Bol. Soc. Esp. Mat. Apl. n^{o}0 (0000), 1–8

Discrete approximation of stochastic differential equations

Raphael Kruse*

University of Bielefeld

rkruse@math.uni-bielefeld.de

Abstract

It is shown how stochastic Itô-Taylor schemes for stochastic ordinary differential equations can be embedded into standard concepts of consistency, stability and convergence. An appropriate choice of function spaces and norms, in particular a stochastic generalization of Spijker's norm (1968), leads to two-sided estimates for the strong error of convergence under the usual assumptions.

Keywords: *SODE, stochastic differential equations, Itô-Taylor schemes, discrete approximation, bistability, two-sided error estimates, stochastic Spijker norm*
AMS subject classifications: *65C20, 65C30, 65J15, 65L20, 65L70.*

1 Introduction

The invention of Itô-Taylor schemes was a major breakthrough in numerical analysis of stochastic ordinary differential equations (SODEs). We refer to the pioneering book [7] and the influential monographs [9] and [10].

In this paper we show how the strong convergence theory of these schemes can be embedded into the standard framework of consistency, stability and convergence as it is formulated in abstract terms in the theory of discrete approximations (see [14]). Moreover, by a special choice of norms, namely a stochastic version of the deterministic Spijker norm (see [12],[13],[6, Ch.III.8]), we are able to derive two-sided estimates for the strong convergence error.

While our notion of consistency and (numerical) stability goes back to the work of F. Stummel [14] there already exist other concepts in the literature. One can find notions of consistency and local truncation errors in the books [7, 9, 10]. We refer to [3] for a discussion. Other authors, who have considered the question of stability, are for instance [2, 4].

*Department of Mathematics, Bielefeld University, P.O. Box 100131, 33501 Bielefeld, Germany, supported by CRC 701 'Spectral Analysis and Topological Structures in Mathematics'.

To be more precise, we deal with the numerical approximation of $\mathbb{R}^d$-valued stochastic processes X, which satisfy an ordinary Itô stochastic differential equation of the form

$$\begin{aligned} dX(t) &= b^0(t, X(t))dt + \sum_{k=1}^{m} b^k(t, X(t))dW^k(t),\ t \in [0,T], \\ X(0) &= X_0. \end{aligned} \tag{1}$$

We assume that the initial value X_0 has finite second moment. By W^k, $k = 1, \dots, m$, we denote real and pairwise independent standard Brownian motions and we also assume that the drift and diffusion coefficient functions $b^k : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ fulfill the usual global Lipschitz and linear growth conditions such that (1) has a unique solution [1].

Note that the corresponding integral form of the SODE (1) has the representation

$$X(t) = X_0 + \int_0^t b^0(s, X(s))ds + \sum_{k=1}^{m} \int_0^t b^k(s, X(s))dW^k(s), \quad t \in [0,T]. \tag{2}$$

Itô-Taylor schemes are based on an iterated application of Itô's formula on the integrands of (2), provided that all appearing integrals and derivatives exist. Again, we refer to the books [7, 9, 10] for a rigorous derivation.

Let $\mathcal{M}$ be the set of all multi-indices $\alpha = (j_1, \dots, j_l)$, $l \in \mathbb{N}$, $j_i \in \{0, \dots, m\}$, $i = 1, \dots, l$. By $\ell(\alpha) \in \mathbb{N}$ and $n(\alpha) \in \mathbb{N}$ we denote the length of $\alpha \in \mathcal{M}$ and the number of zeros in $\alpha \in \mathcal{M}$ respectively. For $\gamma \in \{\frac{n}{2} :\ n \in \mathbb{N}\}$ consider the finite set of multi-indices (c.f. [7])

$$\mathcal{A}_\gamma = \left\{ \alpha \in \mathcal{M} : 1 \leq \ell(\alpha) + n(\alpha) \leq 2\gamma \text{ or } \ell(\alpha) = n(\alpha) = \gamma + \frac{1}{2} \right\}.$$

