Stochastic Hydrol. Hydraul. 3 (1989) 155–178

Stochastic Hydrology and Hydraulics

Review

A survey of numerical methods for stochastic differential equations

P. E. Kloeden

School of Mathematical and Physical Sciences, Murdoch University, Murdoch 6150, Western Australia

E. Platen

Karl-Weierstrass-Institut für Mathematik, Akademie der Wissenschaften der DDR, Mohrenstraße 39, DDR-1086 Berlin

Abstract: The development of numerical methods for stochastic differential equations has intensified over the past decade. The earliest methods were usually heuristic adaptations of deterministic methods, but were found to have limited accuracy regardless of the order of the original scheme. A stochastic counterpart of the Taylor formula now provides a framework for the systematic investigation of numerical methods for stochastic differential equations. It suggests numerical schemes, which involve multiple stochastic integrals, of higher order of convergence. We shall survey the literature on these and on the earlier schemes in this paper. Our discussion will focus on diffusion processes, but we shall also indicate the extensions needed to handle processes with jump components. In particular, we shall classify the schemes according to strong or weak convergence criteria, depending on whether the approximation of the sample paths or of the probability distribution is of main interest.

Key words: Stochastic differential equations, Stochastic Taylor formula, Numerical methods, Simulations, Strong convergence, Weak convergence.

1 Introduction

An Ito stochastic process $X=\{X_t;t\geq 0\}$ with jump component has the form

$$X_t=X_0+\int_0^t a(X_s)ds+\int_0^t b(X_s)dW_s+\int_0^t\int_U c(X_s,u)M(du,ds) \tag{1}$$

for $t\in[0,T]$. See Ito (1951). It consists of an initial value $X_0=x_0$, which may be random, a slowly varying continuous component called the drift, a rapidly varying continuous random component called the diffusion, and a jump component representing shot noise or discontinuous noise. The second integral in (1) is an Ito stochastic integral with respect to the Wiener process $W=\{W_t;0\leq t\leq T\}$ and the third is defined with respect to a Poissonian martingale measure M. The integral equation (1) is often written in differential form and is then called an Ito stochastic differential equation. When the jump component in (1) is absent the resulting process is a diffusion process, called an Ito diffusion.

Stochastic differential equations, both with and without jumps, are being used to model a rapidly expanding variety of random dynamical processes. Of particular interest to engineers are applications to hydrology (eg Unny (1984); Unny and Karmeshu (1983)), structural engineering and seismology (eg Kozin (1977); Shinozuka and Sato (1967)), biological waste treatment (eg Harris (1975); Michail et al. (1987)), fatigue

cracking (eg Sobcyzk (1986)), turbulence (eg Bywater and Chang (1973); Haworth and Pope (1986); Yaglom (1980)), satellite dynamics (eg Balakrishnan (1985); Sagirow (1970)), helicopter stability (eg Pardoux and Pignol (1986)), telecommunication (Viterbi (1965)), filtering (eg Kallianpur (1980)), and control (eg Krylov (1980)). They are widely used in the basic sciences such as astronomy (eg Legland (1981)), chemistry and physics (eg Horsthemke and Lefevre (1984); Van Kampen (1981)), genetics (eg Kimura and Ohta (1971)) and population dynamics (eg Gard (1988)). They are also starting to be used in the social sciences, such as economics (eg Karatzas and Shreve (1988); Merton (1971)), and experimental psychology (eg Schoener et al. (1986)). In addition stochastic differential equations have been used in essentially nonstochastic problems, for example in stochastic annealing (eg Geman and Hwang (1986)) where diffusive noise of small intensity is added to the gradient drift of a function for which the global minimum is sought.

Unfortunately explicit solutions of Ito stochastic differential equations are rare in practical applications. There are however a number of papers which deal with numerical methods for these equations. Nevertheless there exists a wide gap between the well developed theory of stochastic differential equations and its application. The crucial task in bridging this gap is the development of efficient numerical methods, which obviously should be implementable on modern digital computers. In view of the need to simulate a large number of different sample paths in order to estimate various statistical features of the solutions, vector or super computers will play an rapidly increasing role here.

In this paper we shall review time discretized numerical methods which are appropriate for the simulation of Ito processes or functionals of Ito processes on digital computers. First we briefly look at Ito stochastic differential equations and discuss a variety of different approaches that have been proposed for their numerical solution. Then we consider strong and weak convergence criteria for time discretized methods and introduce the stochastic Taylor formula. In the remainder of paper we present some basic strong and weak approximation methods for Ito stochastic differential equations, in particular truncated Taylor approximations and related Runge-Kutta methods. We do this firstly in some detail for Ito diffusions and then note the modifications needed to handle Ito processes with jumps. Finally we conclude with some remarks on how to choose and implement an appropriate numerical scheme. These and related matters will be discussed extensively in a forthcoming book Kloeden and Platen (1990).

2 Ito stochastic differential equations

We assume that the reader is familiar with the contents of the review article on Ito stochastic differential equations by Bodo et al. (1987), which appeared in this journal, or have a similar background knowledge. In addition we require the following notation. We denote by $C[0,T]$ the class of functions $f:[0,T]\to R^d$ which are continuous and by $D[0,T]$ those which are piecewise continuous, to be specific continuous from the right with finite limits from the left. Here $d=1,2,3,...$ corresponds to the dimension of the problem being discussed, and obviously $C[0,T]\subset D[0,T]$ for the same dimension d. The triplet (Ω, A, P) denotes the underlying probability space and consists of a sample space Ω of possibilities, a collection A of subsets of Ω called events, and a probability measure P which assigns to each event $A\in A$ a number $P(A)\in[0,1]$ called its probability.

Technically A is a σ-algebra of subsets and contains the empty set $\varnothing$ (nonevent) and the sample space Ω (sure event). Moreover if A is an event, so is its complement $A^c=\Omega-A$ and $P(A^c)=1-P(A)$. Events are those subsets of elementary events (or samples) that can be detected by their occurrence or nonoccurrence. Further we denote by $\mathbf{A}=\{A_t;t\in[0,T]\}$ a family of sub-σ-algebras of A which, in what follows, will be generated by the Wiener process under consideration. Essentially A_t consists of the events that can be detected, with respect to the Wiener process, by their occurrence or

nonoccurrence up to and including time t. We say that a stochastic process $X=\{X_t;t\in[0,T]\}$ is A-*adapted* if the random variable X_t is measurable (ie detectable) with respect to the σ-algebra of events A_t for each time instant $t\in[0,T]$.

The Ito process $X=\{X_t;t\in[0,T]\}$ in (1) with sample paths in $D[0,T]$ can be expressed equivalently in terms of the Ito stochastic differential equation:

$$dX_t=a(X_t)dt+b(X_t)dW_t+\int_U c(X_t,u)M(du,dt) \tag{2}$$

for $0\le t\le T$ with the initial condition (possible random)

$$X_0=x_0. \tag{3}$$

Here $U=R^r\setminus\{0\}$ for some integer $r=1,2,3,...$ and $W=\{W_t;t\in[0,T]\}$ is an m-dimensional standard A-adapted Wiener process with independent components $\{W_t^j\}_{j=1}^m$. In addition

$$M(du,dt)=p(du,dt)-\pi(du)dt \tag{4}$$

is a Poisson martingale measure on $U\times[0,T]$ which is defined as follows. For each Borel subset B of U the process $\{p(B,[0,t]);t\in[0,T]\}$ is a Poisson process with finite expectation

$$\mathrm{E}\,p(B,[0,t])=\int_0^t\int_B \pi(du)ds\le K<\infty \tag{5}$$

for all $0\le t\le T$, and is independent of the Wiener process W. Here $\pi(du)ds$ denotes the intensity measure which can be chosen in the standard form as

$$\pi(du)ds = duds/|u|^{r+1} \tag{6}$$

In (2) the drift coefficient $a(x)=\{a^i(x)\}_{i=1}^d$ and the jump coefficient $c(x,u)=\{c^i(x,u)\}_{i=1}^d$ are d-dimensional vectors and the diffusion coefficient $b(x)=\{b^j(x)\}_{j=1}^m$ is a $d\times m$ matrix with d-dimensional column vectors $b^j(x)$ for $j=1,2,...,m$. Time dependent drift and diffusion coefficients can be included in this formulation by considering the first component of $X_t=\{X_t^i\}_{i=1}^d$ as the time component $X_t^1=t$, which requires $a^1\equiv1$ and $c^1\equiv0$, $b^{1,j}\equiv0$ for $j=1,2,...,m$. Note that in the absence of jumps, that is when $c^i\equiv0$ for $i=1,2,...,d$, the Ito process satisfying (2) and (3) is a diffusion process with sample paths in $C[0,T]$.

An accessible textbook on this kind of stochastic differential equation is Arnold (1972). In addition see Gard (1988), which has a chapter on numerical methods for these equations. For the general case with jumps the standard references are Gikhmann and Skorokhod (1972) and Ikeda and Watanabe (1981). These books present complete definitions and a detailed investigation of the properties of Ito processes both with and without jumps, including existence and uniqueness theorems. Such theorems hold for instance if the coefficient functions a, b and c are Lipschitz continuous and satisfy a linear growth bound for large x. A mathematical statement of these conditions can also be found in Bodo et al. (1987).

3 Numerical approaches to stochastic differential equations

To begin we shall briefly mention some different approaches that have been suggested for the numerical solution of stochastic differential equations (SDEs). On the very general level there is a method due to Boyce (1978) by means of which one can investigate, in principle, general random systems numerically by Monte Carlo simulations. For SDEs this method is somewhat inefficient because it does not use the special structure of these equations, in particular that of the drift, diffusion and jump coefficients

Kushner (1977) (see also Kushner and Di Masi (1978)) proposed the discretization of both time and space variables, so the approximating processes are then discrete-time finite state Markov chains. These can be handled on digital computers via their probability transition matrices. In comparison with the information encompassed succinctly in the drift, diffusion and jump coefficients of an SDE, the transition matrices contain considerable superfluous information. Moreover all of this information must be processed when implementing numerical algorithms. Consequently this Markov chain approach seems to be applicable to low dimensional problems on bounded domains. Similar disadvantages, at least in higher dimensions, also arise when standard numerical methods are used to solve the parabolic integro-differential equations, such as the Fokker-Planck equation and its adjoint, related to functionals of the solution processes of SDEs. We shall not say any more about these methods here.

Dashevski and Liptser (1966) and Fahrmeier (1976) used analog computers to obtain solutions of Ito SDEs without jumps. The use of such computers is however restricted to those SDEs for which it is technically possible to construct the corresponding electronic circuits. Consequently they cannot be used in many important situation.

