TWO-SIDED ERROR ESTIMATES FOR THE STOCHASTIC THETA METHOD

WOLF-JÜRGEN BEYN* AND RAPHAEL KRUSE*

Abstract. Two-sided error estimates are derived for the strong error of convergence of the stochastic theta method. The main result is based on two ingredients. The first one shows how the theory of convergence can be embedded into standard concepts of consistency, stability and convergence by an appropriate choice of norms and function spaces. The second one is a suitable stochastic generalization of Spijker's norm (1968) that is known to lead to two-sided error estimates for deterministic one-step methods. We show that the stochastic theta method is bistable with respect to this norm and that well-known results on the optimal $\mathcal{O}(\sqrt{h})$ order of convergence follow from this property in a natural way.

1. Introduction

There is a well established theory for the strong convergence of one-step methods for stochastic ordinary equations (SODEs), cf. [6],[10],[5]. The purpose of this paper is to add two new aspects to this theory.

The first one is of conceptual type. By a proper choice of norms and functions spaces we show that strong convergence results 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 [15, 16, 17, 18], [14]). In Section 3 we will discuss why and in which sense our notions deviate from those used in [6], [10].

Our second contribution is concerned with a special choice of norms that allows to prove bistability in the sense of [18] and, as a consequence, to derive two-sided estimates for the convergence error. We show that this can be achieved by a suitable stochastic version of the deterministic Spijker norm (see [13], [14, Ch.2.2], [4, Ch.III.8]). It also turns out that well-known results on the optimal order of convergence [3] follow in a natural way from these two-sided estimates.

In order to present the basic ideas in a simplified technical framework we consider only the semi-implicit Euler method [6] or stochastic theta method (STM) [5]. Extensions of the results to other methods (in particular higher order methods) will be published in a subsequent paper. In the following we give a more technical outline of the paper.

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

2000 *Mathematics Subject Classification.* Primary 65C20, 65C30, 65L20 ; Secondary 65L70, 65J15.

Key words and phrases. SODE, stochastic theta method, two-sided error estimate, consistency, stochastic Spijker norm.

We consider the numerical approximation of $\mathbb{R}^d$-valued stochastic processes, which satisfy an ordinary Itô stochastic differential equation [1, 9, 11] of the form

$$\begin{aligned} dX(t) &= b^0(t, X(t))dt + \sum_{k=1}^{m} b^k(t, X(t))dW^k(t),\ t \in [0,T], \\ X(0) &= X_0, \end{aligned} \tag{1.1}$$

where $W^k, k = 1, \dots, m$ denote real and pairwise independent standard Brownian motions, adapted to the filtration $(\mathcal{F}_t)_{t\in[0,T]}$ on the underlying probability space $(\Omega, \mathcal{F}, P)$. The initial value X_0 is assumed to be $\mathcal{F}_0$-measurable and to have finite second moment. We also assume that the drift and diffusion coefficient functions $b^k : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ are measurable and fulfill the usual Lipschitz conditions such that (1.1) has a unique solution (see Section 2 for details).

The stochastic theta method ($\theta \in [0,1]$) on a grid

$$\tau_h = \{t_i : i = 0, \dots, N\}, \quad 0 = t_0 < t_1 < \cdots < t_{N-1} < t_N = T \tag{1.2}$$

is given by the recursion

$$\begin{aligned} X_h(t_i) &= X_h(t_{i-1}) + h_i \left((1-\theta)b^0(t_{i-1}, X_h(t_{i-1})) + \theta b^0(t_i, X_h(t_i))\right) \\ &\quad + \sum_{k=1}^{m} b^k(t_{i-1}, X_h(t_{i-1}))\Delta_h W^k(t_i), \\ X_h(0) &= X_0, \end{aligned} \tag{1.3}$$

where $h_i = t_i - t_{i-1}$ is the length of the i-th interval and $\Delta_h W^k(t_i) = W^k(t_i) - W^k(t_{i-1})$ denotes the i-th increment of the Wiener process W^k. We collect step-sizes and define

$$h = (h_i)_{i=1}^N \in \mathbb{R}^N, \quad |h| = \max_{i=1,\dots,N} h_i. \tag{1.4}$$

It is well-known (see for example [6, 10]) that the STM converges at least with order $\gamma = \frac{1}{2}$ in the strong sense, i.e. there exists a constant $C > 0$ such that

$$\max_{0\le i\le N} \left(\mathbb{E}\left(|X(t_i) - X_h(t_i)|^2\right)\right)^{\frac{1}{2}} \le C|h|^\gamma, \tag{1.5}$$

where X is the analytic solution and X_h is the numerical solution. Moreover, it was shown by J.M.C. Clark and R.J. Cameron [3] that, in general, $\gamma = \frac{1}{2}$ is the maximum rate of convergence for the STM (and for any method that only uses the Brownian motion at grid points).

The strong convergence in (1.5) is written in terms of the norm

$$\|Y_h\|_{0,h} = \max_{0\le i\le N} \|Y_h(t_i)\|_{L^2(\Omega)}. \tag{1.6}$$

In this paper we show that the following generalization of Spijker's norm also plays an important role

$$\|Y_h\|_{-1,h} = \max_{0\le i\le N} \|\textstyle\sum_{j=0}^{i} Y_h(t_j)\|_{L^2(\Omega)}. \tag{1.7}$$

Writing the equations (1.3) as $A_h(X_h) = R_h$ with a suitable operator A_h and right-hand side R_h, one of our main results is the following bistability inequality

$$C_1\|A_h(Y_h) - A_h(Z_h)\|_{-1,h} \le \|Y_h - Z_h\|_{0,h} \le C_2\|A_h(Y_h) - A_h(Z_h)\|_{-1,h}. \tag{1.8}$$

A precise formulation and the proof will be given in Sections 3 and 5.

In Section 2 we summarize the main assumptions and collect some prerequisites from stochastic analysis. Then the main results on two-sided error estimates are stated in Section 3. In Section 4 we show that the STM is consistent with respect to the stochastic Spijker norm (1.7). We conclude the paper in Section 6 by showing that the results on optimal convergence rates from [3] follow directly from our two-sided error estimates.

2. Main assumptions and some results from stochastic analysis

In this section we collect the main assumptions and some useful results from stochastic analysis.

Let $(\Omega, \mathcal{F}, P)$ be the underlying probability space and denote by $\mathbb{E}$ the expectation with respect to P. As in [1, 9, 11] we assume for the SODE (1.1) that the drift and diffusion coefficient functions $b^k : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$, $k = 0, \dots, m$, are measurable. We also assume that the following assumptions hold:

(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): There exists a constant $K > 0$ such that

$$|b^k(t,x)| \leq K(1+|x|)$$

and

$$|b^k(t,x) - b^k(t,y)| \leq K|x-y|$$

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

(A3): There exists a constant $K > 0$ such that

$$|b^k(t,x) - b^k(s,x)| \leq K(1+|x|)\sqrt{|t-s|}$$

for all $k = 0, \dots, m$, $t, s \in [0,T]$ and $x \in \mathbb{R}^d$.