For a time grid $0 = t_0 < t_1 < \ldots < t_N = T$ with (for simplicity) equidistant step size $h = \frac{T}{N}$, $N \in \mathbb{N}$, the Itô-Taylor scheme of order γ is given by

$$\begin{aligned} &X_h(t_0) = X_0, \\ &X_h(t_k) = X_h(t_{k-1}) + \textstyle\sum_{\alpha \in \mathcal{A}_\gamma} f_\alpha(t_{k-1}, X_h(t_{k-1})) I_{\alpha,k},\ k \geq 1, \end{aligned} \tag{3}$$

with the iterated (stochastic) integrals

$$I_{\alpha,k} := \int_{t_{k-1}}^{t_k} \int_{t_{k-1}}^{s_1} \cdots \int_{t_{k-1}}^{s_{l-1}} dW^{j_1}(s_l) \dots dW^{j_l}(s_1), \tag{4}$$

where $\alpha = (j_1, \dots, j_l)$ and $dW^0(s) = ds$. For the same α the coefficient function $f_\alpha : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ is defined by

$$f_\alpha(t,x) = (L^{j_1} \cdots L^{j_j} f)(t,x), \tag{5}$$

where $f : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ is the projection with respect to the second coordinate, i.e. $f(t,x) = x$, and the L^k are differential operators of the form

$$L^0 = \frac{\partial}{\partial t} + \sum_{i=1}^{d} b^{0,i} \frac{\partial}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^{d} \sum_{k=1}^{m} b^{k,i} b^{k,j} \frac{\partial^2}{\partial x_i \partial x_j},$$

$$L^k = \sum_{j=1}^{d} b^{k,j} \frac{\partial}{\partial x_j}, \ k = 1, \dots, m.$$

Example 1 *If we choose $\gamma = \frac{1}{2}$ then the set $\mathcal{A}_{\frac{1}{2}}$ just consists of all multi-indices of length* 1*, i.e. $\mathcal{A}_{\frac{1}{2}} = \{(0), (1), \dots (m)\}$, and the coefficient functions f_α simplify to the drift and diffusion coefficient functions of the SODE* (1)*, i.e. $f_{(k)} = b^k$ for $k = 0, \dots, m$. Since $I_{(0),k} = h$ and $I_{(j),k} = W^j(t_k) - W^j(t_{k-1})$, the Itô-Taylor scheme of order $\gamma = \frac{1}{2}$ is the well-known Euler-Maruyama scheme. One also easily checks that the choice $\gamma = 1$ leads to the Milstein method.*

It is well-known (see for example [7, 9, 10]) that the Itô-Taylor scheme of order γ converges at least with order γ in the strong sense, i.e. there exists a constant $C > 0$, independent of the step size h, such that

$$\max_{0 \le i \le N} \left(\mathbb{E} \left(|X(t_i) - X_h(t_i)|^2 \right) \right)^{\frac{1}{2}} \le C h^\gamma, \tag{6}$$

where X is the analytic solution to (1) and X_h denotes the numerical solution. Note that [7, 9, 10] use an even stronger norm, where max occurs inside the expectation. It is an open problem whether our approach can handle this norm as well.

In order to embed the Itô-Taylor scheme into the discrete approximation framework, we will write the equations (3) as $A_h(X_h) = R_h$ with a suitable operator A_h and right-hand side R_h. We use the norm

$$\|Y_h\|_{0,h} = \max_{0 \le i \le N} \|Y_h(t_i)\|_{L^2(\Omega)}, \tag{7}$$

and the following generalization of Spijker's norm

$$\|Y_h\|_{-1,h} = \max_{0 \le i \le N} \|\textstyle\sum_{j=0}^{i} Y_h(t_j)\|_{L^2(\Omega)}. \tag{8}$$

Here $\|\cdot\|_{L^2(\Omega)}$ denotes the L^2-norm of random variables.

The key to our two-sided error estimate is the following bistability inequality

$$C_1 \|A_h(Y_h) - A_h(Z_h)\|_{-1,h} \le \|Y_h - Z_h\|_{0,h} \le C_2 \|A_h(Y_h) - A_h(Z_h)\|_{-1,h}. \tag{9}$$

In the following section we show how the Itô-Taylor scheme fits into the discrete approximation theory. In Section 3 we give a precise formulation of our main result together with all assumptions.

2 Writing Itô-Taylor schemes as discrete approximations

In the discrete approximation theory the concepts of consistency, (numerical) stability and convergence are defined in a very general way. Our notions of bistability and of the local truncation error are directly related to the abstract framework invented by F. Stummel [14]. We present the basic ideas behind Stummel's theory in this section. Simultaneously we embed the Itô-Taylor scheme into the framework.

The starting point of the discrete approximation theory is an equation of the form $A(X) = Y$. Here, the operator $A : E \to F$ is a mapping between two sets E and F. For a given $Y \in F$ our aim is to find a discrete approximation of the solution X. To this end we assume the existence of two sequences of metric spaces $(E_h)_{h\in\mathcal{I}}$ and $(F_h)_{h\in\mathcal{I}}$ and operators $A_h : E_h \to F_h$, $h \in \mathcal{I}$, for some index set $\mathcal{I}$. With the help of two sequences of restriction operators $r_h^E : E \to E_h$ and $r_h^F : F \to F_h$, for $h \in \mathcal{I}$, the discrete spaces E_h and F_h are connected to the original spaces E and F respectively. Figure 1 visualizes the setting.