The most efficient and widely applicable approach to solving SDEs seems to be the simulation of sample paths of time discrete approximations on digital computers. This simulation approach is based on a finite discretization of the time interval $[0,T]$ under consideration and generates step by step approximate values to the sample paths of the solution process at the discretization times. These simulated sample paths can then be analyzed by the usual statistical methods to determine how good an approximation is, and in which sense it is close to the exact solution. Here the state variables are not discretized as in Kushner's Markov chain approach and the structure of the SDE as provided by the drift, diffusion and jump coefficients is fully used in a natural way. Simulation studies of, for example Pardoux and Talay (1985) and Liske and Platen (1987), show the efficiency of, in particular, higher order time discretization methods. An advantage of considerable practical importance with this approach is that computational costs such as CPU time and required memory increase only proportionally with the dimension of the problem. These methods are particularly suited to super or vector array computers, because the parallel structure allows the simultaneous computation of independent sample paths, so their use and development will be greatly stimulated by as these computers become more and more widespread. See Petersen (1987).

4 Time discrete approximations

Simulation studies and theoretical investigations by Clements and Anderson (1973), Clark and Cameron (1980), Fahrmeier (1976), Rumelin (1982), Wright (1974) and others show that not all heuristic time discrete approximations of an SDE (2.1) converge in a useful sense to the solution process as the maximum step size δ tends to zero. In particular these papers show that one cannot simply use some deterministic numerical method for ordinary differential equations, such as a higher-order Runge-Kutta method. Consequently a careful and systematic investigation of different methods is needed in order to choose a sufficiently efficient method for the task at hand.

For simplicity we shall consider an equidistant time discretization $(\tau)_\delta$ with

$$0=\tau_0<\tau_1<\ldots<\tau_n=T \tag{7}$$

of an interval $[0,T]$ with step size

$$\delta=T/n. \tag{8}$$

We note that one can use more general time discretizations, which may even be random. See Platen (1981) and Mikulevicius and Platen (1988). For instance they may include jump times of the Poisson jump measure p, as we will assume in Section 10. In any case

it is not necessary to use an equidistant time discretization, but usually a maximum step size δ must be specified.

The simplest heuristic time discrete approximation is the stochastic generalization of the Euler approximation, which is sometimes call the *Euler-Maruyama approximation.* See Maruyama (1955). It has the form

$$Y_{i+1}=Y_i+a(Y_i)\Delta_i+b(Y_i)\Delta W_i+\int_{\tau_i}^{\tau_{i+1}}\int_u c(Y_i,u)p(ds,du)-\int_u c(Y_i,u)\pi(du)\Delta_i \tag{9}$$

for $i=0,1,...,n-1$ with the initial value

$$Y_0=x_0, \tag{10}$$

where

$$\Delta_i=\tau_{i+1}-\tau_i \tag{11}$$

and

$$\Delta W_i=W_{\tau_{i+1}}-W_{\tau_i} \tag{12}$$

for $i=0,1,...,n-1$. Essentially it is formed by fixing the integrands in (9) to their values at the beginning of each discretization time interval. Obviously the recursive scheme (9) only gives values at the discretization times. If values are required at intermediate instants either piecewise constant values from the preceding discretization point or some interpolation, especially linear, of the values at the two immediate enclosing discretization times could be used.

The random variables ΔW_i (12) are independent m-dimensional random vector $\Delta W_i=\{\Delta W_i^j\}_{j=1}^m$ with independent $N(0,\Delta_i)$-normally distributed components, that is with means $\mathbf{E}\Delta W_i^j=0$ and variances $\mathbf{E}(\Delta W_i^j)^2=\Delta_i$ for $i=1,2,...,m$ and $j=1,2,...,n-1$. In simulations we can generate such random numbers from independent uniformly distributed random variables on [0,1]. The latter are usually provided by a pseudo-random number generator on a digital computer. These are generally adequate for the simulation of stochastic differential equations. See Ermakov (1975), Morgan (1984) or Rubinstein (1981). We mention two methods here for transforming a pair (U_1,U_2) of independent random variables uniformly distributed on [0,1] into pair (N_1,N_2) of independent standard $N(0,1)$-normally distributed random variables. The first is the Box-Mueller method which uses the transformations

$$N_1=(-2\log U_1)^{1/2}\cos(2\pi U_2),\quad N_2=(-2\log U_1)^{1/2}\sin(2\pi U_2). \tag{13}$$

The second is the Polar-Marsaglia method which avoids the time consuming computation of trigonometric functions. Instead it uses only those pairs (U_1,U_2) for which

$$W=(2U_1-1)^2+(2U_2-1)^2\le 1 \tag{14}$$

holds, with

$$N_1=(2U_1-1)\{(-2\log W)/W\}^{1/2},\quad N_2=(2U_2-1)\{(-2\log W)/W\}^{1/2} \tag{15}$$

While it discards the fraction $1-\pi/4$ of pairs (U_1,U_2), the Polar-Marsaglia method is often more efficient computationally than the Box-Mueller method when a large quantity of numbers is generated. Obviously, from a standard $N(0,1)$-normally distributed random variable N_s we obtain an $N(0,\Delta_i)$-normally distributed random variable ΔW_i^j by the simple transformation

$$\Delta W_i^j = N_s \Delta_i^{1/2} \tag{16}$$

by which means we can simulate the components ΔW_i^j for $j=1,2,...,m$ of the i-th increment of the Wiener process.

Let us now consider how to generate the Poisson jump measure p. The process $p=\{p(U,[0,t]);\ t\in[0,T]\}$ is a Poisson process with intensity $\pi(U)<\infty$. Since a Poisson process has independent increments, the time intervals between successive jumps are exponentially distributed with expectation $\mu=\pi(U)^{-1}$ and are independent of each other.

We can obtain an exponentially distributed random variable B_μ with mean μ from a random variable U_1 uniformly distributed on [0,1] by the inverse transformation method, which gives

$$B_\mu = -\mu \log U_1 \tag{17}$$

In this way we can generate a sequence of jump points $\{\sigma_1,\sigma_2,\ldots,\sigma_k,...\}$ of the Poisson process p. It remains then to generate the random marks $\{u_1,u_2,\ldots,u_k,...\}$ associated with these jumps. This is a sequence of independent identically distributed random variables with distribution

$$F_{u_k}(B)=P(u_k\in B)=\pi(B)/\pi(U) \tag{18}$$

for all Borel subsets B of U. If the set of marks U is discrete it is easy to generate a sequence of marks from random variables which are uniformly distributed on [0,1]. On the other hand, if the distribution function is continuous and invertible we can use the inverse transformation method and obtain an F_{u_k} distributed random variable u^* from a random variable U_1, which is uniformly distributed on [0,1], by

$$u^* = F_{u_k}^{-1}(U_1). \tag{19}$$

We combine the two methods when F_{u_k} consists of both a continuous part and a discrete part.

From the preceding hints it is easy to generate the increments of the Euler approximation (9). In the diffusion term we use the expression (10) for the increments of the components of the Wiener process and in the integral with respect to the Poisson jump measure we use the above jump times σ_k and marks u_k to obtain

$$\int_{\tau_i}^{\tau_{i+1}} \int_u c(Y_i,u)p(du,ds) = \sum_{k=1}^{\infty} c(Y_i,u_k)1_{\{\sigma_k\in(\tau_i,\tau_{i+1}]\}}\ , \tag{20}$$

where 1_A is the indicator function of the set A. The Euler approximation (9) with noise terms generated as described is then obviously a recursive algorithm for calculating approximate values of the Ito process at the specified discretization points.

Maruyama (1955) proved the mean-square convergence of the Euler approximation of the Ito process without jumps, this being one of the first papers on the approximation of Ito processes. The corresponding result for Ito processes with jumps can be found in Gikhmann and Skorokhod (1979). Results concerning the weak convergence, that is in distribution, of the Euler approximation are contained in Grigelonius and Mikulevicius (1981), Jacod and Shirayaev (1987), Platen and Rebolledo (1985) and in other papers.

As in deterministic numerical analysis, the order of convergence plays an important role in the construction of efficient stochastic algorithms. In the stochastic case there are however several different useful types of convergences. The choice of a convergence criterion depends on the character of the problem that is to be solved by a time discrete

simulation method, in particular on whether approximations to the sample paths or to the distributions are required. We shall classify these as *strong* or *weak* convergence criteria, respectively, and shall consider each type in some detail.

5 Strong convergence criterion

In problems involving direct simulations, filtering or testing statistical estimates of Ito processes, amongst other things, it is important that the approximating trajectores, that is the sample paths, are close to those of the Ito process. Thus this implies that some strong convergence criterion should be used. From a mathematical viewpoint it is often advantageous to investigate the absolute error at the terminal instant $t=T$:

$$A_\delta = \mathbf{E}\,|X_T - Y_n| \tag{21}$$

Here, A_δ can be estimated by the root mean square error via the Lyapunov inequality

$$A_\delta = \mathbf{E}\,|X_T - Y_n| \leq \mathbf{E}(|X_T - Y_n|^2)^{1/2} \tag{22}$$

This absolute error is certainly a criterion for the closeness of the sample paths of the Ito process X and the approximation Y at time T.

In what follows we shall say that an approximating process Y *converges in the strong sense with order* $\gamma \in (0,\infty]$ if there exists a constant $K \in (0,\infty)$ such that

$$\mathbf{E}\,|X_T - Y_n| \leq K\delta^\gamma \tag{23}$$

for any time discretization with maximum step size $\delta \in (0,1]$. In the deterministic case with vanishing diffusion and jump coefficients $b \equiv 0$, $c \equiv 0$ this strong convergence criterion reduces to the usual deterministic criterion used for the approximation of ordinary differential equations; see Butcher (1987) and Gear (1971). The order of a scheme is however sometimes less in the stochastic case than in the corresponding deterministic case, essentially because the increments ΔW_i^j are of root mean square order $\delta^{1/2}$ and not δ. In fact the Euler approximation for Ito stochastic differential equations has strong order $\gamma = 0.5$ in contrast with order 1.0 for Euler approximations for ordinary differential equations. See Atalla (1986), Gikhmann and Skorokhod (1979) and Maruyama (1955).

Higher order time discrete strong approximations of Ito diffusion have been proposed and investigated by Chang (1985), Clark (1978), Clark and Cameron (1980), Glorennec (1977), Jannsen (1982, 1984), Milstein (1974), Newton (1986), Nikitin and Razevig (1978), Platen (1980), Rao et al. (1974), Shimizu and Kawachi (1984), Talay (1982, 1983) and Wagner and Platen (1978). Strong approximations for Ito processes with jumps can be found in Dsagnidse and Tschitaschvili (1975), Platen (1982) and Wright (1980). In addition Runge-Kutta type approximations have been considered by Chang (1987), Clements and Anderson (1973), Kloeden and Pearson (1977), McShane (1974), Milstein (1974), Nikitin and Razevig (1978), Rumelin (1982) and Wright (1974).