Here we denote by $|\cdot|$ the Euclidean norm in $\mathbb{R}^d$. Assumptions (A1) and (A2) are sufficient to assure the existence and uniqueness of a strong Itô solution to (1.1) (see [1, 9, 11]), i.e. there exists a unique, P-a.s. continuous and $(\mathcal{F}_t)_{t\in[0,T]}$-adapted process X which satisfies

$$X(t) = X_0 + \int_0^t b^0(s, X(s))ds + \sum_{k=1}^m \int_0^t b^k(s, X(s))dW^k(s) \tag{2.1}$$

for all $t \in [0,T]$ and

$$\mathbb{E}\left(\int_0^T |X(s)|^2 ds\right) < \infty. \tag{2.2}$$

Assumption (A3) is used in [6] to prove convergence of the Euler-Maruyama scheme. We will use it here to prove consistency of the STM. We note that assumption (A3) can be replaced by an L^2-condition on the second order Itô-Taylor coefficient function (see [8] for further details).

From Theorem 7.1.2 and Remark 7.1.5 in [1] we have the following estimates for the second moment of the strong solution.

Theorem 2.1. *Under the assumptions (A1) and (A2) the solution X to (1.1) satisfies*

$$\mathbb{E}\left(|X(t)|^2\right) \le \left(1+\mathbb{E}(|X_0|^2)\right) e^{Ct}$$

and

$$\mathbb{E}\left(|X(t)-X_0|^2\right) \le C\left(1+\mathbb{E}(|X_0|^2)\right) te^{Dt}$$

for all $0 \le t \le T$ and some constants $C, D > 0$ depending only on K and T.

In particular, using the semigroup property of X, one can prove the estimate

$$\mathbb{E}\left(|X(t)-X(s)|^2\right) \le C|t-s|$$

for all $t, s \in [0,T]$ and some constant $C > 0$ depending only on K, T and $\mathbb{E}(|X_0|^2)$.

3. Definitions and main result

In this section we rewrite the STM as an operator equation and introduce the corresponding spaces and norms. We give precise definitions of our notions of consistency and (numerical) stability and compare them to related notions in the literature. Finally, the main convergence theorem and the two-sided error estimate are stated and proved. We note that our notions are largely motivated by the work of Stummel [18]. It will not be necessary to directly invoke results from [18], but in the remark following Definition 3.2 (see also [7]) we will indicate how a formal embedding of our approach into the abstract framework can be achieved.

3.1. **Basic notions.** We introduce an operator that represents the SODE (1.1). Since a unique solution X to (1.1) is guaranteed by assumptions (A1) and (A2) we consider the trivial operator

$$A: \begin{array}{l} E \to F \\ X \mapsto AX \end{array} \tag{3.1}$$

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

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

With each grid τ_h as in (1.2) we associate the space $\mathcal{G}_h := \mathcal{G}(\tau_h, L^2(\Omega, \mathcal{F}, P; \mathbb{R}^d))$ 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 $\mathcal{F}_{t_i}$-measurable and lie in $L^2(\Omega, \mathcal{F}_{t_i}, P; \mathbb{R}^d)$ for all $t_i \in \tau_h, i = 0, \dots, N$. Note that $\mathcal{G}_h$ is a Banach space with respect to the norm (1.6).

Next we define two sequences of restriction operators on the spaces E and F

$$r_h^E: \begin{array}{l} E \to \mathcal{G}_h \\ X \mapsto r_h^E X, \quad [r_h^E X](t_i) = X(t_i) \quad \text{for} \quad t_i \in \tau_h, \end{array} \tag{3.2}$$

$$r_h^F: \begin{array}{ll} F \to \mathcal{G}_h & \\ Y \mapsto r_h^F Y & [r_h^F Y](t_i) = \begin{cases} X_0 & i = 0, \\ 0 & i = 1, \dots N. \end{cases} \end{array} \tag{3.3}$$

In this and the next section we consider three real parameters $L, \rho, \sigma > 0$ that are arbitrary. They will be given specific values later in Section 5 when we prove stability. The first parameter occurs in the norm for grid functions $Z_h \in \mathcal{G}_h$ given by

$$\begin{aligned} \|Z_h\|_{0,h,L} &:= \max_{0\le i\le N} \left(\mathbb{E}\left(|Z_h(t_i)|^2\right)\right)^{\frac{1}{2}} e^{-Lt_i} \\ &= \max_{0\le i\le N} \|Z_h(t_i)\|_{L^2(\Omega)}\, e^{-Lt_i}, \end{aligned} \tag{3.4}$$

where $\|\cdot\|_{L^2(\Omega)}$ denotes the norm in $L^2(\Omega, \mathcal{F}, P; \mathbb{R}^d)$. As usual the exponential weight is needed for the proof of stability via a contraction argument for the Picard iteration. The weight plays no role in the proof of consistency of the STM. Whenever appropriate we abbreviate $\|\cdot\|_0 = \|\cdot\|_{0,h,L}$.

Our underlying complete metric space is the closed ball

$$E_h := \overline{B}_{\rho,L}(r_h^E X) = \left\{Z_h \in \mathcal{G}_h : \|Z_h - r_h^E X\|_{0,h,L} \le \rho\right\} \subset \mathcal{G}_h. \tag{3.5}$$

Note also that the norms $\|\cdot\|_{0,h,L}$ and (1.6) are equivalent (uniformly in h).

We also introduce a second norm for $Z_h \in \mathcal{G}_h$ by

$$\begin{aligned} \|Z_h\|_{-1,h,L} &:= \max_{0\le i\le N} \left(\mathbb{E}\left(\left|\textstyle\sum_{j=0}^{i} Z_h(t_j)\right|^2\right)\right)^{\frac{1}{2}} e^{-Lt_i} \\ &= \max_{0\le i\le N} \left\|\textstyle\sum_{j=0}^{i} Z_h(t_j)\right\|_{L^2(\Omega)} e^{-Lt_i}. \end{aligned} \tag{3.6}$$

Let us write $\|\cdot\|_{-1,h,L} = \|\cdot\|_{-1}$ for short and note that $\|\cdot\|_{-1,h,L}$ is uniformly in h equivalent to the norm (1.7). From the introduction we also recall that these norms are stochastic versions of Spijker's norm [12, 14]. Then our second complete metric space is the closed ball

$$F_h := \overline{B}_{\sigma,L}(r_h^F Y) = \left\{Z_h \in \mathcal{G}_h : \|Z_h - r_h^F Y\|_{-1,h,L} \le \sigma\right\} \subset \mathcal{G}_h. \tag{3.7}$$

In the next step we introduce the residual mapping $\mathrm{res}_h : \mathcal{G}_h \to \mathcal{G}_h$ of the stochastic theta method (1.3) as follows:

$$\begin{aligned} \mathrm{res}_h(Z_h)(t_0) &= Z_h(t_0) - X_0, \\ \mathrm{res}_h(Z_h)(t_i) &= Z_h(t_i) - Z_h(t_{i-1}) - h_i\left((1-\theta)b^0(t_{i-1}, Z_h(t_{i-1})) + \theta b^0(t_i, Z_h(t_i))\right) \\ &\quad - \sum_{k=1}^{m} b^k(t_{i-1}, Z_h(t_{i-1}))\Delta_h W^k(t_i),\ 1 \le i \le N. \end{aligned}$$

The definition shows that $\mathrm{res}_h(Z_h)$ is adapted to the filtration $(\mathcal{F}_{t_i})_{t_i\in\tau_h}$ and assumption (A2) implies that $\mathrm{res}_h(Z_h)(t_i)$ is square-integrable. Therefore, res_h maps $\mathcal{G}_h$ into $\mathcal{G}_h$.

