *the lag* and we note that the delayed argument qt satisfies $qt \to \infty$ as $t \to \infty$ but the lag is unbounded. Equation (1.2a) provides an example of what are frequently called *delay differential equations*. By analogy, (1.1a), which we term the *stochasic pantograph equation* is an example of a *stochastic delay differential equation*.

The problem (1.1a) occurs when we allow *(i)* for external noise acting on the system, i.e. $\sigma_1 \neq 0$, $\sigma_2 = \sigma_3 = 0$, or *(ii)* for noise in the internal parameters a and b corresponding to $\sigma_2 \neq 0$ and $\sigma_3 \neq 0$, respectively, in our model. We observe that (1.1) can be expressed equivalently (incorporating the initial condition) as

$$X(t) = X_0 + \int_0^t aX(s) + bX(qs)\,ds + \int_0^t \sigma_1 + \sigma_2 X(s) + \sigma_3 X(qs)\,dW(s), \tag{1.3}$$

where we employ the Itô integral. This is an initial-value problem with solution defined by X_0. If the problem were defined for $t \geq t_0 > 0$ it would be necessary (compare (1.2c)) to define an initial function $X(t) = \Psi(t)$ for $t \in [qt_0, t_0]$ and the properties of $\Psi(t)$ would affect the discussion. In this article we will be interested in obtaining approximations to strong solutions of the stochastic pantograph equation. To the best of our knowledge, approximation schemes for stochastic equations with varying lag have not previously been discussed in the literature.

We shall refer to (1.2a–b) as the *underlying deterministic version* of (1.1a–b), and to (1.1a–b) as the stochastic analogue of (1.2a–b). The solution trajectories defined by (1.1a–b) are continuous but not smooth. In consequence the smoothness of solutions of the underlying deterministic problem appears to be of no interest. This claim bears closer examination in the small noise case but, for the present, we consider the initial value problem (1.1a–b) for which the underlying deterministic problem has a smooth solution.

Remark 1: Equation (1.2a) and its generalisations possess a wide range of applications. Equation (1.2a) arises, for example, in the analysis of the dynamics of an overhead current collection system for an electric locomotive or in the problem of a one-dimensional wave motion, such as that due to small vertical displacements of a stretched string under gravity, caused by an applied force which moves along the string ([12] and [26]). Existence, uniqueness and asymptotic properties of the solution of (1.2) have been considered in [7, 15]. The equation (1.2) can be used as a paradigm for the construction of numerical schemes for functional-differential equations with unbounded memory, *cf.* [5], [14], [16] (we do not attempt to give a complete list of references here).

1.2. Additional background. For the theoretical prerequisites on probability concepts we refer to [29]. Stochastic calculus and stochastic ordinary differential equation (SODEs) are treated in [1] and [19]; for the theory of stochastic delay differential equation (SDDEs), see (for example) [21, 24, 25]. SDDEs with general infinite memory are treated in [18], the case of fading memory is considered in [23].

One might expect the numerical analysis of delay differential equation (DDEs) and of SODEs to have some bearing upon the problems that concern us here. We refer to [2, 4, 30] for an overview of the issues in the numerical treatment of DDEs. For an overview of applications and objectives of numerical methods for SODEs, see [8], [27] or [28]; for more extensive treatments see [20, 22].

The numerics of evolutionary problems for ordinary differential equations (ODEs), for SODES, for DDEs, and for SDDEs have some features in common. They are associated with a *mesh*

$$\mathcal{T}_N := t_0 < t_1 < t_2 \ldots < t_N, \qquad \text{with } h_n := t_{n+1} - t_n, \tag{1.4}$$

($\max h_n$ is known as the *width* of $\mathcal{T}_N$) such that the computation of a solution (trajectory) proceeds by determining approximate values on the points $\{t_n\}_{n\geq 0}$ (and thence, if necessary, determining a densely-defined approximation on the subinterval $[t_n, t_{n+1}]$). In this paper we shall consider the case of a uniform mesh with $h_n \equiv h$. Since, for $q \in (0,1)$, it is usually the case that $qt_n \notin \mathcal{T}_N$, we shall then require a densely defined approximation. Note that introducing implicitness into the explicit term in (1.5) can produce unbounded solution values in an individual trajectory (see [20], p. 336); Burrage and Tian [6] have proposed and analysed a method that deals with this type of problem. Our numerical scheme is based on a choice of parameter $\Theta \in [0,1]$ and (for additional detail, see §3) the construction of approximations $\widetilde{X}(t) \approx X(t)$ satisfying the equations $\widetilde{X}(t_0) = X_0$ and, for $0 \leq r \leq N-1$ and $\zeta \in [0,1]$,

$$\begin{aligned} \widetilde{X}(t_r + \zeta h) \;&=\; \widetilde{X}(t_r) \\ &+ \underbrace{\int_{t_r}^{t_r+\zeta h} \Big\{\Theta(a\widetilde{X}(t_{r+1}) + b\widetilde{X}(qt_{r+1})) + (1-\Theta)(a\widetilde{X}(t_r) + b\widetilde{X}(qt_r))\Big\} ds}_{\text{implicit if } \Theta \neq 0} \\ &+ \underbrace{\int_{t_r}^{t_\ell+\zeta h} \Big\{\sigma_1 + \sigma_2\widetilde{X}(t_r) + \sigma_3\widetilde{X}(qt_r)\Big\}\, dW(s);}_{\text{explicit}} \end{aligned} \tag{1.5}$$

of course, the above integrals simplify. For comparison, we observe that, by (1.3), $X(t)$ satisfies the corresponding relation

$$X(t) = X(t_r) + \int_{t_r}^{t} aX(s) + bX(qs)\, ds + \int_{t_r}^{t} \sigma_1 + \sigma_2 X(s) + \sigma_3 X(qs)\, dW(s),$$

for $t \in [t_r, t_{r+1}]$. On $\mathcal{T}_N$, the approximations satisfy (for $n \in \{0, 1, \ldots, N-1\}$)

$$\begin{aligned} \widetilde{X}(t_{n+1}) = \widetilde{X}(t_n) + h\Big\{\Theta(a\widetilde{X}(t_{n+1}) + b\widetilde{X}(qt_{n+1})) + (1-\Theta)(a\widetilde{X}(t_n) + b\widetilde{X}(qt_n))\Big\} \\ + \Big\{\sigma_1 + \sigma_2\widetilde{X}(t_n) + \sigma_3\widetilde{X}(qt_n)\Big\}\left(W_{(n+1)h} - W_{nh}\right), \end{aligned}$$

wherein $\Delta W_{n+1} := W_{(n+1)h} - W_{nh}$, denotes independent $N(0,h)$-distributed Gaussian random variables and $\widetilde{X}(qt_n)$ is obtained from (1.5). We shall implement the scheme for computing values of $\widetilde{X}(qt_n)$ in an efficient mode. The corresponding deterministic version is easily recognised as the linear Θ–method.

2. Theoretical analysis of the stochastic pantograph equation. Let $(\Omega, \mathcal{A}, P)$ be a complete probability space with a filtration $(\mathcal{A}_t)$ satisfying the usual conditions, i.e. the filtration $(\mathcal{A}_t)_{t\geq 0}$ is right-continuous, and each $\mathcal{A}_t$, $t \geq 0$, contains all P-null sets in $\mathcal{A}$. 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 ($0 = t_0 < T < \infty$)

$$\begin{aligned} dX(t) = \{a\, X(t) \;+\; b\, X(qt)\}\, dt \\ + \{\sigma_1 \;+\; \sigma_2\, X(t) \;+\; \sigma_3\, X(qt)\}\, dW(t), \quad t \in [0,T], \end{aligned} \tag{2.1a}$$

ETNA
Kent State University
etna@mcs.kent.edu

$$X(0) \;=\; X_0, \tag{2.1b}$$

with $0 < q < 1$ and $\mathcal{E}|X_0|^2 < \infty$. The first term on the right-hand-side of (2.1a) is usually called the *drift* function and the second term is called the *diffusion* function. The integral version of equation (2.1a) is given by

$$X(t) = X_0 + \int_0^t aX(s) + bX(qs)\,ds + \int_0^t \sigma_1 + \sigma_2 X(s) + \sigma_3 X(qs)\,dW(s), \tag{2.2}$$

for $t \in [0,T]$. The second integral in (2.2) is a stochastic integral which is to be interpreted in the Itô sense. If $\sigma_2 = \sigma_3 = 0$ the equation has *additive noise*, otherwise the equation has *multiplicative noise*.

For an $\mathbb{R}$-valued stochastic process $U(t) : [0,T] \times \Omega \to \mathbb{R}$, we write

$$\mathcal{E}(U) = \int_\Omega U\,dP,$$

and we say that

$$U \in \mathcal{L}^2 = \mathcal{L}^2(\Omega, \mathcal{A}, P) \quad \text{if} \quad \mathcal{E}\left(|U|^2\right) \;<\; \infty;$$

we define the norm

$$\|U\|_2 \;=\; \left(\mathcal{E}\left(|U|^2\right)\right)^{\frac{1}{2}},$$

and in this article we will prove convergence in the norm $\|.\|_2$ of a numerical approximation, as the underlying mesh is refined.

DEFINITION 2.1. *An $\mathbb{R}$-valued stochastic process $X(t) : [0,T] \times \Omega \to \mathbb{R}$ is called a strong solution of (2.1), if it is a measurable, sample-continuous process such that $X|[0,T]$ is $(\mathcal{A}_t)_{0 \le t \le T}$-adapted, and X satisfies (2.1a) or (2.2), almost surely, and satisfies the initial condition $X(0) = X_0$. A solution $X(t)$ is said to be path-wise unique if any other solution $\widehat{X}(t)$ is stochastically indistinguishable from it, i.e.*

$$P\Big(X(t) \;=\; \widehat{X}(t) \text{ for all } 0 \le t \le T\Big) \;=\; 1.$$

THEOREM 2.2. *If $0 < q < 1$ and $\mathcal{E}|X_0|^2 < \infty$, then there exists a path-wise unique strong solution to problem (2.1).*

Proof. Existence and uniqueness can be established in the usual way by successive approximations, setting

$$\begin{aligned} X_0(t) &= X_0, \\ X_n(t) &= X_0 + \int_0^t \{a\, X_{n-1}(s) \;+\; b\, X_{n-1}(qs)\}\; ds \\ &\quad + \int_0^t \{\sigma_1 \;+\; \sigma_2\, X_{n-1}(s) \;+\; \sigma_3\, X_{n-1}(qs)\}\; dW(s). \end{aligned}$$

More general cases have been treated in [18]. □

We adapt a theorem from Mao ([21], Lemma 5.5.2), that we will make use of in our analysis.