6 Weak convergence criterion

In many practical situations it is not necessary to have a pathwise approximation of the Ito process. Often one may only be interested in the expectation of some function of the value of the Ito process at a given terminal time T, for instance the first two moments $\mathbf{E}X_T$ and $\mathbf{E}(X_T)^2$. Generally one may be interested in the expectation $\mathbf{E}\,g(X_T)$ for some function g. In the simulation of such a functional it is not necessary to approximate the exact path of the Ito process X. Rather, it is sufficient to approximate the probability distribution of the random variable X_T. Consequently we only require an approximation of the Ito process which is much weaker than the strong convergence criterion (23).

We shall say that a time discrete approximation Y *converges in the weak sense* with

order $\beta \in (0,\infty]$ if for any polynomial g there exists a constant $K \in (0,\infty)$ such that

$$|\mathrm{E}\, g(X_T) - \mathrm{E}\, g(Y_n)| \leq K \delta^{\beta} \tag{24}$$

for any time discretization with maximum step size $\delta \in (0,1]$. Obviously this weak convergence criterion reduces to the usual convergence in the deterministic case with $b \equiv 0$, $c \equiv 0$, $g(x) \equiv x$, just as it does for the strong convergence criterion.

Under assumptions of sufficient smoothness of the drift and diffusion coefficients, it was shown by Milstein (1978), Platen (1984) and Talay (1984) that an Euler approximation of an Ito diffusion has weak order of convergence $\beta = 1.0$. (Compare this with the strong order $\gamma = 0.5$). For Holder continuous coefficients, that is Lipschitz like to a fractional power, the order of weak approximation is lower; see Mikulevicius and Platen (1986). For discontinuous coefficients refer to Jannsen (1982). Higher order weak approximations have been extensively investigated by Milstein (1986), Pardoux and Talay (1985), Platen (1984) and Talay (1984). In particular, weak approximations of the Runge-Kutta type have been proposed and studied by Artemev (1985), Averina and Artemev (1986), Greenside and Helfand (1981), Haworth and Pope (1986), Helfand (1979), Klauder and Petersen (1985), Milstein (1986), Platen (1984), Petersen (1987) and Talay (1984). Higher order weak approximations for Ito processes with jumps can be found in Mikulevicius and Platen (1988). In addition, Wagner (1987) has investigated unbiased weak approximations, that is with $\beta = \infty$, to estimate functionals of Ito diffusions. Finally, Chang (1987) and Wagner (1988) have applied variance reduction techniques to weak approximations of Ito diffusions.

7 The stochastic Taylor formula

In the last two sections we grouped the literature on numerical methods for stochastic differential equations according to strong and weak convergence criteria. Another natural means of classification is to compare them with strong and weak Taylor approximations. The increments of such approximations are obtained by truncating the stochastic Taylor formula, which was derived in Wagner and Platen (1978) by the iterated application of the Ito formula. It was generalized and investigated in Azencott (1982), Platen (1982) and Platen and Wagner (1982).

The stochastic Taylor formula allows a function of an Ito process, that is $f(X_t)$, to be expanded about $f(X_{t_0})$ in terms of multiple stochastic integrals weighted by coefficients which are evaluated at X_{t_0}. These coefficients are formed from the drift, diffusion and jump coefficients of the Ito process and their derivatives up to some specified order. The remainder term in the formula contains a finite number of multiple stochastic integrals of the next highest multiplicity, but with nonconstant integrands. For example, in the one dimensional case $d=m=1$ a stochastic Taylor expansion for $f(X_t)$ about $f(X_{t_0})$ for $t_0 \in [t_0,T]$ and $t \in [t_0,T]$ may have the form

$$f(X_t) = f(X_{t_0}) + \left[a(X_{t_0}) f'(X_{t_0}) + \tfrac{1}{2}\{b(X_{t_0})\}^2 f''(X_{t_0})\right] \int_{t_0}^{t} ds$$

$$+ \int_U \{f(X_{t_0} + c(X_{t_0},u)) - f(X_{t_0}) - c(X_{t_0},u) f'(X_{t_0})\} \pi(du) \int_{t_0}^{t} ds + b(X_{t_0}) f'(X_{t_0}) \int_{t_0}^{t} dW_s$$

$$+\int_{t_0+U}^{t}\int\{f(X_{t_o}+c(X_{t_0},u))-f(X_{t_0})\}dM(du,ds)$$

$$+b(X_{t_0})\{b(X_{t_0})f''(X_{t_0})+b'(X_{t_0})f'(X_{t_0})\}\int_{t_0}^{t}\int_{t_0}^{s_2}dW_{s_1}dW_{s_2}+R \tag{25}$$

Here the remainder term R consists of stochastic integrals of higher multiplicity than those appearing in the expansion part of the formula. The stochastic Taylor formula can thus be thought of as a generalization of both the Ito formula and the deterministic Taylor formula.

By truncating stochastic Taylor expansions about successive discretization points, we can form time discrete Taylor approximations as functions of a Wiener process, which we may interpret as basic numerical schemes. In addition we can compare other schemes, such as those of the Runge-Kutta type, with the corresponding time discrete Taylor approximations. As we shall see, in order to get a higher order of strong or weak convergence in a numerical scheme we must use additional terms from the stochastic Taylor expansion.

8 Strong approximations of Ito diffusions

In the following two sections we shall consider approximations of Ito diffusions and then in Section 10 consider approximations for Ito processes with jump component.

8.1 Strong Taylor approximations

The simplest strong Taylor approximation of an Ito diffusion process is the *Euler scheme*. In the one dimensional case d=m=1 it has form

$$Y_{i+1}=Y_i+a\Delta_i+b\Delta W_i \tag{26}$$

for $i=0,1,2,...,n-1$ with initial value

$$Y_0=x_0. \tag{27}$$

Here $\Delta_i=\tau_{i+1}-\tau_i$ denotes the step size and $\Delta W_i=W_{\tau_{i+1}}-W_{\tau_i}$ the $N(0,\Delta_i)$ normally distributed increment of the standard Wiener process on the partition subinterval $[\tau_i, \tau_{i+1}]$. For example, the Euler scheme for the Ito stochastic differential equation

$$dX_t=-(\sin 2X_t+\tfrac{1}{4}\sin 4X_t)dt+\sqrt{2}\cos^2 X_t dW_t \tag{28}$$

is

$$Y_{i+1}=Y_i-(\sin 2Y_i+\tfrac{1}{4}\sin 4Y_i)\Delta_i+\sqrt{2}\cos^2 Y_i\Delta W_i \tag{29}$$

We shall indicate the results of a comparative study of different numerical schemes for this particular stochastic differential equation in Section 11.

In the general multidimensional case with $d, m=1,2,3,...$, the drift coefficient a is a d-dimensional vector, the diffusion coefficient b is a $d\times m$-matrix and the increment $\Delta W_i=\{\Delta W_i^j\}_{j=1}^m$ an m-dimensional random vector of independent $N(0,\Delta_i)$ normally distributed random variables ΔW_i^j for $j=1,2,...,m$. To simplify our notation we abbreviate $f(Y_i)$ to f for any function f (usually components of a and b, or their derivatives) in our following statements of numerical schemes. In addition we omit the standard initial condition (27). Then the k-th component of the multidimensional Euler scheme is given by

$$Y_{i+1}^k=Y_i^k+a^k\Delta_i+\sum_{j=1}^m b^{k,j}\Delta W_i^j \tag{30}$$

for $i=0,1,...,n-1$ where $k=1,2,...,d$. Note that the different components are coupled through the coefficients as in the corresponding stochastic differential equation. Under Lipschitz and linear growth conditions on a and b, the Euler scheme is shown in Gikhmann and Skorokhod (1972) to converge with strong order $\gamma=0.5$.

If in the one dimensional case $d=m=1$ we include an additional term from the Taylor expansion, then we obtain the *Milstein scheme* :

$$Y_{i+1}=Y_i+a\Delta_i+b\Delta W_i+\tfrac{1}{2}bb'\{(\Delta W_i)^2-\Delta_i\} \tag{31}$$

for $i=0,1,...,n-1$; see Milstein (1974). This additional term is from the double Wiener integral in (25), which can be easily computed from the increment ΔW_i since

$$\int_{\tau_i}^{\tau_{i+1}}\int_{\tau_i}^{s_2} dW_{s_1}dW_{s_2}=\tfrac{1}{2}\{(\Delta W_i)^2-\Delta_i\} \tag{32}$$

The Milstein scheme (31) for the stochastic differential equation (28) is

$$Y_{i+1}=Y_i-(\sin 2Y_i+\tfrac{1}{4}\sin 4Y_i)\Delta_i+\sqrt{2}\cos^2 Y_i\Delta W_i-2\cos^3 Y_i\sin Y_i\{(\Delta W_i)^2-\Delta_i\} \tag{33}$$

It turns out that the Milstein scheme (31) has strong order of convergence $\gamma=1.0$ under the assumptions that $\mathbf{E}(X_0)^4<\infty$, that a and b are twice continuously differentiable, and that a, a', b, b' and b'' are uniformly Lipschitz.

In the multidimensional case the k-th component of the Milstein scheme is

$$Y_{i+1}^k=Y_i^k+a^k\Delta_i+\sum_{j=1}^m b^{k,j}\Delta W_i^j+\tfrac{1}{2}\sum_{j_1,j_2=1}^m\sum_{l=1}^d b^{l,j_1}\frac{\partial b^{k,j_2}}{\partial x^l}\int_{\tau_i}^{\tau_{i+1}}\int_{\tau_i}^{s_2} dW_{s_1}^{j_1}dW_{s_2}^{j_2} \tag{34}$$

for $i=0,1,...,n-1$ and $k=1,2,...,d$. Again, the different components are coupled thought the coefficients here. For $m\geq 2$ this involves multiple Wiener integrals

$$I_{(j_1,j_2)}=\int_{\tau_i}^{\tau_{i+1}}\int_{\tau_i}^{s_2} dW_{s_2}^{j_1}dW_{s_1}^{j_2} \tag{35}$$

with $j_1\neq j_2$, which cannot be expressed simply in terms of the increments $\Delta W_i^{j_1}$ and $\Delta W_i^{j_2}$ of the corresponding components of the Wiener process, as in (32). At the end of this section we will suggest one possible way of approximating higher order multiple stochastic integrals like (35).

Suppose that the diffusion coefficient b satisfies the *commutativity condition*

$$\sum_{l=1}^d b^{l,j_1}(x)\frac{\partial b^{k,j_2}}{\partial x^l}(x)=\sum_{l=1}^d b^{l,j_2}(x)\frac{\partial b^{k,j_1}}{\partial x^l}(x) \tag{36}$$

for all $x\in\mathbf{R}^d$, $k=1,2,...,d$ and $j_1,j_2=1,2,...,m$. See Clark and Cameron (1980) and Sussman (1978). Then the Milstein scheme (34) reduces to the form

$$Y_{i+1}^k=Y_i^k+a^k\Delta_i+\sum_{j=1}^m b^{k,j}\Delta W_i^j$$

$$+\tfrac{1}{2}\sum_{j_1,j_2=1}^{m}\sum_{l=1}^{d} b^{l,j_1}\frac{\partial b^{k,j_2}}{\partial x^l}\Delta W_{s_1}^{j_1}\Delta W_{s_2}^{j_2}-\tfrac{1}{2}\sum_{j=1}^{m}\sum_{l=1}^{d} b^{l,j}\frac{\partial b^{k,j}}{\partial x^l}\Delta_i \tag{37}$$

for $i=0,1,...,n-1$ where $k=1,2,...,d$, which involves no multiple stochastic integrals of the form (35). The commutativity condition (36) is satisfied, for instance, when the diffusion coefficient is a constant or when each component of the Ito process is disturbed only by noise from another component of the Wiener process.