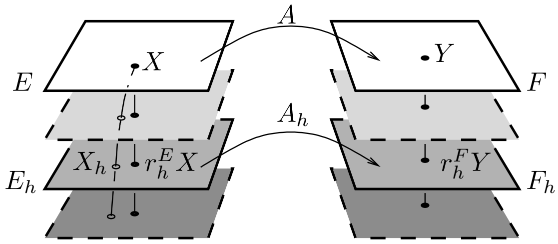

Figure 1: Visualisation of the discrete approximation theory

By solving equations of the form $A_h(X_h) = r_h^F Y$ we obtain a sequence of discrete approximations $(X_h)_{h\in\mathcal{I}}$. Now, the theory of F. Stummel answers the questions, in which sense and under which conditions the sequence $(X_h)_{h\in\mathcal{I}}$ converges to the solution X. Let us first show how the SODE (1) and the Itô-Taylor scheme (3) can be embedded into figure 1.

Since the existence of a unique solution X to (1) is guaranteed by our assumptions we consider the trivial operator

$$A : \quad \begin{aligned} E &\to F \\ X &\mapsto A(X) \end{aligned} \tag{10}$$

where $E := \{X\}$ and $F := \{Y = (X_0, 0)\}$ are singletons (with the second component of Y being the stochastic process which is P-a.s. equal to $0 \in \mathbb{R}^d$) and the operator A is given by

$$A(X) = \Big(X(0), \\ \Big(X(t) - X(0) - \int_0^t b^0(s, X(s))ds - \sum_{k=1}^m \int_0^t b^k(s, X(s))dW^k(s) \Big)_{0\le t\le T} \Big).$$

In order to define the discrete metric spaces we denote the time grid by $\tau_h := \{t_i = ih \,|\, i = 0, \ldots, N\}$. As our underlying discrete space we consider the set $\mathcal{G}_h := \mathcal{G}(\tau_h, L^2(\Omega, \mathcal{F}, P; \mathbb{R}^d))$ of all adapted and $L^2(\Omega)$-valued grid functions, that is, for $Z_h \in \mathcal{G}_h$, the random variables $Z_h(t_i)$ are square-integrable and $\mathcal{F}_{t_i}$-measurable random variables for all $t_i \in \tau_h$. Here $(\mathcal{F}_t)_{t\in[0,T]}$ denotes the filtration which is generated by the Wiener processes W^k, $k = 1, \ldots, m$. Now, we choose the metric spaces E_h and F_h to be the vector space $\mathcal{G}_h$ endowed with the metric induced by the norm

$$\|Z_h\|_{0,h} = \max_{0\le i\le N} \|Z_h(t_i)\|_{L^2(\Omega)} \tag{11}$$

and the stochastic version of Spijker's norm

$$\|Z_h\|_{-1,h} = \max_{0\le i\le N} \|\textstyle\sum_{j=0}^{i} Z_h(t_j)\|_{L^2(\Omega)}, \tag{12}$$

respectively. Note that E_h and F_h are Banach spaces.

Next, define the two sequences of restriction operators

$$r_h^E : \quad \begin{array}{l} E \to E_h \\ X \mapsto r_h^E X, \quad [r_h^E X](t_i) = X(t_i) \quad \text{for} \quad t_i \in \tau_h, \end{array} \tag{13}$$

$$r_h^F : \quad \begin{array}{l} F \to F_h \\ Y \mapsto r_h^F Y \end{array} \qquad [r_h^F Y](t_i) = \left\{ \begin{array}{cc} X_0 & i = 0, \\ 0 & i = 1, \ldots N. \end{array} \right. \tag{14}$$

Finally, for $h > 0$, we introduce the operator

$$A_h : \quad \begin{array}{l} E_h \to F_h \\ X_h \mapsto A_h(X_h) \end{array}$$

by the relationship

$$\begin{array}{l} [A_h(X_h)](t_0) = X_h(t_0), \\ [A_h(X_h)](t_i) = X_h(t_i) - X_h(t_{i-1}) - \sum_{\alpha\in\mathcal{A}_\gamma} f_\alpha(t_{i-1}, X_h(t_{i-1})) I_{\alpha,i}, \end{array} \tag{15}$$

for $1 \le i \le N$. Under the assumption that all Itô-Taylor coefficient functions f_α satisfy a linear growth condition, $[A_h(X_h)](t_i)$ is an adapted and mean-square integrable random variable. Therefore, A_h maps E_h into F_h. See Section 3 for a complete statement of all assumptions.