ETNA
Kent State University
etna@mcs.kent.edu

THEOREM 2.3. *The solution of problem (2.1) has the property*

$$\mathcal{E}\left(\sup_{0\le t\le T}|X(t)|^2\right) \le C_1(T), \tag{2.3}$$

with

$$C_1(T) := (\frac{1}{2} + 3\,\mathcal{E}|X_0|^2)\exp(18\,K\,(T+4)\,T), \tag{2.4a}$$

$$K := \max(|a|, |b|, |\sigma_1|, |\sigma_2|, |\sigma_3|). \tag{2.4b}$$

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

$$\mathcal{E}|X(t) - X(s)|^2 \quad \le \quad C_2(T)(t-s), \tag{2.5}$$

where $C_2(T) = 12K(1 + 2C_1(T))$*.*

Proof. Using the inequality $(a+b+c)^2 \le 3(a^2+b^2+c^2)$, Hölder's inequality and Doob's martingale inequality (both may be found in [20, 21]), we can derive for every $t \in [0, T]$

$$\begin{aligned}
\mathcal{E}\left(\sup_{0\le s\le t}|X(s)|^2\right) &\le 3\,\mathcal{E}|X_0|^2 + 3\,K\,\mathcal{E}\Big(\sup_{0\le s\le t}|\int_0^s 1\;+\;X(r)\;+\;X(qr)\,dr|^2\Big)\\
&\quad +3\,K\mathcal{E}\Big(\sup_{0\le s\le t}|\int_0^s 1\;+\;X(r)\;+\;X(qr)\,dW(r)|^2\Big)\\
&\le 3\,\mathcal{E}|X_0|^2 + 9\,K\,T\int_0^t 1\;+\;\mathcal{E}|X(s)|^2\;+\;\mathcal{E}|X(qs)|^2\,ds\\
&\quad +36\,K\int_0^t 1\;+\;\mathcal{E}|X(s)|^2\;+\;\mathcal{E}|X(qs)|^2\,ds\\
&\le 3\,\mathcal{E}|X_0|^2 + 9\,K\,(T+4)\int_0^t 1\;+\;2\mathcal{E}\left(\sup_{0\le r\le s}|X(r)|^2\right)\,ds.
\end{aligned}$$

Hence

$$\begin{aligned}
&\frac{1}{2} + \mathcal{E}\left(\sup_{0\le s\le t}|X(s)|^2\right)\\
&\quad\le \frac{1}{2} + 3\,\mathcal{E}|X_0|^2 + 18\,K\,(T+4)\int_0^t \frac{1}{2} + \mathcal{E}\left(\sup_{0\le r\le s}|X(r)|^2\right)\,ds,
\end{aligned}$$

from which the Gronwall inequality (see, *e.g.* [20, 21]) gives

$$\frac{1}{2}\;+\;\mathcal{E}\left(\sup_{0\le s\le t}|X(s)|^2\right) \le (\frac{1}{2} + 3\,\mathcal{E}|X_0|^2)\exp(18\,K\,(T+4)\,t),$$

and the required inequality (2.3) follows.

The proof of (2.5) is straightforward: For $t - s < 1$ we obtain

$$\begin{aligned}
\mathcal{E}|X(t) - X(s)|^2 &\le 2\,K\,\mathcal{E}|\int_s^t X(r)\;+\;X(qr)\,dr|^2\\
&\quad + 2\,K\,\mathcal{E}|\int_s^t 1\;+\;X(r)\;+\;X(qr)\,dW(r)|^2
\end{aligned}$$

$$\begin{aligned}
&\le 6\,K\,(t-s)\int_s^t 1\;+\;\mathcal{E}|X(r)|^2\;+\;\mathcal{E}|X(qr)|^2\,dr\\
&\quad +6\,K\int_s^t 1\;+\;\mathcal{E}|X(r)|^2\;+\;\mathcal{E}|X(qr)|^2\,dr\\
&\le 6\,K\,(t-s+1)\int_s^t 1\;+\;\mathcal{E}|X(r)|^2\;+\;\mathcal{E}|X(qr)|^2\,dr\\
&\le 12\,K\,(1\;+\;2\,C_1(T))(t-s).\;\square
\end{aligned}$$

LEMMA 2.4. *As a consequence of the preceding theorem we obtain the following estimate for all* $t \in [0,T]$

$$\mathcal{E}|aX(t)+bX(qt)| \quad\le\quad \sqrt{2\,L\,C_1(T)}, \tag{2.6}$$

with $L = 2\ \max(|a|,|b|)$.

Proof. We have

$$\begin{aligned}
|aX(t)+bX(qt)| &= \sqrt{|aX(t)+bX(qt)|^2}\\
&\le \sqrt{L(|X(t)|^2+|X(qt)|^2)}\\
&\le \sqrt{2\,L\sup_{0\le r\le T}|X(r)|^2}.
\end{aligned}$$

Since $\mathcal{E}(\sqrt{\eta}\cdot 1) \le (\mathcal{E}(\sqrt{\eta})^2)^{\frac{1}{2}}\cdot 1$ holds for any $\eta\in\mathbb{R}^+$ we get

$$\begin{aligned}
\mathcal{E}|aX(t)+bX(qt)| &\le \sqrt{2\,L\,\mathcal{E}(\sup_{0\le r\le T}|X(r)|^2)}\\
&\le \sqrt{2\,L\,C_1(T)}.\;\square
\end{aligned}$$

3. Numerical analysis. Whilst we discussed in [3] the numerical solution of an SDDE with a fixed lag, there appears to be no analysis of numerical methods for SDDEs with varying lag. We propose and analyze a continuous Θ-method to obtain strong approximations of the solution to (2.1). We have a mesh

$$\mathcal{T}_N := \{t_0,t_1,\dots,t_N\},\quad t_n = n\cdot h,\ n=0,\dots,N,\quad h=T/N, \tag{3.1}$$

with a fixed uniform step h on the interval $[0,T]$. Since, in general, the points qt_n will not be mesh-points and in fact we can define a second non-uniform mesh, which consists of all the $t_n\in\mathcal{T}_N$ and all the points qt_n:

$$\mathcal{S}_{2N} := \{s_0,s_1,\dots,s_{2N}\},\quad s_0=t_0,\ s_1=qt_1,\dots,s_{2N}=t_N. \tag{3.2}$$

We can also represent any point $s_\ell\in\mathcal{S}_{2N}$, with $t_n<s_\ell\le t_{n+1}$, by

$$s_\ell = t_n+\zeta h,\qquad \text{where } t_n,t_{n+1}\in\mathcal{T}_N \text{ and } \zeta\equiv\zeta(s_\ell)\in(0,1]. \tag{3.3}$$

We define an approximate solution $\widetilde{X}(t)$ for all $t\ge 0$, which depends upon the parameter Θ with $\Theta\in[0,1]$:

$$\widetilde{X}(t_0) = X_0, \tag{3.4a}$$

ETNA
Kent State University
etna@mcs.kent.edu

$$
\begin{aligned}
&\widetilde{X}(t_{n+1}) = \widetilde{X}(t_n) \\
&\qquad + h\left\{\Theta\left(a\widetilde{X}(t_{n+1}) + b\widetilde{X}(qt_{n+1})\right) + (1-\Theta)\left(a\widetilde{X}(t_n) + b\widetilde{X}(qt_n)\right)\right\} \\
&\qquad + \left\{\sigma_1 + \sigma_2\widetilde{X}(t_n) + \sigma_3\widetilde{X}(qt_n)\right\}\, \Delta W_{n+1},
\end{aligned} \tag{3.4b}
$$

for $0 \le n \le N-1$ and with $\Delta W_{n+1} := W_{(n+1)h} - W_{nh}$, denoting independent $N(0,h)$-distributed Gaussian random variables. We can express (3.4b) equivalently as

$$
\begin{aligned}
&\widetilde{X}(t_{n+1}) = \widetilde{X}(t_n) \\
&\quad + \int_{t_n}^{t_{n+1}} \left\{\Theta\left(a\widetilde{X}(t_{n+1}) + b\widetilde{X}(qt_{n+1})\right) + (1-\Theta)\left(a\widetilde{X}(t_n) + b\widetilde{X}(qt_n)\right)\right\}\, ds \\
&\quad + \int_{t_n}^{t_{n+1}} \left\{\sigma_1 + \sigma_2\widetilde{X}(t_n) + \sigma_3\widetilde{X}(qt_n)\right\}\, dW(s),
\end{aligned} \tag{3.5}
$$

for $0 \le n \le N-1$.

As the approximations of $X(qt_n)$ are not provided by (3.4b), we need a continuous extension that permits evaluation of $\widetilde{X}(s_r)$ at any point $s_r \in \mathcal{S}_{2N}$. For this purpose we modify (3.5) to compute the values $\widetilde{X}(qt_n)$ at $qt_n = t_\ell + \zeta h$ with $\zeta \in (0,1]$:

$$
\begin{aligned}
&\widetilde{X}(t_\ell + \zeta h) = \widetilde{X}(t_\ell) \\
&\quad + \int_{t_\ell}^{t_\ell+\zeta h} \left\{\Theta\left(a\widetilde{X}(t_{l+1}) + b\widetilde{X}(qt_{l+1})\right) + (1-\Theta)\left(a\widetilde{X}(t_\ell) + b\widetilde{X}(qt_\ell)\right)\right\}\, ds \\
&\quad + \int_{t_\ell}^{t_\ell+\zeta h} \left\{\sigma_1 + \sigma_2\widetilde{X}(t_\ell) + \sigma_3\widetilde{X}(qt_\ell)\right\}\, dW(s).
\end{aligned} \tag{3.6}
$$

We require $\widetilde{X}(t_n)$ to be $\mathcal{A}_{t_n}$-measurable at the meshpoints $t_n,\ n = 0, \dots N$.