Generally speaking we can obtain more accurate strong Taylor approximations by including further multiple stochastic integrals from the stochastic Taylor expansion. Such integrals contain additional information about the sample paths of the Wiener process. Their presence is a fundamental difference between the numerical analysis of stochastic differential equations and ordinary differential equations. In the one dimensional case $d=m=1$ we obtain the following *strong Taylor approximation of order* $\gamma=1.5$

$$Y_{i+1}=Y_i+a\Delta_i+b\Delta W_i+\tfrac{1}{2}bb'\{(\Delta W_i)^2-\Delta_i\}+ba'\Delta Z_i+\tfrac{1}{2}\{aa'+\tfrac{1}{2}b^2a''\}\Delta_i^2$$
$$+\{ab'+\tfrac{1}{2}b^2b''\}\{\Delta W_i\Delta_i-\Delta Z_i\}+\tfrac{1}{2}b\{bb''+(b')^2\}\{\tfrac{1}{2}(\Delta W_i)^2-\Delta_i\}\Delta W_i \tag{38}$$

for $i=0,1,...,n-1$. Here an additional random variable ΔZ_i is required to represent the double integral

$$\Delta Z_i=\int_{\tau_i}^{\tau_{i+1}}\int_{\tau_i}^{s_2} dW_{s_1}ds_2 \tag{39}$$

It is normally distributed with mean $\mathrm{E}\Delta Z_i=0$, variance $\mathrm{E}(\Delta Z_i)^2=1/3(\Delta_i)^3$ and correlation $\mathrm{E}(\Delta Z_i\Delta W_i)=\tfrac{1}{2}(\Delta_i)^2$. All other multiple stochastic integrals appearing in the truncated Taylor expansion used to derive (38) can be expressed in terms of Δ_i, ΔW_i and ΔZ_i, thus resulting in (38). Under Lipschitz and linear growth conditions and sufficient smoothness on the coefficients a and b, the strong order of convergence $\gamma=1.5$ for (38) has been established in Wagner and Platen (1978) and Platen (1981). We remark that there is no difficulty in generating the pair $(\Delta W_i,\Delta Z_i)$ of correlated normally distributed random variables using the transformation

$$\Delta W_i=\xi_{i,1}\Delta_i^{1/2},\quad \Delta Z_i=1/2(\xi_{i,1}+\frac{1}{\sqrt{3}}\xi_{i,2})\Delta_1^{3/2} \tag{40}$$

where $\xi_{i,1}$ and $\xi_{i,2}$ are independent $N(0,1)$ random variables.

In order to express the multidimensional version of the strong order of 1.5 scheme (38) in a compact form we need some extra terminology. We introduce the operators:

$$L^0=\sum_{l=1}^{d}a^l\frac{\partial}{\partial x^l}+\tfrac{1}{2}\sum_{r,l=1}^{d}b^{r,l}b^{l,j}\frac{\partial^2}{\partial x^r\partial x^l} \tag{41}$$

and

$$L^j=\sum_{l=1}^{d}b^{l,j}\frac{\partial}{\partial x^l} \tag{42}$$

for $j=1,2,...,m$, and write $dW_s^0=ds$ and $b^{k,0}=a^k$ for $k=1,2,...,d$. In addition we denote multiple integrals up to multiplicity three by

$$I_{(j)}=\int_{\tau_i}^{\tau_{i+1}} dW_s^j, \tag{43}$$

$$I_{(j_1,j_2)}=\int_{\tau_i}^{\tau_{i+1}}\int_{\tau_i}^{s_2} dW_{s_1}^{j_1} dW_{s_2}^{j_2}, \tag{44}$$

and

$$I_{(j_1,j_2,j_3)}=\int_{\tau_i}^{\tau_{i+1}}\int_{\tau_i}^{s_3}\int_{\tau_i}^{s_2} dW_{s_1}^{j_1} dW_{s_2}^{j_2} dW_{s_3}^{j_3} \tag{45}$$

for $j,j_1,j_2,j_3=0,1,...,m$. Then the k-th component of the multidimensional strong Taylor approximation of order 1.5 satisfies

$$Y_{i+1}^k=Y_i^k+\sum_{j=0}^{m} b^{k,j} I_{(j)}+\sum_{j_1,j_2=0}^{m} L^{j_1} b^{k,j_2} I_{(j_1,j_2)}+\sum_{j_1,j_2,j_3=1}^{m} L^{j_1} L^{j_2} b^{k,j_3} I_{(j_1,j_2,j_3)} \tag{46}$$

for $i=0,1,...,n-1$ and $k=1,2,...,m$. See Platen (1981).

In Wagner and Platen (1978) and Platen (1981) it is described how schemes of any desired order of convergence can be constructed from the corresponding strong Taylor approximations. The practical implementation of such schemes involves the generation of multiple stochastic integrals such as $I_{(j_1,j_2)}$ and $I_{(j_1,j_2,j_3)}$, and of higher multiplicity, which is usually not easily done. If one is not willing to use such higher order multiple stochastic integrals, one could follow Clark (1978) and Newton (1986), who propose schemes which only use increments of the Wiener process. These numerical schemes are similar to the above strong Taylor approximations with modified random variables. In addition they are optimal in a sense within the class of strong order $\gamma=0.5$, or $\gamma=1.0$ if the commutativity condition (36) holds, time discrete approximations.

8.2 Strong Runge-Kutta approximations

A disadvantage of the above Taylor approximations is that the derivatives of various orders of the drift and diffusion coefficients must be evaluated at each step in addition to the drift and diffusion coefficients themselves. There are time discrete strong approximation schemes which avoid the use of derivatives. In analogy with ordinary differential equations we shall call these Runge-Kutta approximations, but we remark that they cannot always be obtained as simple generalizations of the widely used deterministic Runge-Kutta schemes because of the difference between Ito and ordinary calculus. See for example Clements and Anderson (1973) and Wright (1974).

In the one dimensional case with $d=m=1$ *a Runge-Kutta scheme of strong order* $\gamma=1.0$ is given by

$$Y_{i+1}=Y_i+a\Delta_i+b\Delta W_i+\tfrac{1}{2}\Delta_i^{-1/2}\{b(\hat{Y}_i)-b\}\{(\Delta W_i)^2-\Delta_i\} \tag{47}$$

with $\hat{Y}_i=Y_i+b\Delta_i^{1/2}$, for $i=0,1,..,n-1$. See Platen (1984). For the stochastic differential equation (29) this scheme is takes the form

$$Y_{i+1}=Y_i-(\sin 2Y_i+\tfrac{1}{4}\sin 4Y_i)\Delta_i+\sqrt{2}\cos^2 Y_i\Delta W_i$$

$$+\frac{1}{\sqrt{2\Delta_i}}(\cos^2(Y_i+\sqrt{2\Delta_i}\cos^2 Y_i)-\cos^2 Y_i)\{(\Delta W_i)^2-\Delta_i\} \tag{48}$$

The multidimensional version of the above first order strong Runge-Kutta scheme is component wise

$$Y_{i+1}^k = Y_i^k + a^k \Delta_i + \sum_{j=1}^{m} b^{k,j} \Delta W_i^j + \sum_{j_1,j_2=1}^{m} \{b^{k,j_2}(\hat{Y}_{i,j_1}^k) - b^{k,j_2}\} \Delta_i^{-1/2} I_{(j_1,j_2)} \tag{49}$$

with $\hat{Y}_{i,j_1}^k = Y_i^k + b^{k,j_1} \Delta_i^{1/2}$ for $i=0,1,...,n-1$ where $k=1,2,...,d$. Here the $I_{(j_1,j_2)}$ are the multiple stochastic integrals defined by (44).

Heuristically these Runge-Kutta schemes (47) and (49) can be obtained from the Milstein schemes (31) and (34) simply by replacing the derivatives there by the corresponding finite differences. From Clark and Cameron (1980) and Rumelin (1984) it follows with the restriction of using only the increments ΔW_i^j of the Wiener process to approximate the double integrals (44) that the Runge-Kutta scheme (49) has strong order of convergence $\gamma=0.5$, or $\gamma=1$ if the commutativity condition (36) holds. In Rumelin (1984) more general Runge-Kutta schemes can be found. However they do not give a higher order of strong convergence than (49) unless higher order multiple stochastic integrals are used, but that also applies to (49).

A detailed discussion of multistep and implicit methods is given in Kloeden and Platen (1990). Also see Petersen (1987).

Another type of strong approximation was investigated in Gorostiza (1980) and Newton (1986). In the one dimensional case $d=1$ time is discretized in such a way that a random walk takes place on a given set of levels in the state space. By this it is meant that the approximating process remains on a fixed level for a random time and then switches with given intensity to the next level above or below. For other more theoretical studies on strong approximations of Ito diffusions the reader is also referred to Doss (1977), Roemisch and Wakolbinger (1987), Sussman (1978) and Talay (1982).