A direct comparison shows that the stochastic theta method (1.3) and the residual condition $\mathrm{res}_h(X_h) = 0 \in \mathcal{G}_h$ are equivalent. This leads to the idea of taking the residual $\mathrm{res}_h(r_h^E X)$, where X is the solution of (2.1), as a measure of the local truncation error.

We are now in the position to define the sequence of operators which represent the stochastic theta method. Define the operator

$$A_h : \begin{array}{c} D(A_h) \subset E_h \to F_h \\ Z_h \mapsto A_h Z_h \end{array} \tag{3.8}$$

on its domain of definition

$$D(A_h) = \left\{ Z_h \in E_h \ : \ \|\mathrm{res}_h(Z_h)\|_{-1} \le \sigma \right\} \tag{3.9}$$

by the relation

$$A_h Z_h = \mathrm{res}_h(Z_h) + r_h^F Y. \tag{3.10}$$

We recall that $r_h^F Y = (X_0, 0, \ldots, 0)$ by (3.3). Thus $A_h Z_h$ and $\mathrm{res}_h(Z_h)$ differ only at the first grid point.

Using the operators A_h, the stochastic theta method (1.3) can equivalently be written as the equation $A_h X_h = r_h^F Y$. Still we have to prove solvability of this equation and to estimate the global error $\|r_h^E(X) - X_h\|_0$. This motivates the following definitions.

Definition 3.1. *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*

$$\left\| A_h r_h^E X - r_h^F A X \right\|_{-1,h,L} \le C|h|^\gamma \tag{3.11}$$

holds for all grids τ_h with $|h| \le \overline{h}$, where X denotes the analytic solution of (1.1).

Note that

$$\left\| A_h r_h^E X - r_h^F A X \right\|_{-1} = \left\| \mathrm{res}_h(r_h^E X) \right\|_{-1},$$

and, therefore, we refer to the left hand side of (3.11) as the *local truncation error* or the *consistency error*. The local truncation error is meaningful even if $r_h^E X \notin D(A_h)$, but for a consistent one-step method we know that $\|\mathrm{res}_h(r_h^E X)\|_{-1} \to 0$ as $|h| \to 0$ and hence $r_h^E X \in D(A_h)$ for $|h|$ sufficiently small.

The second ingredient in a convergence theory for numerical methods is the concept of (numerical) stability. We use here the stronger notion of bistability.

Definition 3.2. *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 $A_h : D(A_h) \to F_h$ are bijective and the estimate*

$$C_1 \left\| A_h Z_h - A_h \tilde{Z}_h \right\|_{-1,h,L} \le \left\| Z_h - \tilde{Z}_h \right\|_{0,h,L} \le C_2 \left\| A_h Z_h - A_h \tilde{Z}_h \right\|_{-1,h,L} \tag{3.12}$$

holds for all $Z_h, \tilde{Z}_h \in D(A_h)$ and for all grids τ_h with $|h| \le \overline{h}$.

Remarks. 1. Our notions of consistency and stability are directly related to the abstract framework invented by F. Stummel [18]. In the general theory of discrete approximations he proves that bistability of a numerical method can be characterized by the equicontinuity of the operators $(A_h)_h$ and the equicontinuity of $(A_h^{-1})_h$. It is easy to see that our definition of bistability is a sufficient condition for bistability in the sense of [18]. The same is true for the consistency error: Our definition of consistency appears in [18, §2, (6)] as a sufficient condition for Stummel's notion of consistency. We refer to [7] for a derivation of the results of this paper from the theory of discrete approximations. Our current definitions turn out to be convenient for providing a direct approach to the convergence of numerical methods for stochastic differential equations.

2. Note that the definition of bistability depends on the three constants L, ρ, σ which appear in the norms in (3.12) and in the domain of definition $D(A_h)$ via (3.5), (3.9).

3. One obtains an equivalent version of our notions when dividing equation (1.3) by h_i and defining the operators res_h and A_h accordingly. Then the stochastic Spijker norm (3.6) is replaced by

$$\|Z_h\|_{-1,h,L} = \max\left(\|Z_h(0)\|_{L^2(\Omega)}, \max_{1\le i\le N} \left\| \sum_{j=1}^{i} h_j Z_h(t_j) \right\|_{L^2(\Omega)} e^{-Lt_i} \right).$$

This form shows that the norm is of $W^{-1,\infty}$ Sobolev type.

3.2. **Main results.** Now we formulate the main results. The proofs of the first two theorems will be deferred to the following two sections.

Theorem 3.3. *Under the assumptions (A1)-(A3) the stochastic theta method* (1.3) *is consistent with order* $\gamma = \frac{1}{2}$.

Theorem 3.4. *Under the assumptions (A1)-(A3) there exist parameter values* $L, \rho, \sigma > 0$ *such that the stochastic theta method* (1.3) *is bistable.*

Theorem 3.5. *Let the assumptions (A1)-(A3) hold and choose parameter values* $L, \rho, \sigma > 0$ *such that the stochastic theta method is bistable. Then the two-sided error estimate*

$$C_1 \left\| A_h r_h^E X - r_h^F A X \right\|_{-1} \le \left\| r_h^E X - X_h \right\|_0 \le C_2 \left\| A_h r_h^E X - r_h^F A X \right\|_{-1} \tag{3.13}$$

holds for all grids τ_h *with* $|h| \le \overline{h}$. *In particular, the numerical solution* X_h *of the stochastic theta method converges uniformly with order* $\gamma = \frac{1}{2}$ *at each grid point to the restriction of the analytic solution* X *to* (1.1).

Proof. By theorem 3.4 we know that there exists an upper step size bound $\overline{h} > 0$ such that the operators $A_h : D(A_h) \to F_h$ are bijective for $|h| \le \overline{h}$. Thus the equation $A_h X_h = r_h^F A X = r_h^F Y$ has a unique solution for $|h| \le \overline{h}$. After possibly reducing $\overline{h}$ further, the consistency of the STM shows that $r_h^E X \in D(A_h)$ for all $|h| \le \overline{h}$. Then the two-sided error estimate (3.13) follows directly from (3.12) by setting $Z_h = r_h^E X$ and $\tilde{Z}_h = X_h$. □

Remark. In our approach we avoid any interpolation and work with grid functions only. Therefore, our estimates of the mean-square error

$$\mathbb{E}\left(|X(t_i) - X_h(t_i)|^2 \right), \quad t_i \in \tau_h$$

are also restricted to grid points. According to [6, Ch.10.2 (2.16)] one can interpolate the numerical approximation to an adapted, right-continuous stochastic process with existing left limits on the interval $[0, T]$ by

$$X_h(t) = X_h(t_i) + \int_{t_i}^{t} b^0(t_i, X_h(t_i))ds + \sum_{k=1}^{m} \int_{t_i}^{t} b^k(t_i, X_h(t_i))dW^k(s), \quad t \in (t_i, t_{i+1}). \tag{3.14}$$

In the case $\theta = 0$ this interpolation is continuous. In [6, Ch.10.2] it is also shown that this interpolation converges uniformly to $X(t), t \in [0, T]$ with the same order that holds at the grid points. The interpolation above can be considered as an a-posteriori process and formula (3.14) may be called a method for dense output, see [4].