Remark 2: (i) When $\Theta \neq 0$, the formula in (3.4b) is an implicit equation for $\widetilde{X}(t_{n+1})$. *(ii)* Such a formula (that is, with $\Theta \neq 0$) would normally be called "semi-implicit" because it is implicit only in the drift but is explicit in the diffusion. In the deterministic case, $\Theta \in [\frac{1}{2}, 1]$ gives better stability properties than $\Theta \in [0, \frac{1}{2})$. However, one would suspect that, in the stochastic case, the semi-implicit methods with $\Theta \in [\frac{1}{2}, 1]$ need not have *robust* stability properties (in particular when the diffusion coefficients are large). We provide some indication of stability in our numerical examples but we do not, in this paper, analyze stability properties theoretically. *(iii)* For $qt_{n+1} \in [t_n, t_{n+1}]$ (this happens for example at $n = 0$, when we wish to compute $\widetilde{X}(t_1)$), the formula in (3.6) is an implicit equation for $\widetilde{X}(qt_{n+1})$, containing the unknown quantity $\widetilde{X}(t_{n+1})$. To proceed, one solves (3.6) for $\widetilde{X}(qt_{n+1})$ and inserts the resulting equation into (3.4b). The latter then can be solved for $\widetilde{X}(t_{n+1})$. *(iv)* Using the fixed stepsize h for the approximations $\widetilde{X}(t_n),\ t_n \in \mathcal{T}_N$ makes it possible to construct the path of the Wiener process from left to right on the fixed, but non-uniform mesh $\mathcal{S}_{2N}$.

DEFINITION 3.1. *(i) The local truncation error for the continuous Θ-method is defined, for a given Θ, $t_n \in \mathcal{T}_N$ and $\zeta \in [0,1]$ as the sequence of random variables*

$$\delta_h(t_n, \zeta) = X(t_n + \zeta h)$$

ETNA
Kent State University
etna@mcs.kent.edu

$$- \Big\{ X(t_n) + \int_{t_n}^{t_n+\zeta h} \Big(\Theta \, (a\, X(t_{n+1}) + b\, X(qt_{n+1})) \Big) \, ds$$

$$+ \int_{t_n}^{t_n+\zeta h} \Big((1-\Theta) \, (a\, X(t_n) + b\, X(qt_n)) \Big) \, ds$$

$$+ \int_{t_n}^{t_n+\zeta h} \Big(\sigma_1 + \sigma_2\, X(t_n) + \sigma_3\, X(qt_n) \Big) \, dW(s) \Big\}.$$

(ii) The Θ-method is said to be consistent with order p_1 in the mean and with order p_2 in the mean-square sense *if the following estimates hold with $p_2 \geq \frac{1}{2}$ and $p_1 \geq p_2 + \frac{1}{2}$:*

$$\max_{0 \leq n \leq N-1} \sup_{\zeta \in [0,1]} |\mathcal{E}(\delta_h(t_n, \zeta))| \quad \leq \quad C\, h^{p_1} \;\; \text{as } h \to 0, \tag{3.7}$$

and

$$\max_{0 \leq n \leq N-1} \sup_{\zeta \in [0,1]} \left(\mathcal{E}|\delta_h(t_n, \zeta)|^2 \right)^{\frac{1}{2}} \quad \leq \quad C\, h^{p_2} \;\; \text{as } h \to 0, \tag{3.8}$$

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

(iii) We denote by $\epsilon(t_n)$ the global error of the Θ-method, i.e. the sequence of random variables

$$\epsilon(t_n) \;=\; X(t_n) \;-\; \widetilde{X}(t_n), \qquad n = 1, \ldots, N. \tag{3.9}$$

Note that $\epsilon(t_n)$ is $\mathcal{A}_{t_n}$-measurable since both $X(t_n)$ and $\widetilde{X}(t_n)$ are $\mathcal{A}_{t_n}$-measurable random variables.

(iv) For fixed $T < \infty$ the approximations $\widetilde{X}(\cdot)$ are convergent in the mean-square sense on meshpoints with order p if

$$\max_{1 \leq n \leq N} \left(\mathcal{E}|\epsilon(t_n)|^2 \right)^{\frac{1}{2}} \quad \leq \quad C\, h^{p} \;\; \text{as } h \to 0. \tag{3.10}$$

THEOREM 3.2. *If $0 < q < 1$, the continuous Θ-method for problem (2.1) is consistent (a) with order 2 in the mean and (b) with order 1 in the mean square.*

We now state the main theorem of this article.

THEOREM 3.3. *Suppose that $0 < q < 1$. Then*

1. *The recurrence relation (3.4b) has a unique solution, if*

$$h < h^* = \begin{cases} \dfrac{1}{\Theta\,(|a| + q|b|)}, & \Theta \in (0, 1], \\ +\infty & \Theta = 0, \end{cases} \tag{3.11}$$

and

2. *The approximation (3.4) of the strong solution of (2.1) on $[0, T]$ is convergent on mesh-points in the mean-square sense with order $\frac{1}{2}$.*

The proofs of Theorems 3.2 and 3.3 are given in Section 5. As a corollary to Theorem 3.3 we have the following result.

THEOREM 3.4. *If equation (2.1a) has additive noise, then the Θ-method is consistent with order $p_1 = 2$ in the mean and order $p_2 = 3/2$ in the mean square, which implies an order of convergence $p = 1$ in the mean-square-sense.*

ETNA
Kent State University
etna@mcs.kent.edu

Proof. In this case the term containing the Itô integral in the expression of the local error $\delta_h(t_n, \zeta)$ vanishes, an appropriate modification of the proof of Theorem 3.2, part b), yields the desired result. $\square$

Remark 3: In [10] it was shown that a method must have an order of convergence of at least 1 to guarantee convergence to the correct solution of a stochastic ordinary differential equation if adaptive step-size techniques are applied. This implies that one cannot hope to employ adaptive step-size techniques successfully, with the Θ-method.

4. Numerical experiments. We have performed numerical experiments in which the method (3.4) is applied to

$$\begin{aligned} dX(t) &= \{a\, X(t) \;+\; b\, X(qt)\}\, dt \\ &\quad + \{\sigma_1 \;+\; \sigma_2\, X(t) \;+\; \sigma_3\, X(qt)\}\, dW(t), \quad t \in [0, T], \end{aligned} \tag{4.1}$$

with initial value $X(0) = 1$. In the examples I, II and III we have used the parameters $T = 2,\ a = -1.5,\ b = 1,\ q = 0.5$ in (4.1) and method (3.4) with $\theta = 1$. Each test was performed with two levels of noise intensity in one of the parameters $\sigma_i,\ i = 1, 2, 3$. In the examples IV, V, and VI we have used the parameters $T = 2,\ a = -16,\ b = 2,\ q = 0.8$ in (4.1) and method (3.4) with the stepsize $h = 1/8$.

One of our tests concerned the illustration of the theoretical order of convergence. If we square both sides of (3.10) we obtain the mean-square error $\mathcal{E}|X(T) - \widetilde{X}_N|^2$ which should be bounded by $C\, h^{2p}$. The mean-square-error at the final time T was estimated in the following way. A set of 20 blocks each containing 100 outcomes $(\omega_{i,j}; 1 \leq i \leq 20, 1 \leq j \leq 100)$, were simulated and for each block the estimator $\epsilon_i \;=\; \frac{1}{100} \sum_{j=1}^{100} |X(T, \omega_{i,j}) - \widetilde{X}_N(\omega_{i,j})|^2$ was formed. The 'explicit solution' was computed on a very fine mesh (usually 2048 steps). In the table below ϵ denotes the mean of this estimator, which was itself estimated in the usual way: $\epsilon = \frac{1}{20} \sum_{i=1}^{20} \epsilon_i$.

In deterministic numerical analysis one can sometimes establish the existence of an expansion of an error $\varepsilon(h)$ in terms of powers of the step size h, of the form (say) $\varepsilon(h) = \mu_1(t) h^{\nu_1} + \mathcal{O}(h^{\nu_2})$ with $\nu_2 > \nu_1 > 0$ and $\mu_1 \neq 0$. One can then estimate the asymptotic rate of convergence ν_1 by computing $|\varepsilon(h)|$ for differing, sufficiently small, $h > 0$; in particular, $|\varepsilon(h/2)|/|\varepsilon(h)| \to \frac{1}{2}^{\nu_1}$ as $h \to 0$. In the case of stochastic differential equations, the existence of an expansion of the error is, to the best of our knowledge, only established for weak approximations. In spite of that, the 'ratio' of errors, given below for the approximations to the pantograph equation in the tables are consistent with the theoretical order of convergence, as stated in Theorem 3.3 (see examples I and II in Tables 4.1 and 4.2, respectively) and Theorem 3.4 (see example III in Table 4.3). It may therefore be possible to show theoretically that, under appropriate conditions, $\{\mathcal{E}|X(T) - \widetilde{X}_N|^2\}^{\frac{1}{2}} = \mu_p h^p + \mathcal{O}(h^{p+1})$.

The Figures 4.1, 4.2 and 4.3 (results for examples IV, V, and VI) suggest an improvement in the stability behaviour of the Θ-methods, when changing from explicit ($\Theta = 0$) to implicit ($\Theta = 0.5$).