8.3 Approximation of multiple stochastic integrals

It seems to be numerically impossible to generate exact values of multiple stochastic integrals such as $I_{(j_1,j_2)}$ defined by (44) for $j_1 \neq j_2$ with $j_1,j_2=1,..,m$. However for numerical purposes it often sufficies to have a good approximation of these integrals. Such an approximation can be obtained from the series expansion of the Wiener process. For example the j-th component of a Wiener process can be written as the series

$$W_{\tau_i+s}^j - W_{\tau_i}^j = s\Delta_i^{-1/2}\xi_{i,0}^j + \frac{(2\Delta_i)^{1/2}}{\pi}\sum_{r=1}^{\infty}\frac{1}{r!}\xi_{i,r}^j \sin\left(\frac{r\pi s}{\Delta_i}\right) \tag{50}$$

for $s \in [0,\Delta_i]$, $j=1,2,...,m$ and $i=0,1,...,n-1$, where the $\xi_{i,r}^j$ are independent standard normally distributed random variables. Following Liske et al. (1982) we then have

$$I_{(j_1,j_2)} = \int_{\tau_1}^{\tau_{i+1}}\int_{\tau_1}^{s_2} dW_{s_1}^{j_1} dW_{s_2}^{j_2} = \tfrac{1}{2}\Delta_i \xi_{i,0}^{j_1}\xi_{i,0}^{j_2} + \frac{2^{3/2}\Delta_i}{\pi^2}\sum_{k=1}^{\infty}\frac{(\xi_{i,2k-1}^{j_1}\xi_{i,0}^{j_2} - \xi_{i,0}^{j_1}\xi_{i,2k-1}^{j_2})}{(2k-1)^2}$$

$$+\frac{4\Delta_i}{\pi^2}\sum_{k,l=1}^{\infty}\frac{(\xi_{i,2k-1}^{j_1}\xi_{i,2l}^{j_2} - \xi_{i,2l}^{j_1}\xi_{i,2k-1}^{j_2})}{(2k-1)^2-(2l)^2} \tag{51}$$

If we truncate this series by using only the first $p=1,2,...$ normally distributed random variables, we obtain the following approximation $I_{(j_1,j_2)}^p$ of $I_{(j_1,j_2)}$

$$I^p_{(j_1,j_2)}=\tfrac{1}{2}\Delta_i\xi^{j_1}_{i,0}\xi^{j_2}_{i,0}+\frac{2^{3/2}\Delta_i}{\pi^2}\sum_{k=1}^{[(p+1)/2]}\frac{(\xi^{j_1}_{i,2k-1}\xi^{j_2}_{i,0}-\xi^{j_1}_{i,0}\xi^{j_2}_{i,2k-1})}{(2k-1)^2}$$

$$+\frac{4\Delta_i}{\pi^2}\sum_{k=1}^{[(p+1)/2]}\sum_{l=1}^{[p/2]}\frac{(\xi^{j_1}_{i,2k-1}\xi^{j_2}_{1,2l}-\xi^{j_1}_{i,2l}\xi^{j_2}_{i,2k-1})}{(2k-1)^2-(2l)^2} \tag{52}$$

where $[q]$ is the integer part of q. When $p=5$, for example, the mean square error estimate $\mathbf{E}\,|I_{(j_1,j_2)}-I^p_{(j_1,j_2)}|^2\le 0.01(\Delta_i)^2$ can be derived for such an approximation. In principle other multiple stochastic integrals can be approximated in the same way.

9. Weak approximations of Ito diffusions

9.1 Weak Taylor approximations

When we are interested only in weak approximations of an Ito process, that is a process with approximately the same probability distribution, we have many more degrees of freedom than with strong approximations. For example it suffices to use an initial value $Y_0=\hat{X}_0$ with a convenient probability distribution which approximates that of X_0 in an appropriate way. In addition the random increments ΔW_i of the Wiener process can be replaced by other more convenient approximation $\Delta\hat{W}_i$ which have similar moment properties to the ΔW_i. In a weak approximation of order $\beta=1$ we could, for instance, choose independent $\Delta\hat{W}_i$ for $i=0,1,...,n-1$ and $j=1,2,...,m$ with

$$\mathbf{E}(\hat{W}_i)^r=\begin{cases}0 & \text{for } r=1 \text{ or } 3\\ \Delta_i & r=2\\ Z_r(\Delta_i) & \text{for } r=4,5,...\end{cases} \tag{53}$$

where

$$|Z_r(\Delta_i)|\le K\Delta_i^2 \tag{54}$$

for $r=4,5,...$ and some constant $K\in(0,\infty)$. This means we could use an easily generated telegraphic noise process instead of the normally distributed increments ΔW_i, that is a two point process $\Delta\hat{W}_i$ taking values $\pm\Delta_i^{1/2}$ with equal probabilities:

$$P(\Delta\hat{W}_i=\pm\Delta_i^{1/2})=\tfrac{1}{2} \tag{55}$$

The simplest useful weak Taylor approximation is again the Euler approximation introduced in Section 8.1, that is (26) for the one dimensional case and (27) for the multidimensional case. Then it follows from results in Talay (1984) that the *weak Euler scheme*

$$Y^k_{i+1}=Y^k_i+a^k\Delta_i+\sum_{j=1}^{m}b^{k,j}\Delta\hat{W}^j_i \tag{56}$$

for $i=0,1,...,n-1$ and $k=1,2,...,d$ has weak order $\beta=1.0$. This holds if the coefficients a and b are four times continuously differentiable with these derivatives satisfying a growth condition. Compare this with the strong order of convergence $\gamma=0.5$ of (30). If the drift and diffusion coefficients are not so smooth or even if they are only Holder continuous, then from a result of Mikulevicius and Platen (1986) the scheme (56) is still weakly convergent but with a smaller order $\beta<1$.

Using the truncated stochastic Taylor expansion we can also construct weak Taylor approximations of higher order $\beta=2,3,...$ In the one dimensional case $d=m=1$ the *weak*

Taylor approximation of order $\beta=2$ has, following Milstein (1978) and Talay (1984), the form

$$Y_{i+1}=Y_i+a\Delta_i+b\Delta\hat{W}_i+\tfrac{1}{2}bb'\{(\Delta\hat{W}_i)^2-\Delta_i\}$$
$$+ba'\Delta\hat{Z}_i+\tfrac{1}{2}\{aa'+\tfrac{1}{2}b^2a''\}\Delta_i^2+\{ab'+\tfrac{1}{2}b^2b''\}\{\Delta\hat{W}_i\Delta_i-\Delta\hat{Z}_i\} \tag{57}$$

for $i=0,1,...,n-1$. Here $\Delta\hat{W}_i$ approximates ΔW_i and $\hat{Z}_i$ the multiple stochastic integral (39), that is $\Delta Z_i=\int_{\tau_i}^{\tau_{i+1}}\int_{\tau_i}^{s_2} dW_{s_1}ds_2$

As with the weak Euler scheme (56) we can use random variables $\Delta\hat{W}_i$ and $\Delta\hat{Z}_i$ which have approximately the same moment properties of ΔW_i and ΔZ_i. For example, we could choose

$$\Delta\hat{W}_i=\Delta W_i, \quad \Delta\hat{Z}_i=\tfrac{1}{2}\Delta_i\Delta W_i \tag{58}$$

See Platen (1984) and Talay (1984). Alternately we could use a three point random variable T_i with probability distribution

$$P(T_i=\pm\sqrt{3})=1/6, \quad P(T_i=0)=2/3 \tag{59}$$

and set

$$\Delta\hat{W}_i=\Delta_i^{1/2}T_i, \quad \Delta\hat{Z}_i=\tfrac{1}{2}\Delta_i^{3/2}T_i \tag{60}$$

The multidimensional version of the second order weak Taylor approximation can be written compactly as

$$Y_{i+1}^k=Y_i^k+\sum_{j=0}^{m}b^{k,j}\hat{I}_{(j)}+\sum_{j_1,j_2=0}^{m}L^{j_1}b^{k,j_2}\hat{I}_{(j_1,j_2)} \tag{61}$$

for $i=1,2,...,n-1$ and $k=1,2,...,d$ using the operators L^j defined in (41)-(42) and the other notation introduced there. (See Platen (1982)). Here the random variables $\hat{I}_{(j)}$ and $\hat{I}_{(j_1,j_2)}$ are approximations of the corresponding multiple stochastic integrals $I_{(j)}$ and $I_{(j_1,j_2)}$ defined by (43) and (44), respectively. We can generate them in the following way. For $i=0,1,...,n-1$ and $j=1,2,...,m$ we generate independent standard normally distributed random variable $Z_{i,j}$. Alternately we could use the three point random variables with the distribution (59). In addition for $i=0,1,...,n-1$ and $r=1,2,...,j-1$ with $j=1,2,...,m$ we generate independent telegraphic random variables $V_{i,j,r}$ with

$$P(V_{i,j,r}=\pm1)=\tfrac{1}{2} \tag{62}$$

and for $r=j+1,...,m$ we define

$$V_{i,j,r}=-V_{i,r,j} \tag{63}$$

In terms of these random variables we then form the following approximations of the multiple stochastic integrals for $i=0,1,...,n-1$ and $j_1,j_2=1,2,...,m$:

$$\hat{I}_{(0)}=\Delta_i, \quad \hat{I}_{(j_1)}=\Delta_i^{1/2}Z_{i,j_1}, \quad \hat{I}_{(0,0)}=\tfrac{1}{2}\Delta_i^2, \quad \hat{I}_{(0,j_1)}=\hat{I}_{(j_1,0)}=\tfrac{1}{2}\Delta_i^{3/2}Z_{i,j_1}$$
$$\hat{I}_{(j_1,j_1)}=\tfrac{1}{2}\Delta_i\{Z_{i,j_1}^2-\Delta_i\}, \quad \hat{I}_{(j_1,j_2)}=\tfrac{1}{2}\Delta_i\{Z_{i,j_1}Z_{i,j_2}+V_{i,j_1,j_2}\} \quad \text{for } j_1\neq j_2 \tag{64}$$

The second order weak convergence of the numerical scheme (61) was verified in Platen (1984) and Talay (1984) under the assumptions that $E(X_0)^6<\infty$, that the coefficients a

and b are six times continuously differentiable with bounded derivatives, and that the function b^{k,j_1} and $L^{j_1}b^{k,j_2}$ for $k=1,2,...,d$ and $j_1,j_2=0,1,...,m$ satisfy a linear growth bound.

In order to derive the third order weak Taylor approximation we need to consider the third order multiple stochastic integrals

$$I_{(j_1,j_2,j_3)}=\int_{\tau_i}^{\tau_{i+1}}\int_{\tau_i}^{s_2}\int_{\tau_i}^{s_1} dW_{s_1}^{j_1}dW_{s_2}^{j_2}dW_{s_3}^{j_3} \tag{65}$$

or some convenient approximation $\hat{I}_{(j_1,j_2,j_3)}$ for $j_1,j_2,j_3=0,1,...,m$. (Recall that we denote $dW_s^0=ds$.) Following Platen (1984) the *third order weak Taylor approximation* has component wise form

$$Y_{i+1}^k=Y_i^k+\sum_{j=0}^{m} b^{k,j}\hat{I}_{(j)}+\sum_{j_1,j_2=0}^{m} L^{j_1}b^{k,j_2}\hat{I}_{(j_1,j_2)}+\sum_{j_1,j_2,j_3=0}^{m} L^{j_1}L^{j_2}b^{k,j_3}\hat{I}_{(j_1,j_2,j_3)} \tag{66}$$

for $i=1,2,...,n-1$ and $k=1,2,...,d$. As above the random variables $\hat{I}_{(j)}$ and $\hat{I}_{(j_1,j_2)}$ are convenient approximations of the first and second order multiple stochastic integrals $I_{(j)}$ and $I_{(j_1,j_2)}$. For instance in the case of a single component Wiener process, that is when $m=1$, we can use independent standard normally distributed random variables $N_{i,1}$ and $N_{i,2}$ for $i=0,1,...,n-1$ and take

$$\hat{I}_{(0)}=\Delta_i,\quad \hat{I}_{(0,0)}=\tfrac{1}{2}\Delta_i^2,\quad \hat{I}_{(0,0,0)}=\frac{1}{6}\Delta_i^3,$$

$$\hat{I}_{(1)}=\Delta_i^{1/2}N_{i,1},\quad \hat{I}_{(1,1)}=\tfrac{1}{2}\Delta_i\{N_{i,1}^2-1\},\quad \hat{I}_{(0,1)}=\tfrac{1}{2}\Delta_i^{3/2}\{N_{i,1}+\frac{1}{\sqrt{3}}N_{i,2}\}, \tag{67}$$

$$\hat{I}_{(1,0)}=\tfrac{1}{2}\Delta_i^{3/2}\{N_{i,1}-\frac{1}{\sqrt{3}}N_{1,2}\},\quad \hat{I}_{(0,0,1)}=\hat{I}_{(0,1,0)}=\hat{I}_{(1,0,0)}=\frac{1}{6}\Delta_i^{5/6}N_{i,1},$$

$$\hat{I}_{(0,1,1)}=\hat{I}_{(1,0,1)}=\hat{I}_{(1,1,0)}=\frac{1}{6}\Delta_i^2\{N_{i,1}^2-1\},\quad \hat{I}_{(1,1,1)}=\frac{1}{6}\Delta_i^{3/2}N_{i,1}\{N_{i,1}^2-3\}$$