3.3. **Comparison with other consistency concepts.** In this subsection we compare, for the case of constant step-size h, our notion of consistency to other approaches in the literature. First, we consider the concept of consistency introduced by G.N. Milstein [10] that was used later, for example, by C.T.H. Baker and E. Buckwar [2] for stochastic delay equations. Then we compare with the approach from the book of P.E. Kloeden and E. Platen [6, Ch.9.6].

The use of a stochastic version of Spijker's norm (3.6) in our definition of the local truncation error is the key to the two-sided error estimate. But apart from this technical issue there is also a more fundamental difference between our notion of local truncation error and the concepts used in the literature.

By Definition 3.1 a one-step method is consistent if the restriction of the analytic solution X to the grid points produces a small residual for the discrete operator equation $A_h X_h = r_h^F Y$. Then a stability estimate allows to measure the distance between $r_h^E X$ and X_h in terms of the residual $\mathrm{res}_h(r_h^E X) = A_h r_h^E X - r_h^F Y$.

For a comparison with G.N. Milstein's notion of consistency we have to extend the notation. By $X(\cdot;\tau_0,x_0)$, where $\tau_0 \in [0,T]$ and $x_0 \in \mathbb{R}^d$, we denote the solution of (1.1) with initial condition $X(\tau_0;\tau_0,x_0) = x_0$. Similarly, $\bar{X}_{h,t_i,x_0}$, $0 \le i \le N$, denotes the solution to the discrete system

$$
\begin{aligned}
\bar{X}_h(t_i) &= x_0, \\
\bar{X}_h(t_{j+1}) &= \bar{X}_h(t_j) + h\left((1-\theta)b^0(t_j,\bar{X}_h(t_j)) + \theta b^0(t_{j+1},\bar{X}_h(t_{j+1}))\right) \\
&\quad + \sum_{k=1}^{m} b^k(t_j,\bar{X}_h(t_j))\Delta_h W^k(t_{j+1}),\ i \le j \le N-1.
\end{aligned} \tag{3.15}
$$

Then the convergence theorem in [10, Ch.1] assumes the following conditions

$$
\begin{aligned}
\left|\mathbb{E}\left(X(t_{i+1};t_i,x_0) - \bar{X}_{h,t_i,x_0}(t_{i+1})\right)\right| &\le K(1+|x_0|^2)^{\frac{1}{2}}h^{p_1} \\
\left[\mathbb{E}\left(\left|X(t_{i+1};t_i,x_0) - \bar{X}_{h,t_i,x_0}(t_{i+1})\right|^2\right)\right]^{\frac{1}{2}} &\le K(1+|x_0|^2)^{\frac{1}{2}}h^{p_2}
\end{aligned} \tag{3.16}
$$

for some constant $K > 0$ and for all $x_0 \in \mathbb{R}^d$, $0 \le i \le N-1$. In addition, the exponents p_1, p_2 have to satisfy the constraints

$$
p_2 \ge \frac{1}{2}, \quad p_1 \ge p_2 + \frac{1}{2}.
$$

Under this assumption G.N. Milstein proves that the one-step method converges with order $p = p_2 - \frac{1}{2}$. The idea behind (3.16) is that the one-step method causes a small error (in the mean and in the mean-square) after exactly one step when compared to the analytic solution at every grid point t_i and for every initial value x_0.

The main difference between our local truncation error (3.11) and assumption (3.16) lies in the meaning of the word 'local'. While G.N. Milstein localizes the convergence error in time but considers the whole phase space, we localize along the trajectory of the analytic solution X to (1.1) but consider the error over the whole time interval at once. We also emphasize that we do not solve a discrete equation of the form (3.15) in order to define consistency. Rather, our notions of consistency and solvability are separated in the case of implicit methods ($\theta > 0$).

However, for explicit one-step methods this difference is only minor. In [7, App.A] it is shown, that if the Euler-Maruyama method ($\theta = 0$) satisfies assumption (3.16) (which is true for $p_2 = 1$, $p_1 = 2$ [10, Ch.1]) then it is also consistent in the sense of definition 3.1 with order $\gamma = \frac{1}{2}$.

In [6, Ch.9.6] the authors P.E. Kloeden and E. Platen use another concept of consistency. They call a one-step method "strongly consistent" if there exists a nonnegative function $c = c(h)$ with

$$\lim_{h \searrow 0} c(h) = 0$$

such that

$$\mathbb{E}\left(\left|\mathbb{E}\left(\frac{\bar{X}_{h,t_i,x_0}(t_{i+1}) - \bar{X}_{h,t_i,x_0}(t_i)}{h}\middle| \mathcal{F}_{t_i}\right) - b^0\left(t_i, \bar{X}_{h,t_i,x_0}(t_i)\right)\right|^2\right) \leq c(h)$$

and

$$\mathbb{E}\left(\frac{1}{h}\middle|\bar{X}_{h,t_i,x_0}(t_{i+1}) - \bar{X}_{h,t_i,x_0}(t_i) - \mathbb{E}\left(\bar{X}_{h,t_i,x_0}(t_{i+1}) - \bar{X}_{h,t_i,x_0}(t_i)\middle|\mathcal{F}_{t_i}\right) - \sum_{k=1}^{m} b^k\left(t_i, \bar{X}_{h,t_i,x_0}(t_i)\right)\Delta_h W^k(t_{i+1})\right|^2\right) \leq c(h)$$

for all initial values $x_0 \in \mathbb{R}^d$ and $0 \leq i \leq N-1$. Again, this notion focuses on the one-step method (3.15) started at arbitrary vectors rather than at the specific solution X as in definition 3.1.

In a sense, strong consistency of a method requires that it does not differ too much from one step of the Euler-Maruyama scheme which is taken as a prototype of a strongly consistent scheme. Consequently, the Euler-Maruyama method itself is strongly consistent with $c \equiv 0$. This interpretation also shows that the function c is of limited use when estimating the order of convergence. Contrary to this, our notion of consistency provides a strong link to the order of convergence. As Theorem 3.5 shows, the local truncation error (3.11) lies in the class $\mathcal{O}(h^\gamma)$ if and only if the strong error of convergence does.