Since the Itô-Taylor schemes are explicit, the operators A_h are bijective, i.e. there exists a unique solution $\tilde{X}_h$ to the equation $A_h(\tilde{X}_h) = Z_h$ for all $Z_h \in F_h$. In particular, the Itô-Taylor approximation X_h to (1) is equivalently written as the solution to the equation $A_h(X_h) = r_h^F Y$.

Next, we introduce our notion of consistency, bistability and convergence.

Definition 1 *Consider a one-step method given by a sequence of operators $(A_h)_h$. The method is called consistent of order $\gamma > 0$, if there exists a constant $C > 0$ and an upper step size bound $\overline{h} > 0$, such that the estimate*

$$\|A_h(r_h^E X) - r_h^F A(X)\|_{-1,h} \le Ch^\gamma \tag{16}$$

holds for all grids τ_h with $h \le \overline{h}$, where X denotes the analytic solution of (1).

The left hand side of (16) is called *local truncation error* or *consistency error*. Therefore, a one-step method is consistent if the diagram in Figure 1 commutes up to an error of order γ, that is $r_h^F \circ A \approx A_h \circ r_h^E$ for h small enough.

The second ingredient in the convergence theory is the concept of (numerical) stability. In [14] F. Stummel introduces the stronger notion of bistability and he proves that bistability of a numerical method can be characterized by the equicontinuity of the operators $(A_h)_h$ and $(A_h^{-1})_h$. In this sense the following definition is a sufficient condition for Stummel's notion of bistability.

Definition 2 *A one-step method defined by operators $(A_h)_h$ is called bistable, if there exist constants $C_1, C_2 > 0$ and an upper step size bound $\overline{h} > 0$ such that the operators A_h are bijective and the estimate*

$$C_1\|A_h(Z_h) - A_h(\tilde{Z}_h)\|_{-1,h} \le \|Z_h - \tilde{Z}_h\|_{0,h} \le C_2\|A_h(Z_h) - A_h(\tilde{Z}_h)\|_{-1,h}$$

holds for all $Z_h, \tilde{Z}_h \in E_h$ and for grids τ_h with $h < \overline{h}$.

Finally, we define the error of convergence in terms of the norm $\|\cdot\|_{0,h}$, the space E_h and the restriction operators r_h^E.

Definition 3 *A one-step method is called convergent of order $\gamma > 0$ if there exist an upper step size bound $\overline{h} > 0$ and a constant $C > 0$ such that the corresponding operators A_h are bijective and*

$$\|X_h - r_h^E X\|_0 \le Ch^\gamma \tag{17}$$

for all $h \le \overline{h}$. Here X_h denotes the solution to $A_h(X_h) = r_h^F Y$.

3 Main result

In this section we give a precise formulation of the underlying assumptions and our main result.

(A1) The initial value X_0 is an $\mathcal{F}_0$-measurable and $\mathbb{R}^d$-valued random variable satisfying $\mathbb{E}(|X_0|^2) < \infty$.

(A2) For all $\alpha \in \mathcal{A}_\gamma$ there exists a constant $L_\alpha > 0$ such that

$$|f_\alpha(t,x) - f_\alpha(t,y)| \le L_\alpha|x-y| \quad \text{and} \quad |f_\alpha(t,x)| \le L_\alpha(1+|x|)$$

for all $x, y \in \mathbb{R}^d$ and $t \in [0,T]$.

(A3) For a given order γ the Itô-Taylor expansion of $X(t)$ with respect to $\mathcal{A}_\gamma$ exists for all $t \in [0, T]$.

(A4) For all $\alpha \in \mathcal{B}(\mathcal{A}_\gamma)$ we have

$$\int_0^T \mathbb{E}\left(|f_\alpha(s, X(s))|^2\right) ds < \infty.$$

The first two assumptions are used, for example, in [1] to assure the existence and uniqueness of the solution X on $[0, T]$, such that $X(t)$ is mean-square integrable for all $t \in [0, T]$. The assumption (A2) also assures that the operators A_h are well-defined and bistable. In (A3) we assume that the Itô-Taylor expansion exists up to a given order γ. Assumption (A4) is needed in order to prove the consistency of the Itô-Taylor schemes. There we use the notation of the remainder set $\mathcal{B}(\mathcal{A}_\gamma)$ of the Itô-Taylor expansion which is given by

$$\mathcal{B}(\mathcal{A}_\gamma) = \{\alpha = (j_1, j_2, \ldots, j_l) \in \mathcal{M} \,:\, (j_2, \ldots, j_l) \in \mathcal{A}_\gamma\} \subset \mathcal{M}$$

(c.f. [7]). Now we formulate our main result, which is proven in [8].