It was shown in Platen (1984) that the scheme (56) has third order weak convergence under the assumptions that the coefficients a and b are eight times continuously differentiable with bounded derivatives, and that the functions b^{k,j_1}, $L^{j_1}b^{k,j_2}$ and $L^{j_1}L^{j_2}b^{k,j_3}$ for $k=1,2,...,d$ and $j_1,j_2,j_3=0,1,...,m$ satisfy a linear growth bound.

Under corresponding assumptions it was shown in Mikulevicius and Platen (1988) and Platen (1984) that a Taylor approximation of a general weak order of convergence $\beta=1,2,3,...$ Using the preceeding terminology we can write such a *weak Taylor approximation of order* β as

$$Y_{i+1}^k=Y_i^k+\sum_{r=1}^{\beta}\sum_{j_1,j_2,j_3=0}^{m} L^{j_1}...L^{j_{r-1}}b^{k,r}\hat{I}_{(j_1,j_2,\ldots,j_r)} \tag{68}$$

for $i=0,1,...,n-1$ and $k=1,2,...,d$.

9.2 Weak Runge-Kutta approximations

As with strong approximations it is often desirable practically to have weak approximations of the Runge-Kutta type which avoid the use of derivatives, particularly higher order derivatives, of the drift and diffusion coefficients. One of the earliest schemes of this kind is the *second order weak Runge-Kutta approximation* proposed by Milstein

(1986) for the one dimensional case $d=m=1$. It has the form

$$Y_{i+1}=\tfrac{1}{2}\{a-bb'\}\Delta_i+\tfrac{1}{2}b\Delta W_i+\tfrac{1}{2}bb'(\Delta W_i)^2+\tfrac{1}{2}a(\hat{Y}_i)\Delta_i+\tfrac{1}{4}\{b(\hat{Y}_i^+)+b(\hat{Y}_i^-)\}\Delta W_i \tag{69}$$

with

$$\hat{Y}_i=Y_i+a\Delta_i+b\Delta W_i, \quad \text{and} \quad \hat{Y}_i^{\pm}=Y_i+a\Delta\pm\frac{1}{3}b\Delta W_i$$

for $i=0,1,...,n-1$. Another *second order weak Runge-Kutta scheme* was presented in Talay (1984) and has

$$Y_{i+1}=Y_i+\tfrac{1}{2}\{a(\bar{Y}_i)-\tfrac{1}{2}b(\bar{Y}_i)b'(\bar{Y}_i)\}\Delta_i+b\{A_i-B_i\}[\tfrac{1}{2}\Delta_i]^{1/2}$$
$$+b(\bar{Y}_i)B_i(2\Delta_i)^{1/2}+\tfrac{1}{2}\Delta_i\{b(\bar{Y}_i)b'(\bar{Y}_i)-bb'\}B_i^2-\tfrac{1}{2}bb'A_iB_i\Delta_i \tag{70}$$

with

$$\bar{Y}_i=Y_i+\tfrac{1}{2}\{a-\tfrac{1}{2}bb'\}\Delta_i+bA_i[\tfrac{1}{2}\Delta_i]^{1/2}+\tfrac{1}{4}bb'A_i^2\Delta_i$$

for $i=0,1,...,n-1$. Here the A_i and B_i are independent random variables, which are for example either standard normally distributed or as in (59). A multidimensional version of (70) can be found in Pardoux and Talay (1985).

Both of the above schemes still use the derivative b' of the diffusion coefficient b, but always in the product combination $\tfrac{1}{2}bb'$. This is the correction term relating the Ito and Stratonovich interpretations of a stochastic differential equation and occurs here because, roughly speaking, a deterministic Runge-Kutta scheme applied to a stochastic differential equation converges to the Stratonovich version of the equation. It is however possible to also avoid using the derivative b'. A second order weak Runge-Kutta approximation with this feature is, in the scalar case $d=m=1$, given by

$$Y_{i+1}=Y_i+\tfrac{1}{2}\{a(\hat{Y}_i)+a\}\Delta_i+\tfrac{1}{4}\{b(Y_i^+)+b(Y_i^-)+2b\}\Delta\hat{W}_i$$
$$+\tfrac{1}{4}\{b(Y_i^+)-b(Y_i^-)\}\{(\Delta\hat{W}_i)^2-\Delta_i\}\Delta_i^{-1/2} \tag{71}$$

with

$$\hat{Y}_i=Y_i+a\Delta_i+b\Delta\hat{W}_i, \quad \text{and} \quad Y_i^{\pm}=Y_i+a\Delta_i\pm b\Delta_i^{1/2}$$

for $i=0,1,...,n-1$ where the $\Delta\hat{W}_i$ can be chosen as in (58) and (60). See Platen (1984). The multidimensional generalization of this scheme is

$$Y_{i+1}^k=Y_i^k+\tfrac{1}{2}\{a^k(\bar{Y}_i)+a^k\}\Delta_i$$
$$+\tfrac{1}{4}\sum_{j=1}^{m}[\{b^{k,j}(Y_{i,j}^+)+b^{k,j}(Y_{i,j}^-)+2b^{k,j}\}+\sum_{r=1,r\neq j}^{m}\{b^{k,j}(\hat{Y}_{i,j}^+)+b(\hat{Y}_{i,j}^-)-2b^{k,j}\}]\hat{I}_{(j)}$$
$$+\tfrac{1}{2}\Delta_i^{-1/2}\sum_{j=1}^{m}[\{b^{k,j}(Y_{i,j}^+)-b^{k,j}(Y_{i,j}^-)\}\hat{I}_{(j,j)}+\sum_{r=1,r\neq j}^{m}\{b^{k,j}(\hat{Y}_{i,r}^+)-b^{k,j}(\hat{Y}_{i,r}^-)\}\hat{I}_{(r,j)}] \tag{72}$$

with

$$Y_i^{\pm k}=Y_i^k+a^k\Delta_i\pm b^{k,j}\Delta_i^{1/2}, \text{ and } \hat{Y}_{i,j}^{\pm k}=Y_{i,j}^k\pm b^{k,j}\Delta_i^{1/2}, \quad \bar{Y}_i^k=Y_i^k+a^k\Delta_i+\sum_{j=1}^{m}b^{k,j}\hat{I}_{(j)}$$

for $i=0,1,...,n-1$ and $k=1,2,...,d$, where the random variables $\hat{I}$ can be chosen as in (64). The second order weak convergence of the Runge-Kutta schemes has been established in Platen (1984) and Talay (1984) under the same assumptions as for second order weak

Taylor approximations.

In Greenside and Helfand (1984), Haworth and Pope (1986), Helfand (1979), Klauder and Petersen (1985) and Petersen (1987), amongst others, Runge-Kutta schemes are considered with convergence only in first two moments. This is a weaker convergence criterion than the weak convergence criterion (24) considered above. Obviously a scheme which converges with weak order $\beta=0,1,2,...$ will converge not only in the first two moments but also in all higher moments when they exist.

Another way of constructing weak approximations for diffusion processes has been proposed by Wagner (1987). It is based on the Monte Carlo simulation of functional integrals and uses unbiased, variance reduced approximations to estimate functionals of an Ito diffusion process.

10 Strong and weak approximations of Ito processes with jump components

10.1 Jump adapted time discretizations

The Poisson process $p=\{p(U,[0,t]);t\geq 0\}$ generates with finite intensity $\pi(U)$ a sequence of jump times $\sigma=\{\sigma_1,\sigma_2,\ldots,\sigma_k,\ldots\}$. These jump times are realizations of a random process, so if we include those which are not greater than T in a time discretization of the interval $[0,T]$ we obtain a random time discretization. To be compatible with the Wiener process these random discretization times should be A-adapted. To be specific, we consider a time discretization $(\tau)_\delta=\{\tau_0,\tau_1,...\}$ of $[0,T]$ with maximum step size $\delta>0$ to be a sequence of A-adapted stopping times $\{\tau_0,\tau_1,...\}$ with $0=\tau_0<\tau_1<...<\tau_{i_T}=T$ where $i_T<\infty$ with probability one, which includes all of the jump times from (σ) not greater than T. Here we are using the integer

$$i_t=\max\{i:\tau_i\leq t \quad \text{for } i=0,1,2,...\} \tag{73}$$

for all $t\in[0,T]$. In addition for all $i=0,1,...,i_T-1$ we assume that the stopping time τ_{i+1} is A_{τ_i}-measurable if it is not a jump time of the Poisson process. More precisely, for each $i=0,1,2,...$ we define the sub-σ-algebra of A

$$A_i^p=\sigma(1_{\{p(U,\{\tau_{i+1}\})\neq 0\}})\vee A_{\tau_i}$$

and assume that τ_{i+1} is A_i^p-measurable. Finally we suppose that

$$\max\{\tau_i-\tau_{i-1};i=1,2,...,i_T\}\leq\delta \tag{74}$$

For instance, the time discretization $(\tau)_\delta$ could be the superposition of all jump times of p not greater than T with a deterministic time discretization $[0,T]$ with maximum step size δ. Another example which permits more control over the step size can be constructed as follows. At the initial discretization time $\tau_0=0$ we first compute an A_{τ_0}-measurable next possible discretization time $\hat{\tau}_1$ with $\hat{\tau}_1-\tau_0\leq\delta$. If there is no jump between τ_0 and $\hat{\tau}_1$ we take $\hat{\tau}_1$ to be the next discretization instant τ_1. Otherwise we choose the first jump time of the Poisson process greater than τ_0 to be τ_1. Then we start the procedure again at τ_1 in the same manner, and so on until we reach time T.