In what follows we present two intriguing examples of behaviour of trajectories. In the first one we add noise to the example of the deterministic pantograph equation, that was considered (and computed numerically) in [12] and [17]:

$$y'(t) \;=\; 0.95 y(t) - y(0.99t), \quad y(0) = 1, t \in [0, 150]. \tag{4.2}$$

Our numerical solution decreases rapidly, stays very small for a considerably long period of time and then increases exponentially, as does the exact solution. We have computed

$$dX(t) \;= \{0.95 X(t) - X(0.99t)\}\, dt + \sigma_2\, X(t)\, dW(t), \quad X(0) = 1, \tag{4.3}$$

ETNA
Kent State University
etna@mcs.kent.edu

Time step	Ia ϵ	ratio	Ib ϵ	ratio
0.125	0.00126	*	0.03234	*
0.0625	0.000593	2.1	0.02374	1.4
0.03125	0.000287	2.1	0.01175	2
0.015625	0.000128	2.2	0.00431	2.2

TABLE 4.1
Example Ia $\sigma_2 = 0.5$, *Ib* $\sigma_2 = 1,\ \sigma_i = 0,\ i = 1, 3$

Time step	IIa ϵ	ratio	IIb ϵ	ratio
0.125	0.00131	*	0.02284	*
0.0625	0.00057	2.3	0.01209	1.9
0.03125	0.000245	2.3	0.00497	2.4
0.015625	0.00011	2.2	0.002583	1.9

TABLE 4.2
Example IIa $\sigma_3 = 0.5$, *IIb* $\sigma_3 = 1,\ \sigma_i = 0,\ i = 1, 2$

Time step	IIIa ϵ	ratio	IIIb ϵ	ratio
0.125	0.00074	*	0.002776	*
0.0625	0.000185	4	0.000742	3.7
0.03125	5.189E-05	3.6	0.00018	4.1
0.015625	1.2097E-05	4.3	4.401E-05	4.1

TABLE 4.3
Example IIIa $\sigma_1 = 0.5$, *IIIb* $\sigma_1 = 1,\ \sigma_i = 0,\ i = 2, 3$

for $t \in [0, 160]$ and $\sigma_2 = 0,\ 0.1,\ 0.3,\ 0.5,\ \Theta = 0.5$ for various trajectories of the Wiener process. A typical outcome is presented in Figures 4.4 and 4.5.

In our second example we have changed the value of q in

$$dX(t) \ = \{-X(t) + 0.9X(qt)\}\, dt + \sigma_2\, X(t)\, dW(t), \quad X(0) = 1, \tag{4.4}$$

from 0.1 to 0.9. In Figures 4.6 to 4.9 are shown representatives of typical trajectories. One cannot draw conclusions from a small set of sample trajectories, but (assuming the numerical solutions to be accurate) the results raise some interesting questions about the relative merits of deterministic and stochastic models of the real-life pantograph.

5. Proofs of Theorems 3.2 and 3.3. *Proof of Theorem 3.2*: For the SDDE (2.1) and the Θ-method (3.4) the local error $\delta_h(t_n, \zeta)$ takes the special form:

$$\begin{aligned} \delta_h(t_n,\zeta) = a \int_{t_n}^{t_n+\zeta h} & X(s) - (\Theta X(t_{n+1}) + (1-\Theta)X(t_n))ds \\ +b \int_{t_n}^{t_n+\zeta h} & X(qs) - (\Theta X(qt_{n+1}) + (1-\Theta)X(qt_n))\, ds \\ + \int_{t_n}^{t_n+\zeta h} & \sigma_2(X(s) - X(t_n)) + \sigma_3(X(qs) - X(qt_n))\, dW(s), \end{aligned} \tag{5.1}$$

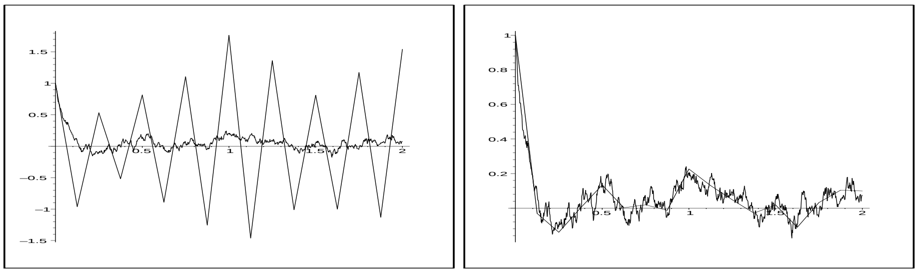

FIG. 4.1. *Example IV* $\sigma_1 = 0.5,\ \sigma_i = 0,\ i = 2, 3$ *left:* $\Theta = 0$*, right:* $\Theta = 0.5$

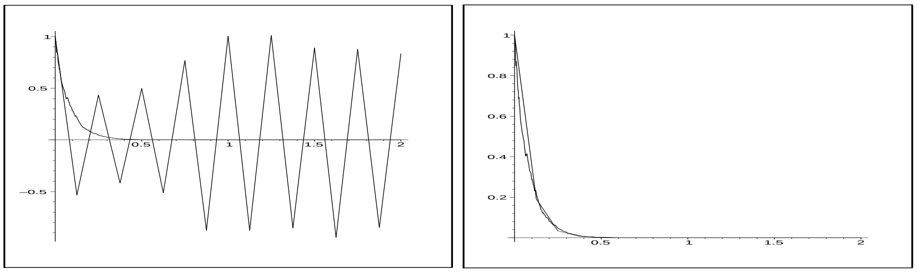

FIG. 4.2. *Example V* $\sigma_2 = 0.5,\ \sigma_i = 0,\ i = 1, 3$ *left:* $\Theta = 0$*, right:* $\Theta = 0.5$

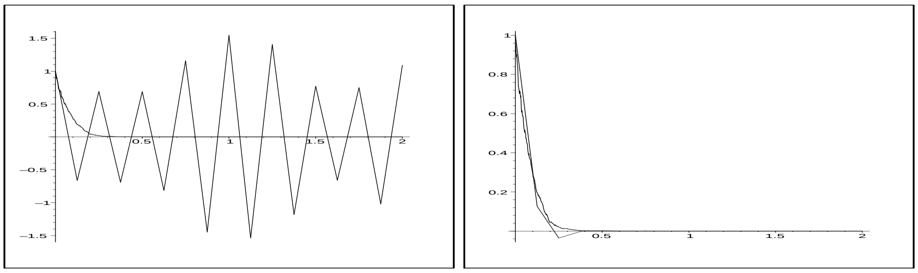

FIG. 4.3. *Example VI* $\sigma_3 = 0.5,\ \sigma_i = 0,\ i = 1, 2$ *left:* $\Theta = 0$*, right:* $\Theta = 0.5$

for $n = 0, \ldots, N-1$. We will frequently make use of the fact that for all $0 \le u \le t \le T$ the equation

$$X(t) - X(0) = \int_0^t aX(s) + bX(qs)\,ds + \int_0^t \sigma_1 + \sigma_2 X(s) + \sigma_3 X(qs)\,dW(s)$$

is equivalent to

$$X(t) - X(u) = \int_u^t aX(s) + bX(qs)\,ds + \int_u^t \sigma_1 + \sigma_2 X(s) + \sigma_3 X(qs)\,dW(s). \tag{5.2}$$

First we prove consistency in the mean with order 2. We have

$$|\mathcal{E}(\delta_h(t_n, \zeta))| \le |a|\;|\mathcal{E}\Big(\int_{t_n}^{t_n+\zeta h} X(s) - X(t_n) + \Theta(X(t_{n+1}) - X(t_n))\,ds\Big)|$$

ETNA
Kent State University
etna@mcs.kent.edu

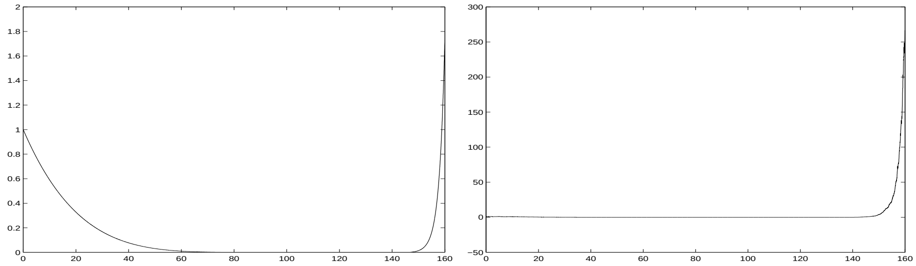

FIG. 4.4. *A path of the solution of (4.3) with (left) $\sigma_2 = 0$ and (right) $\sigma_2 = 0.1$.*

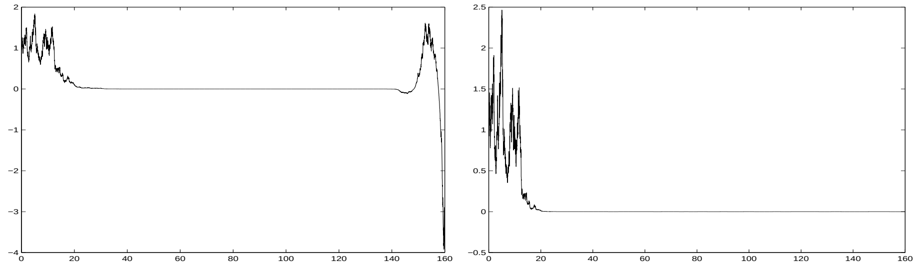

FIG. 4.5. *A path of the solution of (4.3) with (left) $\sigma_2 = 0.3$ and (right) $\sigma_2 = 0.5$.*

$$+ |b|\ |\mathcal{E}\Big(\int_{t_n}^{t_n+\zeta h} X(qs) - X(qt_n) + \Theta(X(qt_{n+1}) - X(qt_n))\ ds\Big)|.$$