Theorem 1 *Let the assumptions (A1)-(A4) hold for $\gamma \in \{\frac{n}{2} \,|\, n \in \mathbb{N}\}$. Then the Itô-Taylor scheme of order γ is*

(i) consistent of order γ,

(ii) bistable with respect to the norms $\|\cdot\|_{0,h}$ and $\|\cdot\|_{-1,h}$,

(iii) convergent of order γ.

Moreover, there exists $\overline{h} > 0$ such that the two-sided error estimate

$$C_1\|A_h(r_h^E X) - r_h^F Y\|_{-1,h} \le \|r_h^E X - X_h\|_{0,h} \le C_2\|A_h(r_h^E X) - r_h^F Y\|_{-1,h}$$

holds for all grids τ_h with $|h| \le \overline{h}$.

Remark 1 *Theorem 1 also holds for implicit methods like the stochastic theta method [3] and for stochastic multi-step methods [8].*

Remark 2 *The two-sided error estimate in Theorem 1 can be used to discuss the optimal order of convergence of the Itô-Taylor methods. J. M. C. Clark and R. J. Cameron [5] constructed the example*

$$dX(t) = \begin{pmatrix} 1 & 0 \\ 0 & X_1(t) \end{pmatrix} d\begin{pmatrix} W^1(t) \\ W^2(t) \end{pmatrix}, \quad X(0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \tag{18}$$

to show that, in general, the maximum order of convergence is equal to $\frac{1}{2}$ if the numerical method, like the Euler-Maruyama scheme, uses only the increments $W^k(t_i) - W^k(t_{i-1})$ of the driving Wiener processes. For this example the local trunction error of the Euler-Maruyama is exactly computed to be $\sqrt{\frac{1}{2}Th}$. Hence, the strong error of convergence is bounded from below by a term of order $\gamma = \frac{1}{2}$.

A suitable generalization of this example gives corresponding results for the higher order schemes [8].

References

[1] L. Arnold. *Stochastic differential equations*. Wiley, New York, 1974.

[2] C. T. H. Baker and E. Buckwar. Numerical analysis of explicit one-step methods for stochastic delay differential equations. *LMS J. Comput. Math.*, 3:315–335, 2000.

[3] W-J. Beyn and R. Kruse. Two-sided error estimates for the stochastic theta method. Technical report, CRC 701, Bielefeld University, 2009. (submitted).

[4] E. Buckwar and R. Winkler. Multi-step methods for SDEs and their application to problems with small noise. *SIAM J. Numer. Anal.*, 44(2):779–803, 2006.

[5] J. M. C. Clark and R. J. Cameron. The maximum rate of convergence of discrete approximations for stochastic differential equations. In B. Grigelionis, editor, *Stochastic Differential Systems – Filtering and Control*, volume 25 of *Lecture Notes in Control and Information Sciences*. Springer-Verlag, 1980.

[6] E. Hairer, S. P. Nørsett, and G. Wanner. *Solving Ordinary Differential Equations I Nonstiff Problems*. Springer, 1993. (2nd ed.).

[7] P. E. Kloeden and E. Platen. *Numerical solution of stochastic differential equations*. Applications of mathematics ; 23. Springer, 1992.

[8] R. Kruse. Two-sided error estimates for stochastic one- and multistep methods of higher order. Technical report, CRC 701, Bielefeld University, 2009. (in preparation).

[9] G. N. Milstein. *Numerical integration of stochastic differential equations*. Kluwer, Dordrecht, 1995. (translated and revised from the 1988 Russian original).

[10] G. N. Milstejn and M. V. Tretjakov. *Stochastic numerics for mathematical physics*. Springer, Berlin [u.a.], 2004.

[11] B. K. Øksendal. *Stochastic differential equations, an introduction with applications*. Springer, Berlin [u.a.], 2007.

[12] M. N. Spijker. On the structure of error estimates for finite difference methods. *Numer. Math.*, 18:73–100, 1971.

[13] H. J. Stetter. *Analysis of discretization methods for ordinary differential equations*. Springer, Berlin, 1973.

[14] F. Stummel. *Approximation methods in analysis*, volume 35 of *Lecture Notes Series*. Aarhus: Mat. Inst., Aarhus, 1973.