10.2 Jump adapted time discrete approximations

For a given deterministic time discretation similar results to those for strong approximations of Ito diffusions based on stochastic Taylor expansions are proved in Platen (1982) for Ito processes with jump components. This paper uses multiple stochastic integrals with respect to the Poisson jump measure p, which are generally difficult to calculate numerically.

For strong time discrete approximations based on jump adapted time discretizations it is also shown in Platen (1984) that any desired order of strong approximation can be obtained under appropriate assumptions on the coefficients. Moreover these approximations are generally much easier to calculate. We recall that for such a jump adapted time discretization the jumps of p only occur at discretization points. Consequently between any two discretization points τ_i and τ_{i+1} we need only approximate the increment in the diffusion until time τ_{i+1} by $Y^-_{i+1}-Y_i$. Then at time τ_{i+1} we approximate the jump which may occur there by an expression $Y_{i+1}-Y^-_{i+1}$, which will be zero if no jump occurs at τ_{i+1}. To be more precise, in the one dimensional case $d=m=1$ we have the following recursion for the *jump adapted Euler approximation*

$$Y^-_{i+1}=Y_i+\{a-\int_U c(Y_i,u)\pi(du)\}\Delta_i+b\Delta W_i, \tag{75}$$

$$Y_{i+1}=Y^-_{i+1}+\sum_{r=1}^{\infty} c(Y^-_{i+1},u_r)1_{\{\sigma_r=\tau_{r+1}\}}$$

for $i=0,1,...,n-1$ where as in (20) σ_r denotes the r-th jump and u_r the r-th jump mark of the Poisson process $p=\{p(U,[0,t]);t\geq 0\}$.

We note that we split the Euler approximation (9) into two parts, with the first approximating a diffusion with modified drift

$$a^*(x)=a(x)-\int_U c(x,u)\pi(du) \tag{76}$$

formed from the original drift a and the jump compensation term, and with diffusion coefficient b. The second part then approximates only the jumps according to the Poisson process p. For Lipschitz and growth conditions on the coefficients a, b, c it follows as in the pure diffusion case that this jump adapted Euler process converges with $\gamma=0.5$.

In order to formulate general jump adapted approximations we denote by $F^k_{a,b,i}$ the increment in the k-th component of a time discrete approximation on the interval $[\tau_i,\tau_{i+1}]$ of a diffusion process with drift a and diffusion coefficient b, starting at x at time τ_i. Then we have:

$$Y^{-k}_{i+1}=Y^k_i+F^k_{a^*,b,i}(Y^k_i) \quad Y^k_{i+1}=Y^{-k}_{i+1}+\sum_{r=1}^{\infty} c(Y^{-k}_{i+1},u_r)1_{\{\sigma_r=\tau_{r+1}\}} \tag{77}$$

for $i=0,1,...,n-1$ and $k=1,2,...,d$. If the $F^k_{a^*,b,i}$ correspond to one of the time discrete diffusion approximations of strong order $\gamma>0$ which were discussed in Section 8, then it has been shown in Platen (1984) that the jump adapted approximation (77) also converges with strong order γ under analogous assumptions on the coefficients a, b, c as for the diffusion approximations of the same order. For weak approximations a similar result was proved in Mikulevicius and Platen (1988). To be specific, if the $F^k_{a^*,b,i}$ correspond to one of the time discrete diffusion approximations of weak order $\beta>0$ as described in Section 9, then the jump adapted process (77) also converges with weak order β under analogous assumptions on the coefficients a, b, c.

11 Choosing and implementing a numerical scheme

The choice and implementation of an appropriate numerical method for solving a particular stochastic differential equation is much more complicated than for ordinary differential equation. This is due not only to the greater complexity of the schemes and general lack of software, but also to the additional demands of the stochastic situation. In the

space available we cannot do more than briefly mention some of the important issues that must be taken into account.

Probably the simplest choice is between a strong and a weak scheme, this being determined by the purpose of the task at hand. Then comes the choice of the order of the scheme to be used and of the actual scheme itself. In making this choice the user must bear in mind the discretization error, the ease of determining the coefficients of the scheme and of programming it, the computational efficiency of the scheme, the computational round-off errors, the sampling errors from the use of pseudo-random number generators and, of course, the computing resources available. When the scheme has been programmed and is ready for implementation, the user must then choose an appropriate time step, time discretization and number of different sample paths to be calculated. Usually calculations are performed with different time steps and sample numbers, and the outputs tested numerically and statistically to obtain some indication of their stability, convergence and validity. Unfortunately there are few definitive answers to these questions, so some amount of experimentation is required.

Much of the literature cited in the preceding sections concentrates on the theoretical aspects of the numerical methods, such as proving a particular order of convergence, although some also give results of test runs on simple examples for the method under discussion. The paper of Liske and Platen (1987) gives a detailed computational and statistical comparison of the mean and mean square errors for the Euler (26), Milstein (31), second order derivative free Runge-Kutta (71) and third order Taylor (57) schemes for the nonlinear stochastic differential equation (29), which is explicitly solvable. Specifically, for each scheme 100 blocks of 100 sample paths were calculated for each of three different time steps, and Student t-tests carried out on the results. In addition a comparison of CPU times is given, being roughly in the ratio 1:1.4:1.6:4.5 for the four schemes and doubling as the time step is halved. The derivative free Runge-Kutta scheme seemed preferable in these experiments in the sense of a balance between programming effort, computational efficiency and statistical reliability, see also Greiner et al. (1987).

An extensive investigation and statistical analysis of sample estimates for the Euler method can be found in Schenkl (1988). Klauder and Petersen (1985) and Petersen (1987) compare test runs on explicitly solvable equations for a specific second order weak Runge-Kutta scheme. In particular, Klauder and Petersen (1985) consider different numbers of sample paths and discuss their simultaneous calculation on vector array computers. They also compare the discretization error due to the time step size with the sampling errors and round-off errors. They found that the latter tended to dominate when the time step was too small, particularly in a third order version of their scheme which contained 65 seperate terms! Haworth and Pope (1987) comment that the error in generating pseudo-random numbers is generally negligible compared to the above mentioned errors, and that the CPU time required to generate such numbers is typically about 2% of the total for fairly large, complicated problems, although in very simple test examples it can be as high as 50%.

Those readers who have worked out the various numerical schemes surveyed in this article for the simple nonlinear stochastic differential equation (29) will have discovered that the coefficients corresponding to higher order terms soon become algebraically complicated. Leblond and Talay (1986) have developed an expert system package called PRESTO which used the algebraic manipulator REDUCE to determine the appropriate scheme coefficients for given drift and diffusion coefficients for a number of different numerical schemes and to generate a Fortran program.

We conclude this survey with the remark that the theoretical understanding and the practical application of numerical methods for stochastic differential equations are still in their infancy. The existing literature is in many ways ad hoc and contains many gaps. The future theoretical development and applicability of the subject will be greatly

stimulated by feedback from those who apply such methods to real problems.

Acknowledgements: Both authors would like to thank Ludwig Arnold of the Institute for Dynamical Systems, University of Bremen, for his hospitality in Bremen during which time they wrote this article. In addition, the first named author would like to thank the Karl-Weierstrass-Institute for Mathematics of the Academy of Sciences of the German Democratic Republic in Berlin for its hospitality during several visits in 1988 and 1989.

References

Arnold, L. 1974: Stochastic differential equations. Wiley, New York

Artemev, S.S. 1985: A variable step algorithm for the numerical solution of stochastic differential equations. Methody Mekhaniki Sploshnoi Sredy 16 (2), 11-23. (In Russian)

Atalla, M.A. 1986: Finite-difference approximations for stochastic differential equations. In Probabilistic Methods for the Investigation of Systems with an Infinite Number of Degrees of Freedom. Collection of Scientific Works. Kiev, pp. 11-16

Averina, T.A.; Artemev, S.S. 1986: A new family of numerical methods for solving stochastic differential equations. Dokl. Akad. Nauk SSSR 288 (4), 777-780. (In Russian)

Azencott, R. 1982: Formule de Taylor stochastique et development as asymptotique d'integrales de Feynmann. Springer Lecture Notes in Mathematics Vol. 921, pp. 237-285

Balakrishnan, A.V. 1985: On a class of stochastic differential equations which do not satisfy a Lipschitz condition. Springer Lecture Notes in Control and Inform. Sc. Vol. 78

Bodo, B.A.; Thompson, M.E.; Unny, T.E. 1987: A review on stochastic differential equations for applications in hydrology. J. Stoch. Hydrol. Hydraulics 1, 81-100

Boyce, W.E. 1978: Approximate solution of random ordinary differential equations. Adv. Appl. Prob. 10, 172-184

Butcher, J.C. 1987: The numerical analysis of ordinary differential equations. Runge-Kutta and general linear methods. Wiley, Chirchester.

Bywater, R.J.; Chung, P.M. 1973: Turbulent flow fields with two dynamically significant scales. AIAA papers 73-646, 991-10

Chang, C.C. 1985: Numerical solution of stochastic differential equations. Ph.D. dissertation, University of California, Berkeley, 1985

Chang, C.C. 1987: Numerical solution of stochastic differential equation with constant diffusion coefficients. Math. Computat. 49, 523-542

Clements, D.J.; Anderson, B.D.O. 1973: Well behaved Ito equations with simulations that always misbehave. IEEE Trans. Autom. Control AC-18, 676-677

Clark, J.M.C. 1978: The design of robust approximations to the stochastic differential equations of nonlinear filtering. In Communication systems and Random process theory, edited by J.K. Skwirzynski. Sijthoff and Noordhoff, Alphe naan de Rijn, pp. 721-734

Clark, J.M.C.; Cameron, R.J. 1980: The maximum rate of convergence of discrete approximations for stochastic differential equations. Springer Lecture Notes In Control And Inform. Sc. Vol. 25, pp. 162-171

Dashevski, M.L.; Liptser, R.S. 1966: Simulation of stochastic differential equations connected with the disorder problem by means of analog computer. Autom. Remote Control 27, 665-673. (In Russian)

Doss, H. 1977: Liens entre equations differentielles stochastiques et ordinaire. Ann. Inst. Henri Poincare 13, 99-125.

Dsagnidse, A.A.; Tschitashvili, R.J. 1975: Approximate integration of stochastic differential equations. Tbilisi State University, Institute of Appl. Math. "Trudy IV", Tibilisi, pp. 267-279. (In Russian)

Ermakov, S.M. 1975: The Monte Carlo method and related questions. Nauka, Moscow. (In Russian)

Fahrmeier, L. 1976: Approximation von stochastischen Differenzialgleichungen auf Digitial- und Hybridrechnern. Computing 16, 359-371

Gard, T.C. 1988: Introduction to stochastic differential equations. Marcel Dekker, New York

Gear, C.W. 1971: Numerical initial value problems in ordinary differential equations. Prentice-Hall, Englewood Cliffs, N.J.