4. Consistency

The aim of this section is to prove Theorem 3.3. To this end we estimate the local truncation error (3.11)

$$\begin{aligned}
&\|A_h r_h^E X - r_h^F A X\|_{-1} \\
&= \max_{0\leq i\leq N}\left\|\sum_{j=1}^{i}\left[X(t_j) - X(t_{j-1}) - h_j\left((1-\theta)b^0(t_{j-1}, X(t_{j-1})) + \theta b^0(t_j, X(t_j))\right)\right.\right. \\
&\qquad\left.\left. - \sum_{k=1}^{m} b^k(t_{j-1}, X(t_{j-1}))\Delta_h W^k(t_j)\right]\right\|_{L^2(\Omega)} e^{-Lt_i}.
\end{aligned}$$

Note that the terms for $i = 0$ vanish and that the exponential weight is bounded by 1. Then from the representation (2.1) and the triangle inequality we obtain the

estimate

$$\leq (1-\theta) \max_{0\leq i\leq N} \left\| \sum_{j=1}^{i} \left[\int_{t_{j-1}}^{t_j} \left[b^0(s,X(s)) - b^0(t_{j-1},X(t_{j-1})) \right] ds \right] \right\|_{L^2(\Omega)} \tag{4.1}$$

$$+\theta \max_{0\leq i\leq N} \left\| \sum_{j=1}^{i} \left[\int_{t_{j-1}}^{t_j} \left[b^0(s,X(s)) - b^0(t_{j},X(t_{j})) \right] ds \right] \right\|_{L^2(\Omega)} \tag{4.2}$$

$$+ \max_{0\leq i\leq N} \sum_{k=1}^{m} \left\| \sum_{j=1}^{i} \left[\int_{t_{j-1}}^{t_j} \left[b^k(s,X(s)) - b^k(t_{j-1},X(t_{j-1})) \right] dW^k(s) \right] \right\|_{L^2(\Omega)} . \tag{4.3}$$

We estimate the terms separately. The square of the first term (4.1) has the form

$$S_1(i) := \mathbb{E}\left(\left| \sum_{j=1}^{i} \int_{t_{j-1}}^{t_j} \left[b^0(s,X(s)) - b^0(t_{j-1},X(t_{j-1})) \right] ds \right|^2 \right), \quad i = 1,\dots,N.$$

Now Jensen's inequality yields

$$\begin{aligned} S_1(i) &= \mathbb{E}\left(t_i^2 \left| \sum_{j=1}^{i} \frac{h_j}{t_i} \int_{t_{j-1}}^{t_j} \frac{1}{h_j} \left[b^0(s,X(s)) - b^0(t_{j-1},X(t_{j-1})) \right] ds \right|^2 \right) \\ &\leq \mathbb{E}\left(t_i \sum_{j=1}^{i} \int_{t_{j-1}}^{t_j} \left| b^0(s,X(s)) - b^0(t_{j-1},X(t_{j-1})) \right|^2 ds \right) \\ &= t_i \sum_{j=1}^{i} \int_{t_{j-1}}^{t_j} \mathbb{E}\left(\left| b^0(s,X(s)) - b^0(t_{j-1},X(t_{j-1})) \right|^2 \right) ds. \end{aligned}$$

Using the assumptions (A2) and (A3) we find for $k = 0,\dots,m$

$$\begin{aligned} &\left| b^k(s,X(s)) - b^k(t_{j-1},X(t_{j-1})) \right|^2 \\ &\leq \left(\left| b^k(s,X(s)) - b^k(s,X(t_{j-1})) \right| + \left| b^k(s,X(t_{j-1})) - b^k(t_{j-1},X(t_{j-1})) \right| \right)^2 \\ &\leq \left(K \left| X(s) - X(t_{j-1}) \right| + K \left(1 + \left| X(t_{j-1}) \right| \right) \sqrt{|s-t_{j-1}|} \right)^2 \\ &\leq 2K^2 \left| X(s) - X(t_{j-1}) \right|^2 + 2K^2 \left(1 + \left| X(t_{j-1}) \right| \right)^2 |s-t_{j-1}| . \end{aligned}$$

Applying Theorem 2.1 leads to the estimate

$$\begin{aligned} &\mathbb{E}\left(\left| b^k(s,X(s)) - b^k(t_{j-1},X(t_{j-1})) \right|^2 \right) \\ &\leq 2K^2 \mathbb{E}\left(\left| X(s) - X(t_{j-1}) \right|^2 \right) + 4K^2 \left(1 + \mathbb{E}\left(\left| X(t_{j-1}) \right|^2 \right) \right) |s-t_{j-1}| \\ &\leq C \left| s - t_{j-1} \right| , \end{aligned} \tag{4.4}$$

where the constant C only depends on K, T and $\mathbb{E}(|X_0|^2)$. Hence we complete our estimate of $S_1(i)$ as follows

$$\begin{aligned} S_1(i) &\leq t_i \sum_{j=1}^{i} \int_{t_{j-1}}^{t_j} C|s-t_{j-1}|ds \\ &= Ct_i \sum_{j=1}^{i} \frac{1}{2} h_j^2 \leq \frac{1}{2} CT^2 |h|. \end{aligned} \tag{4.5}$$

Replacing t_{j-1} by t_j in (4.4) one gets the analogous result for the term in (4.2), i.e.

$$S_2(i) := \mathbb{E}\left(\left|\sum_{j=1}^{i}\int_{t_{j-1}}^{t_j}\left[b^0(s,X(s)) - b^0(t_j,X(t_j))\right]ds\right|^2\right) \leq \frac{1}{2}CT^2|h|. \tag{4.6}$$

Thus it remains to estimate the third term from (4.3)

$$S_3(i,k) := \left\|\sum_{j=1}^{i}\int_{t_{j-1}}^{t_j}\left[b^k(s,X(s)) - b^k(t_{j-1},X(t_{j-1}))\right]dW^k(s)\right\|^2_{L^2(\Omega)}, k = 1,\dots,m.$$

By the martingale property of the stochastic Itô integral (c.f. Corollary (5.2.1) in [1]) we find for all indices $j,\ell = 1,\dots,i$ with $j \neq \ell$

$$\left\langle \int_{t_{j-1}}^{t_j}\left[b^k(s,X(s)) - b^k(t_{j-1},X(t_{j-1}))\right]dW^k(s), \right.$$
$$\left. \int_{t_{\ell-1}}^{t_\ell}\left[b^k(s,X(s)) - b^k(t_{\ell-1},X(t_{\ell-1}))\right]dW^k(s)\right\rangle_{L^2(\Omega)} = 0.$$

Hence, by the Pythagoras theorem we get

$$\begin{aligned} S_3(i,k) &= \sum_{j=1}^{i}\left\|\int_{t_{j-1}}^{t_j}\left[b^k(s,X(s)) - b^k(t_{j-1},X(t_{j-1}))\right]dW^k(s)\right\|^2_{L^2(\Omega)} \\ &= \sum_{j=1}^{i}\mathbb{E}\left(\left|\int_{t_{j-1}}^{t_j}\left[b^k(s,X(s)) - b^k(t_{j-1},X(t_{j-1}))\right]dW^k(s)\right|^2\right) \\ &= \sum_{j=1}^{i}\int_{t_{j-1}}^{t_j}\mathbb{E}\left(\left|b^k(s,X(s)) - b^k(t_{j-1},X(t_{j-1}))\right|^2\right)ds, \end{aligned}$$

where we used the Itô isometry in the last step. Again we apply (4.4) and obtain

$$\begin{aligned} S_3(i,k) &\leq \sum_{j=1}^{i}\int_{t_{j-1}}^{t_j} C|s - t_{j-1}|ds \\ &\leq \frac{1}{2}CT|h|. \end{aligned} \tag{4.7}$$