Invoking (5.2), as well as Lemma 2.4, we obtain

$$\begin{aligned}
|\mathcal{E}(\delta_h(t_n,\zeta))| &\le |a|\ \mathcal{E}\Big\{ \int_{t_n}^{t_n+\zeta h} \int_{t_n}^{s} |aX(u) + bX(qu)|\ du\ \ ds\Big\} \\
&+ |a|\ \Theta\ \mathcal{E}\Big\{ \int_{t_n}^{t_n+\zeta h} \int_{t_n}^{t_{n+1}} |aX(u) + bX(qu)|\ du\ ds\Big\} \\
&+ |b|\ \mathcal{E}\Big\{ \int_{t_n}^{t_n+\zeta h} \int_{qt_n}^{qs} |aX(u) + bX(qu)|\ du\ ds\Big\} \\
&+ |b|\ \Theta\ \mathcal{E}\Big\{ \int_{t_n}^{t_n+\zeta h} \int_{qt_n}^{qt_{n+1}} |aX(u) + bX(qu)|\ du\ ds\Big\} \\
&\le K\ \sqrt{2\ K\ C_1(T)} \Big\{ \int_{t_n}^{t_n+\zeta h} \int_{t_n}^{s} du\ ds + \Theta\ \int_{t_n}^{t_n+\zeta h} \int_{t_n}^{t_{n+1}} du\ ds \\
&+ \int_{t_n}^{t_n+\zeta h} \int_{qt_n}^{qs} du\ ds + \Theta\ \int_{t_n}^{t_n+\zeta h} \int_{qt_n}^{qt_{n+1}} du\ ds\Big\} \\
&= \ K\ \sqrt{2\ K\ C_1(T)}\big(\frac{\zeta^2}{2}(1+q) + \zeta\ \Theta(1+q)\big)\ h^2,
\end{aligned}$$

ETNA
Kent State University
etna@mcs.kent.edu

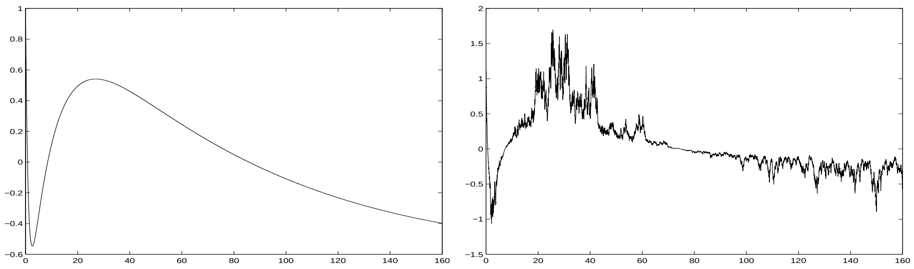

FIG. 4.6. *A path of the solution of (4.4) with* $q = 0.1$ *and (left)* $\sigma_2 = 0$ *and (right)* $\sigma_2 = 0.5$.

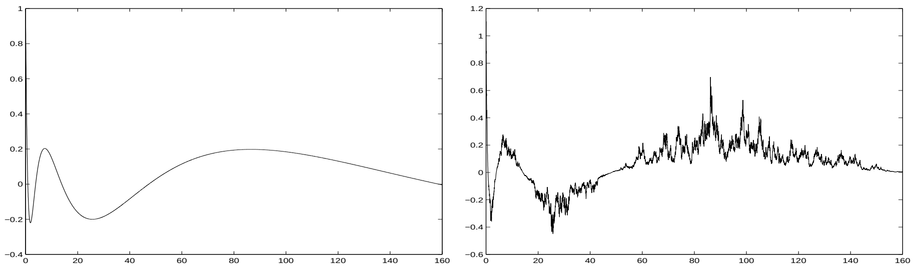

FIG. 4.7. *A path of the solution of (4.4) with* $q = 0.3$ *and (left)* $\sigma_2 = 0$ *and (right)* $\sigma_2 = 0.5$.

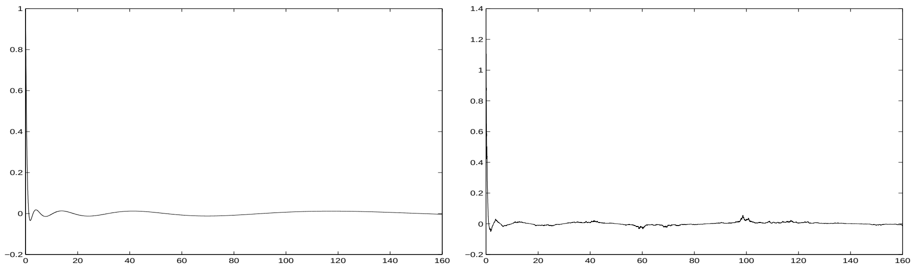

FIG. 4.8. *A path of the solution of (4.4) with* $q = 0.6$ *and (left)* $\sigma_2 = 0$ *and (right)* $\sigma_2 = 0.5$.

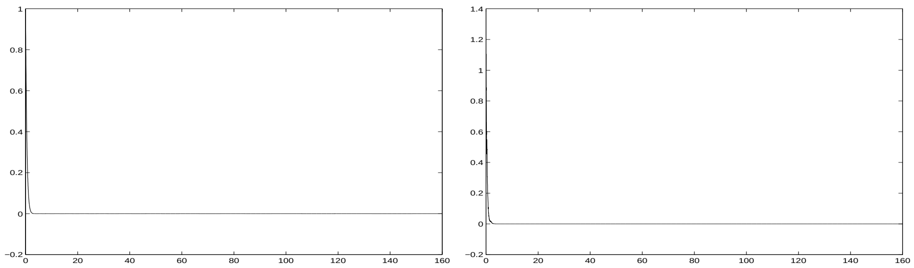

FIG. 4.9. *A path of the solution of (4.4) with* $q = 0.9$ *and (left)* $\sigma_2 = 0$ *and (right)* $\sigma_2 = 0.5$.

with $C_1(T)$ and K given in (2.4). With

$$C_3(T,\zeta) := K\ \sqrt{2\ K\ C_1(T)}(\frac{\zeta^2}{2}(1+q) + \zeta\ \Theta(1+q))$$

ETNA
Kent State University
etna@mcs.kent.edu

the preceding calculations can be expressed as

$$|\mathcal{E}(\delta_h(t_n,\zeta))| \;\le\; C_3(T,\zeta)\, h^2, \tag{5.3}$$

yielding

$$\sup_{\zeta\in(0,1]} |\mathcal{E}(\delta_h(t_n,\zeta))| \;\le\; C_3(T,1)\, h^2,$$

which proves (a). Now we prove (b) consistency in the mean-square with order 1. For ease of exposition we abbreviate

$$\begin{aligned} f(s) &= a\,(X(s)-X(t_n)-\Theta(X(t_{n+1})-X(t_n)))\\ &\quad +b\,(X(qs)-X(qt_n)-\Theta(X(qt_{n+1})-X(qt_n))),\\ g(s) &= \sigma_2\,(X(s)-X(t_n)) \;+\; \sigma_3\,(X(qs)-X(qt_n)). \end{aligned}$$

We use the Hölder inequality, Schwartz inequality for integrals, $2ab \le a^2+b^2$, $(a+b)^2 \le 2(a^2+b^2)$ and (2.4). We have:

$$\begin{aligned} \mathcal{E}|\delta_h(t_n,\zeta)|^2 &\le \mathcal{E}\left(\int_{t_n}^{t_n+\zeta h} |f(s)|\,ds\right)^2 \;+\; \mathcal{E}\left(\int_{t_n}^{t_n+\zeta h} |g(s)|\,dW(s)\right)^2\\ &\quad +2\,\mathcal{E}\left(\int_{t_n}^{t_n+\zeta h} |f(s)|\,ds\right)\times\left(\int_{t_n}^{t_n+\zeta h} |g(s)|\,dW(s)\right)\\ &\le \mathcal{E}\left(\int_{t_n}^{t_n+\zeta h} |f(s)|\,ds\right)^2 \;+\; \int_{t_n}^{t_n+\zeta h} \mathcal{E}|g(s)|^2\,ds\\ &\quad +2\left(\mathcal{E}\left(\int_{t_n}^{t_n+\zeta h} |f(s)|\,ds\right)^2\right)^{1/2}\times\left(\int_{t_n}^{t_n+\zeta h} \mathcal{E}|g(s)|^2\,ds\right)^{1/2}\\ &\le 2\,\mathcal{E}\Big(\int_{t_n}^{t_n+\zeta h} |f(s)|\,ds\Big)^2 \;+\; 2\int_{t_n}^{t_n+\zeta h} \mathcal{E}|g(s)|^2\,ds\\ &\le 2\,h\int_{t_n}^{t_n+\zeta h} \mathcal{E}|f(s)|^2\,ds \;+\; 2\int_{t_n}^{t_n+\zeta h} \mathcal{E}|g(s)|^2\,ds. \end{aligned}$$

By setting $H=\max(|a|^2,|b|^2,|\sigma_2|^2,|\sigma_3|^2)$ and replacing $f(s)$ and $g(s)$, we obtain

$$\begin{aligned} \mathcal{E}|\delta_h(t_n,\zeta)|^2 &\le H\Big\{8\,h\int_{t_n}^{t_n+\zeta h} \mathcal{E}\Big(|X(s)-X(t_n)|^2\Big) \;+\Theta^2\mathcal{E}\Big(|X(t_{n+1})-X(t_n)|^2\Big)\\ &\quad +\mathcal{E}\Big(|X(qs)-X(qt_n)|^2\Big)+\Theta^2\mathcal{E}\Big(|X(qt_{n+1})-X(qt_n)|^2\Big)\,ds\\ &\quad +4\int_{t_n}^{t_n+\zeta h} \mathcal{E}\Big(|X(s)-X(t_n)|^2\Big) \;+\; \mathcal{E}\Big(|X(qs)-X(qt_n)|^2\Big)\,ds\Big\}\\ &\le H\Big\{8\,h\int_{t_n}^{t_n+\zeta h} C_2(T)(s-t_n) \;+\; \Theta^2 C_2(T)(t_{n+1}-t_n) \end{aligned}$$

ETNA
Kent State University
etna@mcs.kent.edu

$$+ C_2(T)q(s-t_n) \;+\; \Theta^2 q C_2(T)(t_{n+1}-t_n)\, ds$$

$$+ 4 \int_{t_n}^{t_n+\zeta h} C_2(T)(s-t_n) + \; q\, C_2(T)(s-t_n)\, ds \Big\}$$

$$\leq 8\, C_2\,(T)\; H(1+q)(\zeta^2 + \Theta^2\, \zeta)\; h^2.$$

With $C_4(T,\zeta) := 8\, C_2\,(T)\; H(1+q)(\zeta^2+\Theta^2\,\zeta)$ we have

$$\mathcal{E}|\delta_h(t_n,\zeta)|^2 \;\leq\; C_4(T,\zeta)\; h^2, \tag{5.4}$$

which implies that

$$\sup_{\zeta\in(0,1]} (\mathcal{E}|\delta_h(t_n,\zeta)|^2)^{\frac{1}{2}} \;\leq\; \sqrt{C_4(T,1)}h. \square$$

Proof of Theorem 3.3:
For 1. we follow [16]. The recurrence relation (3.4b) has a unique solution if and only if

$$\Theta(a + b((n+1)q-n)_+)h \;\neq\; 1, \qquad n = 0,1,...,$$

where $(x)_+ = 0$ if $x \leq 0$ and $(x)_+ = x$ if $x > 0$. It is then not difficult to show, that (3.4b) has a unique solution, when $h < h^*$.