Geman, S.; Hwang, C. 1986: Diffusions for global optimization. SIAM J. Control Optim. 24, 1031-1043

Greiner, A.; Strittmatter, W.; Honerkamp, J. 1987: Numerical integration of stochastic differential equations. J. Stat. Physics 51, 95-108

Gikhmann, I.I.; Skorokhod, A.V. 1972: Stochastic differential equations. Springer, Berlin

Gikhmann, I.I.; Skorokhod, A.V. 1979: The theory of stochastic processes, Vol. I-III. Springer, Berlin.

Glorennec, P.Y. 1977: Estimation a priori des erreurs dans la resolution numerique d'equations differentials stochastiques. Seminare de Probabilite, Universite de Rennes, Facsicule 1, 57-93

Gorostiza, L.G. 1980: Rate of convergence of an approximate solution of stochastic differential equations. Stochastics 3, 267-276. Erratum Stochastics 4 (1981), 85

Greenside, H.S.; Helfand, E. 1981: Numerical integration of stochastic differential equations II. Bell System Tech. J. 60, 1927-1940

Grigelonius B. 1981: On weak convergence of semimartingales. Lietuvos Matem. Rink. 21, no. 3, 9-25

Harris, C.J. 1976: Simulation of nonlinear stochastic equations with applications in modeling water pollution. In Mathematical models for environmental problems, edited by C.A. Brebbi. Pentech Press, London, pp. 269-282

Haworth, D.C.; Pope, S.B. 1986: A second-order Monte Carlo method for the solution of the Ito stochastic differential equation. Stoch. Anal. Appl. 4, 151-186

Helfand, E. 1979: Numerical integration of stochastic differential equations. Bell System Tech. J. 58, 2239-2299

Horsthemke, W.; Lefevre, R. 1984: Noise induced transitions. Springer, Berlin.

Ikeda, N.; Watanabe, S. 1981: Stochastic differential equations and diffusion processes. North-Holland, Amsterdam.

Ito, K. 1951: On stochastic differential equations. Memoirs Amer. Math. Soc. 4

Jacod, J.; Shirayaev, A.N. 1987: Limit theorems for stochastic processes. Springer, Berlin

Jannsen, R. 1982: Difference-methods for stochastic differential equations. Preprint no. 50, FB Mathematik, Universitaet Kaiserslautern

Jannsen, R. 1984: Discretization of the Wiener process in difference methods for stochastic differential equations. Stoch. Processes Appl. 18, 361-369

Kallianpur, B. 1980: Stochastic filtering theory. Springer, Berlin

Karatzas, I.; Shreve, S.E. 1988: Brownian motion and stochastic calculus. Springer, Berlin

Kimura, M.; Ohta, T. 1971: Theoretical aspects of population genetics. Princeton University Press, Princeton

Klauder, J.R.; Petersen, W.P. 1985: Numerical integration of multiplicative-noise stochastic differential equations. SIAM J. Numer. Anal. 22, 1153-1166

Kloeden, P.E.; Pearson, R.A. 1977: The numerical solution of stochastic differential equations. J. Austral. Math. Soc., Series B, 20, 8-12.

Kloeden, P.E.; Liske, H.; Platen, E. 1990: The numerical solution of stochastic differential equations through computer experiments. (to appear)

Kloeden, P.E.; Platen, E. 1990: The numerical solution of stochastic differential equations. (to appear)

Kozin, F. 1977: An approach to characterizing, modeling and analyzing earthquake excitation records. CISM Lecture Notes Vol. 225, pp. 77-109. Springer, Berlin

Kushner, H.J. 1977: Probability methods for approximations in stochastic control and for elliptic equations. Academic press, New York.

Kushner, H.J.; Di Masi, G. 1978: Approximation for functionals and optimal control on jump diffusion processes. J. Math. Anal. Appl. 63, 772-800.

Leblond, J.; Talay, D. 1986: Simulation of diffusion processes with PRESTO building systems like PRESTO with ADAGIO. Proc. French- Belgian Satisticians' Congress, November 1986, Rouen. Edited by J.P. Raoult

Legland, F. 1981: Estimation de parametres dans les processes stochastiques en observation incomplete: Application a un probleme de radio-astronomie. Dr. Ing. Thesis, L'Universite de Paris IX (Dauphine)

Liske, H.; Platen, E.; Wagner, W. 1982: About mixed multiple Wiener integrals. Preprint P-Math-23/82, IMath, Akad. der Wiss. der DDR, Berlin

Liske, H.; Platen, E. 1987: Simulation studies on time discrete diffusion approximations. In Mathematics and Computers in Simulation 29, 253-260. North-Holland, Amsterdam

Maruyama, G. 1955: Continuous Markov processes and stochastic equations. Rend. Circolo Math. Palemo 4, 48-90

McShane, E.J. 1974: Stochastic calculus and stochastic models. Academic Press, New York

Merton, R.C. 1971: Optimum consumption and portfolio rules in a continuous-time model. J. Econ. Theory 3, 373-413

Mikulevicius R.; Platen, E. 1986: Rate of convergence of the Euler approximation for diffusion processes. Preprint P-Math-38/86, IMath, Akad. der Wiss. der DDR, Berlin

Mikulevicius R.; Platen, E. 1988: Time discrete Taylor approximations for Ito processes with jump component. Math. Nachr. 138, 93-104

Milstein, G.N. 1974: Approximate integration of stochastic differential equations. Theor. Prob. Appl. 19, 557-562

Milstein, G.N. 1978: A method of second-order accuracy for the integration of stochastic differential equations. Theor. Prob. Appl. 23, 396-401

Milstein, G.N. 1986: Weak approximation of the solution of systems of stochastic differential equations. Theor. Prob. Appl. 30, 750-766

Morgan, B.J.J. 1984: Elements of simulation. Chapman and Hall, London

Newton, N.J. 1986: An asymptotically efficient difference formula for solving stochastic differential equations. Stochastics 19, 175-206

Nikitin, N.N.; Rasevig, V.D. 1978: Methods of computer simulation of stochastic differential equation. Vytshical. Matem. i Matem. Fisiki 18, 106-117 (In Russian)

Pardoux, E.; Pignol, M. 1984: Etude de la stabilite de la solution d'une EDS bilinear a coefficients periodiques: Application au movement d'une pale helicopter. Springer Lecture Notes in Control and Inform. Sc. Vol. 63, pp. 92-103

Pardoux, E.; Talay, D. 1985: Discretization and simulation of stochastic differential equations. Acta Appl. Math 3, 23-47

Petersen, W.P. 1987: Numerical simulation of Ito stochastic differential equations on supercomputers. Proc. IMA Conference, 18-24th September 1985

Platen, E. 1980: Approximation of Ito integral equation. Springer Lecture Notes in Control and Inform. Sc. Vol 25, pp. 172-176

Platen, E. 1981: An approximation method for a class of Ito processes. Lietuvos Matem. Rink. 21, 121-133

Platen, E. 1982a: A generalized Taylor formula for solutions of stochastic differential equations. Sankhya, A 44, 163-172

Platen, E. 1982b: An approximation method for a class of Ito processes with jump component, Lietuvos Matem. Rink. 21, 121-133

Platen, E. 1984: Zur zeitdiskreten Approximation von Itoprozessen. Diss. B., IMath, Akad. der Wiss. Der DDR, Berlin

Platen, E.; Rebolledo, R. 1985: Weak convergence of semimartingales and discretization methods. Stoch. Processes Appl. 20, 41-58

Platen, E.; Wagner, W. 1982: On a Taylor formula for a class of Ito processes. Probability Math. Statistics 3, 37-51

Rao, N.J.; Borwankar, J.D.; Ramakrishna, D. 1974: Numerical solution of Ito integral equations. SIAM J. Control 12, 124-139

Roemisch, W.; Wakolbinger, A. 1987: On the convergence rates of approximate solutions of stochastic equations. Preprint, Sektion Mathematik, Humboldt Universitaet, Berlin

Rubinstein, R.Y. 1981: Simulation and the Monte Carlo method. Wiley, New York.

Rumelin, W. 1982: Numerical treatment of stochastic differential equations. SIAM J. Numer. Anal. 19, 604-613

Sagirow, P. 1970: Stochastic methods in the dynamics of statellites. ICMS Lecture Notes Vol. 57. Springer, Berlin

Schenkl, F. 1988: Bericht ueber die Loesung stochastischer Ito-Differentialgleichungen durch Simulation. Preprint, Universitaet der Bundeswehr Muenchen

Schoener, G., Haken, H.; Kelso, J.A.S. 1986: A stochastic theory of phase transitions in human hand movement. Biol. Cybernetics 53, 247-257

Shimizu, A.; Kawachi, T. 1984: Approximate solutions of stochastic differential equations. Bull Nagoya Inst. Tech. 36, 105-108

Shinozuka, M.; Sato, Y. 1967: Simulation of nonstationary random processes. J. Eng. Mech Div. ASCE 93 EM1, 11

Sobcyzk, K. 1986: Modelling of random fatigue crack growth. Eng. Fracture Mech. 24, 609-623

Sussmann, H. 1978: On the gap between deterministic and stochastic differential equations. Ann. Prob. 6, 19-41

Talay, D. 1982: Convegence pour chaque trajectoire d'un schema d'approximation des EDS. Comptes Rendus Acad. Sc. Paris 295, Serie I. 249-252

Talay, D. 1983: Resolution trajectorielle et analysis numerique des equations differentielles stochastiques. Stochastics 9, 275-306

Talay, D 1984: Efficient numerical schemes for the approximation of expectation of functionals of the solution of an SDE and applications. Springer Lecture Notes in Control and Inform Sc. Vol. 61, pp. 294-313

Unny, T.E. 1984: Numerical integration of stochastic differential equations in catchment modeling. Water Resources Res 20, 360-368

Unny, T.E.; Karmeshu 1983: Stochastic nature of outputs from conceptual reservoir model cascades. J. Hydrol. 68, 161-180

Van Kampen, N.G. 1981: Stochastic processes in physics and chemistry. North-Holland, Amsterdam.

Viterbi, A.J. 1966: Principles of coherent communication. McGraw-Hill, New York.

Wagner, W. 1987: Unbiased Monte Carlo evaluation of certain functional integrals. J. Comput. Physics 71, 21-23

Wagner, W. 1988: Monte Carlo evaluation of functionals of solutions of stochastic differential equations. Variance reduction and numerical examples. Stochastic Anal. Appl. 6, 447-468

Wagner, W.; Platen, E. 1978: Approximation of Ito integral equations. Preprint ZIMM, Akad. der Wiss. der DDR, Berlin

Wright, D.J. 1974: The digital simulation of stochastic differential equations. IEEE Trans. Autom. Control AC-19, 75-76

Wright, D.J. 1980: Digital simulation of Poisson stochastic differential equations. Int. J. Systems Sc. 11, 781-785

Yaglom, A.M. 1980: Application of stochastic differential equations to the description of turbulent equations. Springer Lecture Notes in Control and Inform. Sc. Vol. 25, pp. 1-13

Accepted June 5, 1989