Combining the estimates (4.5), (4.6) and (4.7) we arrive at the final estimate

$$\begin{aligned} \left\|A_h r_h^E X - r_h^F AX\right\|_{-1} &\leq (1-\theta)\sqrt{\frac{1}{2}CT^2|h|} + \theta\sqrt{\frac{1}{2}CT^2|h|} + \sum_{k=1}^{m}\sqrt{\frac{1}{2}CT|h|} \\ &= \bar{C}|h|^{\frac{1}{2}}, \end{aligned}$$

where the constant $\bar{C}$ only depends on K, T, $\mathbb{E}(|X_0|^2)$ and m. Thus the proof of theorem 3.3 is complete.

Remark. For a given stochastic process f the norm of the deterministic integral

$$\left\|\int_{t_{j-1}}^{t_j} f(s)ds\right\|_{L^2(\Omega)}$$

converges to 0 as $h_j = t_j - t_{j-1} \to 0$ faster than the norm of the stochastic integral

$$\left\| \int_{t_{j-1}}^{t_j} f(s) dW^k(s) \right\|_{L^2(\Omega)} .$$

However, as we have shown, for the stochastic Spijker-norm (3.6) both types of integrals give the same order of convergence. This is the key property of the stochastic Spijker-norm that facilitates the proof of the two-sided error estimate with maximum rate of convergence.

5. Stability

This section is devoted to the proof of bistability for the STM. The main idea is to rewrite the STM as a fixed point problem that is a discrete analog of the integral equation (2.1). Then we choose appropriate parameter values $L, \rho, \sigma > 0$ such that the Banach fixed point theorem applies.

We define the mapping

$$\Phi_h : \begin{array}{l} E_h \times F_h \to \mathcal{G}_h \\ (Y_h, Z_h) \mapsto \Phi_h(Y_h, Z_h) \end{array} \tag{5.1}$$

by

$$\begin{aligned} [\Phi_h(Y_h, Z_h)](t_i) = &\sum_{j=0}^{i} Z_h(t_j) + \sum_{j=1}^{i} h_j \left[(1-\theta) b^0(t_{j-1}, Y_h(t_{j-1})) + \theta b^0(t_j, Y_h(t_j)) \right] \\ &+ \sum_{j=1}^{i} \sum_{k=1}^{m} \left[b^k(t_{j-1}, Y_h(t_{j-1})) \Delta_h W^k(t_j) \right], \quad i = 0, \dots, N. \end{aligned}$$

By induction on i one readily proves the following equivalence for all $Y_h \in D(A_h)$ and $Z_h \in F_h$

$$\Phi_h(Y_h, Z_h) = Y_h \qquad \Longleftrightarrow \qquad A_h Y_h = Z_h. \tag{5.2}$$

Since the inhomogeneities $Z_h(t_j)$ appear as sums in Φ_h we have the following norm relation for $Y_h \in E_h$ and $Z_h, \tilde{Z}_h \in F_h$

$$\left\| \Phi_h(Y_h, Z_h) - \Phi_h(Y_h, \tilde{Z}_h) \right\|_0 = \left\| Z_h - \tilde{Z}_h \right\|_{-1} . \tag{5.3}$$

The following lemma is a generalization of a similar assertion proved in [18, §6, (19)] for the case of deterministic ordinary differential equations.

Lemma 5.1. *For every $\rho > 0$ and $L \geq 2K^2(e\sqrt{T} + m)^2$ there exists an $h_{\rho,L} > 0$ such that the following properties hold with the settings*

$$\sigma_\rho = \frac{\rho}{4}, \quad E_h := \overline{B}_{\rho,L}(r_h^E X), \quad F_h := \overline{B}_{\sigma_\rho,L}(r_h^F Y) \tag{5.4}$$

for all grids with $|h| \leq h_{\rho,L}$

(i) $\left\| \mathrm{res}_h(r_h^E X) \right\|_{-1} \leq \sigma_\rho$,
(ii) $\Phi_h(E_h \times F_h) \subset E_h$,
(iii) *The inequality*

$$\left\| \Phi_h(Y_h, Z_h) - \Phi_h(\tilde{Y}_h, Z_h) \right\|_0 \leq \frac{1}{2} \left\| Y_h - \tilde{Y}_h \right\|_0 \tag{5.5}$$

is satisfied for all $Y_h, \tilde{Y}_h \in E_h, Z_h \in F_h$.

Proof. Let us first prove (5.5). For $Y_h, \tilde{Y}_h \in E_h$ and $Z_h \in F_h$ consider the term

$$\left\|\Phi_h(Y_h, Z_h) - \Phi_h(\tilde{Y}_h, Z_h)\right\|_0 = \max_{0\le i\le N} \left\|[\Phi_h(Y_h, Z_h)](t_i) - [\Phi_h(\tilde{Y}_h, Z_h)](t_i)\right\|_{L^2(\Omega)} e^{-Lt_i}. \tag{5.6}$$

For $0 \le i \le N$ we estimate as follows

$$\left\|[\Phi_h(Y_h, Z_h)](t_i) - [\Phi_h(\tilde{Y}_h, Z_h)](t_i)\right\|_{L^2(\Omega)}$$

$$\le (1-\theta) \left\|\textstyle\sum_{j=1}^i h_j \left[b^0(t_{j-1}, Y_h(t_{j-1})) - b^0(t_{j-1}, \tilde{Y}_h(t_{j-1}))\right]\right\|_{L^2(\Omega)} \tag{5.7}$$

$$+ \theta \left\|\textstyle\sum_{j=1}^i h_j \left[b^0(t_j, Y_h(t_j)) - b^0(t_j, \tilde{Y}_h(t_j))\right]\right\|_{L^2(\Omega)} \tag{5.8}$$

$$+ \sum_{k=1}^m \left\|\textstyle\sum_{j=1}^i \left[b^k(t_{j-1}, Y_h(t_{j-1})) - b^k(t_{j-1}, \tilde{Y}_h(t_{j-1}))\right] \Delta_h W^k(t_j)\right\|_{L^2(\Omega)}. \tag{5.9}$$

Note that this estimate does not depend on Z_h. In the following we treat the terms separately. We apply Jensen's inequality to the square of the term (5.7) and then use assumption (A2) to obtain

$$\begin{aligned}
&\mathbb{E}\left(\left|\textstyle\sum_{j=1}^i h_j \left[b^0(t_{j-1}, Y_h(t_{j-1})) - b^0(t_{j-1}, \tilde{Y}_h(t_{j-1}))\right]\right|^2\right) \\
&\le t_i \sum_{j=1}^i h_j \mathbb{E}\left(\left|b^0(t_{j-1}, Y_h(t_{j-1})) - b^0(t_{j-1}, \tilde{Y}_h(t_{j-1}))\right|^2\right) \\
&\le K^2 t_i \sum_{j=1}^i h_j \mathbb{E}\left(\left|Y_h(t_{j-1}) - \tilde{Y}_h(t_{j-1})\right|^2\right).
\end{aligned}$$