2. By using (5.2) and (3.5), adding and subtracting

$$\int_{t_n}^{t_n+\zeta h} a(\Theta X(t_{n+1}) + (1-\Theta)X(t_n)) + b(\Theta X(qt_{n+1}) + (1-\Theta)X(qt_n))\, ds$$
$$+ \int_{t_n}^{t_n+\zeta h} \sigma_1 + \sigma_2 X(t_n) + \sigma_3 X(qt_n)\, dW(s)$$

and rearranging, we obtain for any point $s_\ell \in \mathcal{S}_{2N}$ with $t_n < s_\ell \leq t_{n+1}$, using the representation (3.3)

$$\begin{aligned} \epsilon(t_n+\zeta h) &= X(t_n+\zeta h) \;-\; \widetilde{X}(t_n+\zeta h) \\ &= \epsilon(t_n) + \delta_h(t_n,\zeta) + U_h(t_n,\zeta), \end{aligned} \tag{5.5}$$

where

$$\begin{aligned} U_h(t_n,\zeta) := &\int_{t_n}^{t_n+\zeta h} a\,(\Theta(X(t_{n+1}) - \widetilde{X}(t_{n+1}) + (1-\Theta)(X(t_n) - \widetilde{X}(t_n)))\, ds \\ &+ \int_{t_n}^{t_n+\zeta h} b\,(\Theta(X(qt_{n+1}) - \widetilde{X}(qt_{n+1}) + (1-\Theta)(X(qt_n) - \widetilde{X}(qt_n)))\, ds \\ &+ \int_{t_n}^{t_n+\zeta h} \sigma_2(X(t_n) - \widetilde{X}(t_n)) + \sigma_3(X(qt_n) - \widetilde{X}(qt_n))\, dW(s). \end{aligned} \tag{5.6}$$

Thus squaring, employing the conditional mean with respect to the σ-algebra $\mathcal{A}_0$ and taking absolute values, we obtain

$$\mathcal{E}(|\epsilon(t_n+\zeta h)|^2|\mathcal{A}_0) \leq \mathcal{E}(|\epsilon(t_n)|^2|\mathcal{A}_0) + \underbrace{\mathcal{E}(|\delta_h(t_n,\zeta)|^2|\mathcal{A}_0)}_{1)} + \underbrace{\mathcal{E}(|U_h(t_n,\zeta)|^2|\mathcal{A}_0)}_{2)}$$

ETNA
Kent State University
etna@mcs.kent.edu

$$+\underbrace{2\,|\mathcal{E}(\delta_h(t_n,\zeta)\cdot\epsilon(t_n)\,|\mathcal{A}_0)|}_{3)}+\underbrace{2\,|\mathcal{E}(\delta_h(t_n,\zeta)\cdot U_h(t_n,\zeta)\,|\mathcal{A}_0)|}_{4)}$$
$$+\underbrace{2\,|\mathcal{E}(\epsilon(t_n)\cdot U_h(t_n,\zeta)\,|\mathcal{A}_0)|}_{5)}, \tag{5.7}$$

which holds almost surely.

We will now estimate the separate terms in (5.7) individually and in sequence; all the estimates hold almost surely. We will frequently use the Hölder inequality, the inequalities $2\,ab \le a^2+b^2$ and $(a+b)^2 \le 2a^2+2b^2$ and properties of conditional expectation. In the calculations the generic constant $C(T,\zeta)$ appears frequently and may have different values.

• For the term labelled 1) in (5.7) we have by (5.4)

$$\mathcal{E}(\,|\delta_h(t_n,\zeta)|^2|\mathcal{A}_0)\;=\;\mathcal{E}(\mathcal{E}(\,|\delta_h(t_n,\zeta)|^2|\mathcal{A}_{t_n})|\mathcal{A}_0)\;\le\;C(T,\zeta)\,h^2.$$

• For the term labelled 2) in (5.7) we have

$$\begin{aligned}
&\mathcal{E}(|U_h(t_n,\zeta)|^2|\mathcal{A}_0)\\
\le\;& 2\,\mathcal{E}\Big(|\int_{t_n}^{t_n+\zeta h} a\,(\Theta(X(t_{n+1})-\widetilde{X}(t_{n+1}))+(1-\Theta)(X(t_n)-\widetilde{X}(t_n)))\\
&+b\,(\Theta\,(X(qt_{n+1})-\widetilde{X}(qt_{n+1})+(1-\Theta)(X(qt_n)-\widetilde{X}(qt_n)))\,ds|^2|\mathcal{A}_0\Big)\\
&+2\,\mathcal{E}\Big(|\int_{t_n}^{t_n+\zeta h}\sigma_2(X(t_n)-\widetilde{X}(t_n))+\sigma_3(X(qt_n)-\widetilde{X}(qt_n))\,dW(s)|^2|\mathcal{A}_0\Big)\\
\le\;& 8\,\zeta\,h\,\Big\{\int_{t_n}^{t_n+\zeta h}|a|^2\,\Theta^2\mathcal{E}(|X(t_{n+1})-\widetilde{X}(t_{n+1})|^2|\mathcal{A}_0)\,ds\\
&+\int_{t_n}^{t_n+\zeta h}|a|^2\,(1-\Theta)^2\mathcal{E}(|X(t_n)-\widetilde{X}(t_n)|^2|\mathcal{A}_0)\,ds\\
&+\int_{t_n}^{t_n+\zeta h}|b|^2\,\Theta^2\mathcal{E}(|X(qt_{n+1})-\widetilde{X}(qt_{n+1})|^2|\mathcal{A}_0)\,ds\\
&+\int_{t_n}^{t_n+\zeta h}|b|^2\,(1-\Theta)^2\mathcal{E}(|X(qt_n)-\widetilde{X}(qt_n)|^2|\mathcal{A}_0)\,ds\Big\}\\
&+4\int_{t_n}^{t_n+\zeta h}|\sigma_2|^2\,\mathcal{E}(|X(t_n)-\widetilde{X}(t_n)|^2|\mathcal{A}_0)+|\sigma_3|^2\,\mathcal{E}(|X(qt_n)-\widetilde{X}(qt_n)|^2|\mathcal{A}_0)\,ds\\
=\;& 8\,\zeta^2\,h^2\,\Big\{|a|^2\,\Theta^2\mathcal{E}(|\epsilon(t_{n+1})|^2|\mathcal{A}_0)+|a|^2\,(1-\Theta)^2\mathcal{E}(|\epsilon(t_n)|^2|\mathcal{A}_0)\\
&+|b|^2\,\Theta^2\mathcal{E}(|\epsilon(qt_{n+1})|^2|\mathcal{A}_0)+|b|^2\,(1-\Theta)^2\mathcal{E}(|\epsilon(qt_n)|^2|\mathcal{A}_0)\Big\}\\
&+4\,\zeta\,h\,\Big\{|\sigma_2|^2\,\mathcal{E}(|\epsilon(t_n)|^2|\mathcal{A}_0)+|\sigma_3|^2\,\mathcal{E}(|\epsilon(qt_n)|^2|\mathcal{A}_0)\Big\}.
\end{aligned}$$

• For the term labelled 3) we have, due to properties of conditional expectation and (5.3)

$$\begin{aligned}
2\,|\mathcal{E}(\delta_h(t_n,\zeta)\cdot\epsilon(t_n)\,|\mathcal{A}_0)| &\le 2\,\mathcal{E}(|\mathcal{E}(\delta_h(t_n,\zeta)\,|\mathcal{A}_{t_n})|\cdot|\epsilon(t_n)|\,|\mathcal{A}_0)\\
&\le 2\,\left(\mathcal{E}(|\mathcal{E}(\delta_h(t_n,\zeta)\,|\mathcal{A}_{t_n})|^2)\right)^{\frac12}\cdot\left(\mathcal{E}(|\epsilon(t_n)|^2|\mathcal{A}_0)\right)^{\frac12}\\
&\le 2\left(\mathcal{E}(C(T,\zeta)\,h^2)^2\right)^{\frac12}\cdot\left(\mathcal{E}(|\epsilon(t_n)|^2|\mathcal{A}_0)\right)^{\frac12}\\
&= 2\left(\mathcal{E}(C(T,\zeta)\,h^3)\right)^{\frac12}\cdot\left(h\,\mathcal{E}(|\epsilon(t_n)|^2|\mathcal{A}_0)\right)^{\frac12}
\end{aligned}$$

$$\leq\ C(T,\zeta)\,h^3 + h\,\mathcal{E}(|\epsilon(t_n)|^2|\mathcal{A}_0).$$

• For the term labelled 4) in (5.7) we obtain, using (5.4) and the calculations for term 2)

$$
\begin{aligned}
&2\,|\mathcal{E}(\delta_h(t_n,\zeta)\cdot U_h(t_n,\zeta)\,|\mathcal{A}_0)|\\
&\leq 2\left(\mathcal{E}(|\delta_h(t_n,\zeta)|^2|\mathcal{A}_0)\right)^{\frac12}\left(\mathcal{E}(|U_h(t_n,\zeta)|^2|\mathcal{A}_0)\right)^{\frac12}\\
&\leq \mathcal{E}(\mathcal{E}(|\delta_h(t_n,\zeta)|^2|\mathcal{A}_{t_n})|\mathcal{A}_0)+\mathcal{E}(|U_h(t_n,\zeta)|^2|\mathcal{A}_0)\\
&\leq C(T,\zeta)\,h^2+8\,\zeta^2\,h^2\,\Big\{|a|^2\,\Theta^2\mathcal{E}(|\epsilon(t_{n+1})|^2|\mathcal{A}_0)+|a|^2\,(1-\Theta)^2\mathcal{E}(|\epsilon(t_n)|^2|\mathcal{A}_0)\\
&+|b|^2\,\Theta^2\mathcal{E}(|\epsilon(qt_{n+1})|^2|\mathcal{A}_0)+|b|^2\,(1-\Theta)^2\mathcal{E}(|\epsilon(qt_n)|^2|\mathcal{A}_0)\Big\}\\
&+4\,\zeta\,h\,\Big\{|\sigma_2|^2\,\mathcal{E}(|\epsilon(t_n)|^2|\mathcal{A}_0)+|\sigma_3|^2\,\mathcal{E}(|\epsilon(qt_n)|^2|\mathcal{A}_0)\Big\}.
\end{aligned}
$$