Since

$$\mathbb{E}\left(\left|Y_h(t_{j-1}) - \tilde{Y}_h(t_{j-1})\right|^2\right) \le \left\|Y_h - \tilde{Y}_h\right\|_0^2 e^{2Lt_{j-1}} \tag{5.10}$$

we complete the estimate of (5.7) by

$$\begin{aligned}
&\le K^2 t_i \left\|Y_h - \tilde{Y}_h\right\|_0^2 \sum_{j=1}^i h_j e^{2Lt_{j-1}} \\
&\le K^2 t_i \left\|Y_h - \tilde{Y}_h\right\|_0^2 \frac{1}{2L}\left(e^{2Lt_i} - 1\right).
\end{aligned} \tag{5.11}$$

If we apply (5.10) with t_j instead of t_{j-1} the same arguments work for the term (5.8) and we arrive at

$$\begin{aligned}
&\left\|\textstyle\sum_{j=1}^i h_j \left[b^0(t_j, Y_h(t_j)) - b^0(t_j, \tilde{Y}_h(t_j))\right]\right\|_{L^2(\Omega)}^2 \\
&\le K^2 t_i \left\|Y_h - \tilde{Y}_h\right\|_0^2 \sum_{j=1}^i h_j e^{2Lt_j} \\
&\le K^2 t_i \left\|Y_h - \tilde{Y}_h\right\|_0^2 e^{2L|h|} \frac{1}{2L}\left(e^{2Lt_i} - 1\right).
\end{aligned} \tag{5.12}$$

It remains to estimate (5.9). As in section 4 we use the martingale property of the Itô-integral to obtain

$$\begin{aligned}
&\left\|\sum_{j=1}^{i}\left[b^k(t_{j-1},Y_h(t_{j-1}))-b^k(t_{j-1},\tilde{Y}_h(t_{j-1}))\right]\Delta_h W^k(t_j)\right\|^2_{L^2(\Omega)}\\
&=\sum_{j=1}^{i}\mathbb{E}\left(\left|\left[b^k(t_{j-1},Y_h(t_{j-1}))-b^k(t_{j-1},\tilde{Y}_h(t_{j-1}))\right]\Delta_h W^k(t_j)\right|^2\right)\\
&\le\sum_{j=1}^{i}h_j\mathbb{E}\left(\left|b^k(t_{j-1},Y_h(t_{j-1}))-b^k(t_{j-1},\tilde{Y}_h(t_{j-1}))\right|^2\right)\\
&\le K^2\sum_{j=1}^{i}h_j\mathbb{E}\left(\left|Y_h(t_{j-1})-\tilde{Y}_h(t_{j-1})\right|^2\right)\\
&\le K^2\left\|Y_h-\tilde{Y}_h\right\|_0^2\sum_{j=1}^{i}h_j e^{2Lt_{j-1}}\\
&\le K^2\left\|Y_h-\tilde{Y}_h\right\|_0^2\frac{1}{2L}\left(e^{2Lt_i}-1\right),\quad k=1,\dots,m. \qquad (5.13)
\end{aligned}$$

By inserting (5.11), (5.12) and (5.13) into (5.6) we get the estimate

$$\begin{aligned}
&\left\|\Phi_h(Y_h,Z_h)-\Phi_h(\tilde{Y}_h,Z_h)\right\|_0\\
&\le K\left\|Y_h-\tilde{Y}_h\right\|_0\max_{0\le i\le N}\left[\left(\frac{1}{2L}(e^{2Lt_i}-1)\right)^{\frac12}\left((1-\theta)\sqrt{t_i}+\theta\sqrt{t_i}e^{L|h|}+m\right)\right]e^{-Lt_i}\\
&\le K\left\|Y_h-\tilde{Y}_h\right\|_0\max_{0\le i\le N}\left(\frac{1}{2L}\left(1-e^{-2Lt_i}\right)\right)^{\frac12}\left(\sqrt{t_i}e^{L|h|}+m\right)\\
&\le\frac{K}{\sqrt{2L}}\left\|Y_h-\tilde{Y}_h\right\|_0\left(\sqrt{T}e^{L|h|}+m\right).
\end{aligned}$$

Taking $h_{\rho,L}\le\frac{1}{L}$ the contraction estimate (5.5) follows by the choice of L.

By Theorem 3.3 the STM is consistent. After possibly reducing $h_{\rho,L}>0$ further we obtain

$$\left\|r_h^E X-\Phi_h(r_h^E X,r_h^F Y)\right\|_0=\left\|\mathrm{res}_h(r_h^E X)\right\|_{-1}\le\sigma_\rho=\frac14\rho \qquad (5.14)$$

for all $|h|\le h_{\rho,L}$.

It remains to show that Φ_h maps $E_h\times F_h$ into E_h. But this follows for all $|h|\le h_{\rho,L}$ from (5.3),(5.5) and (5.14)

$$\begin{aligned}
\left\|r_h^E X-\Phi_h(Y_h,Z_h)\right\|_0&\le\left\|r_h^E X-\Phi_h(r_h^E X,r_h^F Y)\right\|_0\\
&\quad+\left\|\Phi_h(r_h^E X,r_h^F Y)-\Phi_h(Y_h,r_h^F Y)\right\|_0+\left\|\Phi_h(Y_h,r_h^F Y)-\Phi_h(Y_h,Z_h)\right\|_0\\
&\le\sigma_\rho+\frac12\rho+\sigma_\rho=\rho.
\end{aligned}$$

Thus $\Phi_h(Y_h,Z_h)\in E_h=\overline{B}_{\rho,L}(r_h^E X)$ for all $Y_h\in E_h$ and $Z_h\in F_h$. □

Remark. One can improve the lower bound on L to $L>\frac12K^2\left(\sqrt{T}+m\right)^2$ at the expense of larger contraction and stability constants that depend on L and T.

We are now prepared for the proof of Theorem 3.4.

Proof of theorem 3.4. Choose parameter values L, ρ, σ_ρ according to Lemma 5.1. Then for every $Z_h \in F_h$ the mapping $\Phi_h(\cdot, Z_h) : E_h \to E_h$ is a contraction and the Banach fixed point theorem yields the existence of a unique fixed point Y_h in E_h, i.e.

$$\Phi_h(Y_h, Z_h) = Y_h.$$

Therefore, by the relation (5.2), there also exists a unique solution $Y_h \in D(A_h)$ (recall (3.9)) of the equation

$$A_h Y_h = Z_h.$$

Hence the mapping $A_h : D(A_h) \to F_h$ is bijective for $|h| \leq h_{\rho,L}$. Moreover, we have the relationship

$$\Phi_h(Y_h, A_h Y_h) = Y_h \tag{5.15}$$

for all $Y_h \in D(A_h)$.