• For the term labelled 5) in (5.7) we have

$$
\begin{aligned}
&2\,|\mathcal{E}(\epsilon(t_n)\cdot U_h(t_n,\zeta)\,|\mathcal{A}_0)|\\
&\leq 2\,\mathcal{E}(|\mathcal{E}(U_h(t_n,\zeta)\,|\mathcal{A}_{t_n})|\cdot|\epsilon(t_n)|\,|\mathcal{A}_0)\\
&\leq\ 2\,\zeta\,h\,\Big\{|a|\,\Theta\,\mathcal{E}(|\epsilon(t_{n+1})|\cdot|\epsilon(t_n)|\,|\mathcal{A}_0)\ +\ |a|\,(1-\Theta)\,\mathcal{E}(|\epsilon(t_n)|^2\,|\mathcal{A}_0)\\
&\quad+|b|\,\Theta\,\mathcal{E}(|\epsilon(qt_{n+1})|\cdot|\epsilon(t_n)|\,|\mathcal{A}_0)\ +\ |b|\,(1-\Theta)\,\mathcal{E}(|\epsilon(qt_n)|\cdot|\epsilon(t_n)|\,|\mathcal{A}_0)\Big\}\\
&\leq |a|\,\Theta\,\zeta\,h\,\Big\{2\,\left(\mathcal{E}(|\epsilon(t_{n+1})|^2\,|\mathcal{A}_0)\right)^{\frac12}\cdot\left(\mathcal{E}(|\epsilon(t_n)|^2\,|\mathcal{A}_0)\right)^{\frac12}\Big\}\\
&\quad+2\,|a|\,(1-\Theta)\,\zeta\,h\,\mathcal{E}(|\epsilon(t_n)|^2|\mathcal{A}_0)\\
&\quad+|b|\,\Theta\,\zeta\,h\,\Big\{2\,\left(\mathcal{E}(|\epsilon(qt_{n+1})|^2\,|\mathcal{A}_0)\right)^{\frac12}\cdot\left(\mathcal{E}(|\epsilon(t_n)|^2\,|\mathcal{A}_0)\right)^{\frac12}\Big\}\\
&\quad+|b|\,(1-\Theta)\,\zeta\,h\,\Big\{2\,\left(\mathcal{E}(|\epsilon(qt_n)|^2\,|\mathcal{A}_0)\right)^{\frac12}\cdot\left(\mathcal{E}(|\epsilon(t_n)|^2\,|\mathcal{A}_0)\right)^{\frac12}\Big\}\\
&\leq |a|\,\Theta\,\zeta\,h\,\mathcal{E}(|\epsilon(t_{n+1})|^2\,|\mathcal{A}_0)\\
&\quad+\zeta\,h\,(|a|\,\Theta+2\,|a|\,(1-\Theta)+|b|\,\Theta+|b|\,(1-\Theta))\mathcal{E}(|\epsilon(t_n)|^2|\mathcal{A}_0)\\
&\quad+\ |b|\,\Theta\,\zeta\,h\,\mathcal{E}(|\epsilon(qt_{n+1})|^2\,|\mathcal{A}_0)\ +\ |b|\,(1-\Theta)\,\zeta\,h\,\mathcal{E}(|\epsilon(qt_n)|^2\,|\mathcal{A}_0).
\end{aligned}
$$

Combining these results, we obtain

$$
\begin{aligned}
&\mathcal{E}(|\epsilon(t_n+\zeta h)|^2|\mathcal{A}_0)\\
&\leq(16\,\zeta^2\,h^2\,|a|^2\,\Theta^2+|a|\,\Theta\,\zeta\,h)\,\mathcal{E}(|\epsilon(t_{n+1})|^2\,|\mathcal{A}_0)\\
&+(1+16\,\zeta^2\,h^2\,|a|^2\,(1-\Theta)^2+8\,\zeta\,h\,|\sigma_2|^2+h\\
&+h\,(|a|\,\Theta+2\,|a|\,(1-\Theta)+|b|\,\Theta+|b|\,(1-\Theta)))\mathcal{E}(|\epsilon(t_n)|^2|\mathcal{A}_0)\\
&+(16\,\zeta^2\,h^2\,|b|^2\,\Theta^2+|b|\,\Theta\,\zeta\,h)\,\mathcal{E}(|\epsilon(qt_{n+1})|^2\,|\mathcal{A}_0)
\end{aligned}
\tag{5.8}
$$

$$+\,(16\,\zeta^2\,h^2\,|b|^2\,(1-\Theta)^2+8\,\zeta\,h\,|\sigma_3|^2+|b|\,(1-\Theta)\,\zeta\,h)\mathcal{E}(|\epsilon(qt_n)|^2\,|\mathcal{A}_0)$$
$$+\,C(T,\zeta)\,h^2+C(T,\zeta)\,h^3.$$

As $h,\Theta<1$, $\zeta\in(0,1]$ and by setting $a^*=\max(|a|,|a|^2)$ and $b^*=\max(|b|,|b|^2)$, we can obtain from (5.8) the estimate

$$\begin{aligned}
&\mathcal{E}(|\epsilon(t_n+\zeta h)|^2|\mathcal{A}_0)\\
&\leq 17\,\zeta\,h\,a^*\,\Theta\,\mathcal{E}(|\epsilon(t_{n+1})|^2\,|\mathcal{A}_0)\\
&+\,(1+\,\zeta\,h\,(1+18\,a^*\,(1-\Theta)+8\,|\sigma_2|^2+\,a^*\,\Theta+b^*)\,)\mathcal{E}(|\epsilon(t_n)|^2|\mathcal{A}_0)\\
&+\,\zeta\,h\,17\,b^*\,\Theta\,\mathcal{E}(|\epsilon(qt_{n+1})|^2\,|\mathcal{A}_0)\\
&+\,\zeta\,h\,(17\,b^*\,(1-\Theta)+8\,|\sigma_3|^2)\,\mathcal{E}(|\epsilon(qt_n)|^2\,|\mathcal{A}_0)\;+\;C(T,\zeta)\,h^2.
\end{aligned}\tag{5.9}$$

Subsequent calculations involving the term $\mathcal{E}(|\epsilon(qt_{n+1})|^2\,|\mathcal{A}_0)$ depend on whether a) $t_n\leq qt_{n+1}\leq t_{n+1}$ or b) $qt_{n+1}<t_n$.

• Case a)

We set $t_n+\zeta h=qt_{n+1}$ on the left hand side of (5.9), solve for $\mathcal{E}(|\epsilon(qt_{n+1})|^2\,|\mathcal{A}_0)$, and, using that $\zeta\leq q$ in this case, we obtain for $0<q\,h\,17\,b^*\,\Theta<1$

$$\begin{aligned}
&\mathcal{E}(|\epsilon(qt_{n+1})|^2\,|\mathcal{A}_0)\\
&\leq\frac{1}{1-q\,h\,17\,b^*\,\Theta}\Big(17\,q\,h\,a^*\,\Theta\,\mathcal{E}(|\epsilon(t_{n+1})|^2\,|\mathcal{A}_0)\\
&+\,(1+\,q\,h\,(1+18\,a^*\,(1-\Theta)+8\,|\sigma_2|^2+\,a^*\,\Theta+b^*))\,\mathcal{E}(|\epsilon(t_n)|^2|\mathcal{A}_0)\\
&+\;q\,h\,(17\,b^*\,(1-\Theta)+8\,|\sigma_3|^2)\,\mathcal{E}(|\epsilon(qt_n)|^2\,|\mathcal{A}_0)\;+\;C(T,q)\,h^2\Big).
\end{aligned}\tag{5.10}$$

Returning to (5.9), we estimate $\mathcal{E}(|\epsilon(qt_{n+1})|^2\,|\mathcal{A}_0)$ on the right hand side by (5.10), which yields

$$\begin{aligned}
&\mathcal{E}(|\epsilon(t_n+\zeta h)|^2|\mathcal{A}_0)\\
&\leq 17\,\zeta\,h\,a^*\,\Theta\,\mathcal{E}(|\epsilon(t_{n+1})|^2\,|\mathcal{A}_0)\\
&+\,(1+\,\zeta\,h\,(1+18\,a^*\,(1-\Theta)+8\,|\sigma_2|^2+\,a^*\,\Theta+b^*))\,\mathcal{E}(|\epsilon(t_n)|^2|\mathcal{A}_0)\\
&+\,\zeta\,h\,\frac{17\,b^*\,\Theta}{1-q\,h\,17\,b^*\,\Theta}\Big\{17\,q\,h\,a^*\,\Theta\,\mathcal{E}(|\epsilon(t_{n+1})|^2\,|\mathcal{A}_0)\\
&+\,(1+q\,h\,(1+18\,a^*\,(1-\Theta)+8\,|\sigma_2|^2+\,a^*\,\Theta+b^*))\,\mathcal{E}(|\epsilon(t_n)|^2|\mathcal{A}_0)\\
&+\;q\,h\,(17\,b^*\,(1-\Theta)+8\,|\sigma_3|^2)\,\mathcal{E}(|\epsilon(qt_n)|^2\,|\mathcal{A}_0)\;+\;C(T,q)\,h^2\Big\}\\
&+\;\zeta\,h\,(17\,b^*\,(1-\Theta)+8\,|\sigma_3|^2)\,\mathcal{E}(|\epsilon(qt_n)|^2\,|\mathcal{A}_0)\;+\;C(T,\zeta)\,h^2.
\end{aligned}$$

We write

$$R_0=0\quad\text{and}\quad R_k=\sup_{\substack{0\leq i<k\\ \zeta\in(0,1]}}\mathcal{E}\Big(|\epsilon(t_i+\zeta h)|^2|\mathcal{A}_0\Big).\tag{5.11}$$

Thus

$$R_{n+1}\,\frac{1-h\,17\,\Theta(a^*+qb^*)}{1-q\,h\,17\,b^*\,\Theta}$$
$$\leq (1+\frac{17\,h\,b^*\,\Theta}{1-q\,h\,17\,b^*\,\Theta})\,R_n$$
$$+\,h\,\Big(1+\frac{17\,b^*\,\Theta\,q\,h}{1-q\,h\,17\,b^*\,\Theta}\Big)\Big((1+18\,a^*\,(1-\Theta)+8\,|\sigma_2|^2+\ a^*\,\Theta+b^*)\,R_n$$
$$+\ (17\,b^*\,(1-\Theta)+8\,|\sigma_3|^2)\,R_n\Big)+C(T,1)\,h^2.$$

Hence

$$R_{n+1}\,\frac{1-h\,17\,\Theta(a^*+qb^*)}{1-q\,h\,17\,b^*\,\Theta}$$
$$\leq \frac{1}{1-q\,h\,17\,b^*\,\Theta}\Big\{1+h\,\Big(1+18\,(a^*+b^*)-\ 17\,\Theta\,(a^*+qb^*)$$
$$+\,8\,(|\sigma_2|^2+|\sigma_3|^2)\Big)\Big\}R_n\ +\ C(T,1)\,h^2.$$

Provided that $0<h\,17\,\Theta\,(a^*+qb^*)<1$ we get

$$R_{n+1}$$
$$\leq \frac{1}{1-h\,17\,\Theta(a^*+qb^*)}$$
$$\Big\{1+h\,\Big(1+18(a^*+b^*)-17\Theta\,(a^*+qb^*)+8(|\sigma_2|^2+|\sigma_3|^2)\Big)\Big\}\,R_n \tag{5.12}$$
$$+\ \frac{1-q\,h\,17\,b^*\,\Theta}{1-h\,17\,\Theta(a^*+qb^*)}\,C(T,1)\,h^2$$
$$\leq \Big\{1+\ h\,\frac{1+18\,(a^*+b^*)+8\,(|\sigma_2|^2+|\sigma_3|^2)}{1-h\,17\,\Theta\,(a^*+qb^*)}\Big\}\,R_n+C(T,1)\,h^2.$$

• Case b)
We derive from (5.9) by using (5.11), when $0<17\,h\,a^*\,\Theta<1$, that

$$R_{n+1}\,(1-17\,h\,a^*\,\Theta)$$
$$\leq \Big(1+h(1+18(a^*+b^*)-17\,a^*\Theta+8\,(|\sigma_2|^2+|\sigma_3|^2)\Big)\,R_n+C(T,1)h^2,$$
$$R_{n+1}$$
$$\leq \frac{1+h(1+18(a^*+b^*)-17a^*\Theta+8(|\sigma_2|^2+|\sigma_3|^2)}{1-17\,h\,a^*\,\Theta}\,R_n$$
$$+\,C(T,1)\,h^2$$
$$\leq \Big\{1+h\,\frac{1+18\,(a^*+b^*)+8\,(|\sigma_2|^2+|\sigma_3|^2)}{1-17\,h\,a^*\,\Theta}\Big\}\,R_n+C(T,1)h^2. \tag{5.13}$$

ETNA
Kent State University
etna@mcs.kent.edu

Now we take h such that $h\ 17\ \Theta\ (a^* + qb^*) < L_1 < 1$ and $h\ 17\ a^*\ \Theta < L_2 < 1$ and set

$$M = \max_{i=1,2} \frac{1 + 18\,(a^* + b^*) + 8\,(|\sigma_2|^2 + |\sigma_3|^2)}{1 - L_i}.$$

Then we get from (5.12) and (5.13)

$$R_{n+1} \le (1 + h\ M)\ R_n\ +\ C(T,1)\ h^2,$$

and, by iterating and observing that $h(n+1) = t_{n+1} \le T$,

$$\begin{aligned} R_{n+1} &\le (1 + Mh)^{n+1} R_0 + C(T,1) h^2 \sum_{k=0}^{n} (1 + Mh)^k \\ &\le h\, \frac{C(T,1)}{M} ((1 + Mh)^{n+1} - 1)) \\ &\le h\, \frac{C(T,1)}{M} (e^{Mh(n+1)} - 1)\ \le\ h\, \frac{C(T,1)}{M} (e^{MT} - 1). \end{aligned}$$

From this the results of the theorem follow. □

REFERENCES

[1] L. ARNOLD, *Stochastic Differential Equations: Theory and Applications*, Wiley-Interscience, New York, 1974.
[2] C. T. H. BAKER, *Retarded differential equations*, J. Comput. Appl. Math., 125 (2000), pp. 309–335.
[3] C. T. H. BAKER AND E. BUCKWAR, *Introduction to the numerical analysis of stochastic delay differential equations*, LMS J. Comput. Math., 3 (2000), pp. 315–335.
[4] C. T. H. BAKER, C. A. H. PAUL AND D. WILLÉ, *Issues in the numerical solution of evolutionary delay differential equations*, Adv. Comput. Math., 3 (1995), pp. 171–196.
[5] H. BRUNNER, *On the discretization of differential and Volterra integral equations with variable delay*, Preprint, 1995.
[6] K. BURRAGE AND T. TIAN, *The composite Euler method for stiff stochastic differential equations*, Preprint, 1999. Available online from http://www.maths.uq.edu.au/~kb/papers.html .
[7] J. CARR AND J. DYSON, *The functional differential equation $y'(x) = ay(\lambda x) + by(x)$*, Proc. Roy. Soc. Edinburgh Sect. A, 74 (1974/75), pp. 165–174.
[8] J. M. C. CLARK, *The discretization of stochastic differential equations: a primer*, in Road Vehicle Systems and Related Mathematics; Proc. 2nd DMV-GAMM Workshop, Torino, H. Neunzert, ed., Teubner, Stuttgart, and Kluwer Academic Publishers, Amsterdam, 1987, pp. 163–179.
[9] J. G. GAINES, *A basis for iterated stochastic integrals*, Math. Comput. Simulation, 38 (1995), pp.7–11.
[10] 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), pp. 1455–1484.
[11] T. C. GARD, *Introduction to Stochastic Differential Equations*, Marcel-Dekker, New York and Basel, 1988.
[12] L. FOX, D. F. MAYERS, J. R. OCKENDON AND A. B. TAYLER, *On a functional differential equation*, J. Inst. Math. Appl., 8 (1971), pp. 271–307.
[13] A. ISERLES,*On the generalized pantograph functional-differential equation*, European J. Appl. Math., 4 (1993), pp. 1–38.
[14] A. ISERLES, *Exact and discretized stability of the pantograph equation*, DAMTP Technical Report 1996/NA10, University of Cambridge, 1996.
[15] T. KATO AND J. B. MCLEOD, *The functional-differential equation $y'(x) = ay(\lambda x) + by(x)$*, Bull. Amer. Math. Soc., 77 (1971), pp. 891–937.
[16] Y. LIU, *On functional differential equations with proportional delays*, Ph.D. thesis, Department of Mathematics, Cambridge University, Cambridge, 1996.
[17] Y. LIU, *Numerical investigation of the pantograph equation*, Appl. Numer. Math., 24 (1997), pp. 309–317.
[18] K. ITÔ AND M. NISIO, *On stationary solutions of a stochastic differential equation*, J. Math. Kyoto Univ., 4 (1964), pp. 1–75.
[19] I. KARATZAS AND S. E. SHREVE, *Brownian Motion and Stochastic Calculus*, Springer-Verlag, New York, 1991.

[20] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag, Berlin, 1992.
[21] X. Mao, *Stochastic Differential Equations and Their Applications*, Horwood Publishing Limited, Chichester, 1997.
[22] G. N. Milstein, *Numerical Integration of Stochastic Differential Equations*, Kluwer Academic Publishers Group, Dordrecht, 1995, translated and revised from the 1988 Russian original.
[23] V. J. Mizel and V. Trutzer, *Stochastic hereditary equations*: *existence and asymptotic stability*, J. Integral Equations, 7 (1984), pp. 1–72.
[24] S. E. A. Mohammed, *Stochastic differential systems with memory. Theory, examples and applications*, in Stochastic Analysis and Related Topics VI, The Geilo Workshop, 1996, L. Decreusefond, Jon Gjerde, B. Oksendal and A.S. Ustunel, eds., Progress in Probability, Birkhauser, 1998, pp. 1-77.
[25] S. E. A. Mohammed, *Stochastic Functional Differential Equations*, Pitman, Boston, Mass., 1984.
[26] J. R. Ockendon and A. B. Tayler, *The dynamics of a current collection system for an electric locomotive*, Proc. Roy. Soc. London Ser. A, 322 (1971), pp. 447–468.
[27] E. Platen, *An introduction to numerical methods for stochastic differential equations*, Acta Numerica, 8 (1999), pp. 195–244.
[28] D. Talay, *Simulation of stochastic differential systems*, in Probabilistic Methods in Applied Physics, P. Krée and W. Wedig, eds., Lecture Notes in Physics 451, Springer, Berlin, 1995, pp. 54–96.
[29] D. Williams, *Probability with Martingales*, Cambridge University Press, Cambridge, 1991.
[30] M. Zennaro, *Delay differential equations*: *theory and numerics*, in Theory and Numerics of Ordinary and Partial Differential Equations, M. Ainsworth, J. Levesley, W. A. Light and M. Marletta, eds., Oxford University Press, New York, 1995, pp. 291–333.