Now consider arbitrary elements $Y_h, \tilde{Y}_h \in D(A_h)$. With the help of (5.15) and Lemma 5.1 we obtain

$$\begin{aligned}
&\left\| Y_h - \tilde{Y}_h \right\|_0 \\
&= \left\| \Phi_h(Y_h, A_h Y_h) - \Phi_h(\tilde{Y}_h, A_h \tilde{Y}_h) \right\|_0 \\
&\leq \left\| \Phi_h(Y_h, A_h Y_h) - \Phi_h(\tilde{Y}_h, A_h Y_h) \right\|_0 + \left\| \Phi_h(\tilde{Y}_h, A_h Y_h) - \Phi_h(\tilde{Y}_h, A_h \tilde{Y}_h) \right\|_0 \\
&\leq \frac{1}{2} \left\| Y_h - \tilde{Y}_h \right\|_0 + \left\| A_h Y_h - A_h \tilde{Y}_h \right\|_{-1}.
\end{aligned}$$

Hence

$$\left\| Y_h - \tilde{Y}_h \right\|_0 \leq 2 \left\| A_h Y_h - A_h \tilde{Y}_h \right\|_{-1},$$

which is one part of the bistability inequality (3.12). The second part follows from

$$\begin{aligned}
&\left\| A_h Y_h - A_h \tilde{Y}_h \right\|_{-1} \\
&= \left\| \Phi_h(Y_h, A_h Y_h) - \Phi_h(Y_h, A_h \tilde{Y}_h) \right\|_0 \\
&\leq \left\| Y_h - \Phi_h(\tilde{Y}_h, A_h \tilde{Y}_h) \right\|_0 + \left\| \Phi_h(\tilde{Y}_h, A_h \tilde{Y}_h) - \Phi_h(Y_h, A_h \tilde{Y}_h) \right\|_0 \\
&\leq \left\| Y_h - \tilde{Y}_h \right\|_0 + \frac{1}{2} \left\| \tilde{Y}_h - Y_h \right\|_0 \\
&= \frac{3}{2} \left\| Y_h - \tilde{Y}_h \right\|_0.
\end{aligned}$$

□

Remark. Note that the parameter $\rho > 0$ in Lemma 5.1 is arbitrary due to the global Lipschitz condition (A2). For the proof of Lemma 5.1 we only need a Lipschitz condition of the form

$$\mathbb{E}\left(\left| b^k(s, Y_h(s)) - b^k(s, \tilde{Y}_h(s)) \right|^2 \right) \leq K_\rho \mathbb{E}\left(\left| Y_h(s) - \tilde{Y}_h(s) \right|^2 \right). \tag{5.16}$$

Thus the STM is still bistable if we weaken assumption (A2) by assuming constants $\rho, K_\rho > 0$ such that (5.16) holds for all $s \in [0, T]$ and all adapted random variables

$Y_h(s), \tilde{Y}_h(s)$ in the ρ-neighborhood of $X(s)$, i.e. for all $Y_h(s) \in L^2(\Omega, \mathcal{F}_s, P; \mathbb{R}^d)$ with

$$\|Y_h(s) - X(s)\|_{L^2(\Omega)} \leq \rho.$$

In this case the lower bound for the exponential weight parameter L also depends on ρ.

6. Maximum order of convergence

In this section we discuss the maximum order of convergence for the STM. J.M.C. Clark and R.J. Cameron [3] constructed an example to show that, in general, the maximum order of convergence is equal to $\frac{1}{2}$. This extends to all one-step methods which use only the increments $W^k(t_i) - W^k(t_{i-1})$ of the driving Wiener processes. We will show that the same result follows in a natural way for the STM from the two-sided error estimate (3.13).

Unlike e.g. the Milstein method [6, 10], the STM does not use information about the Wiener processes at intermediate times $s \in [0,T] \setminus \tau_h$. Therefore, one cannot expect the STM to give good approximations for an equation with an iterated Wiener process, i.e. with a stochastic integral of the form

$$\int_0^t W^1(s) dW^2(s).$$

As in [3], we consider two real and independent $(\mathcal{F}_t)_{t\geq 0}$-Wiener processes W^1 and W^2. The two-dimensional stochastic differential equation

$$\begin{aligned} dX(t) &= \begin{pmatrix} 1 & 0 \\ 0 & X_1(t) \end{pmatrix} d \begin{pmatrix} W^1(t) \\ W^2(t) \end{pmatrix} \\ X(0) &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \end{aligned} \tag{6.1}$$

has the analytic solution

$$X(t) = \begin{pmatrix} W^1(t) \\ \int_0^t W^1(s) dW^2(s) \end{pmatrix}, \quad \text{for} \quad t \in [0,T]. \tag{6.2}$$

For this equation the local truncation error of the STM is

$$\begin{aligned} &\|A_h r_h^E X - r_h^F A X\|_{-1}^2 \\ &= \max_{0\leq i\leq N} \mathbb{E}\left(\left| \sum_{j=1}^{i} \left[X(t_j) - X(t_{j-1}) - \begin{pmatrix} 1 & 0 \\ 0 & X_1(t_{j-1}) \end{pmatrix} \Delta_h W(t_j) \right] \right|^2 \right) \\ &= \max_{0\leq i\leq N} \mathbb{E}\left(\left| \sum_{j=1}^{i} \left[\begin{pmatrix} 0 \\ \int_{t_{j-1}}^{t_j} W^1(s) dW^2(s) - W^1(t_{j-1}) \left(W^2(t_j) - W^2(t_{j-1}) \right) \end{pmatrix} \right] \right|^2 \right). \end{aligned}$$

As before $\Delta_h W(t_j)$ denotes the j-th increment of the two Wiener processes. Note that the local truncation error is independent of the parameter $\theta \in [0,1]$. Since the

first component is equal to zero the local truncation error equals

$$
\begin{aligned}
&= \max_{0\le i\le N} \mathbb{E}\left(\left|\sum_{j=1}^{i}\int_{t_{j-1}}^{t_j}\left(W^1(s)-W^1(t_{j-1})\right)dW^2(s)\right|^2\right)\\
&= \max_{0\le i\le N}\left[\sum_{j=1}^{i}\mathbb{E}\left(\left|\int_{t_{j-1}}^{t_j}\left(W^1(s)-W^1(t_{j-1})\right)dW^2(s)\right|^2\right)\right]\\
&= \sum_{j=1}^{N}\int_{t_{j-1}}^{t_j}\mathbb{E}\left(\left|W^1(s)-W^1(t_{j-1})\right|^2\right)ds\\
&= \sum_{j=1}^{N}\int_{t_{j-1}}^{t_j}|s-t_{j-1}|\,ds\\
&= \frac{1}{2}\sum_{j=1}^{N}(t_j-t_{j-1})^2,
\end{aligned}
$$

where we used the martingale property of the stochastic integral in the first step and the Itô-isometry in the second step. Note that we have an exact expression for the local truncation error in the Spijker norm.

The Cauchy-Schwarz inequality yields

$$\frac{1}{2}\sum_{j=1}^{N}(t_j-t_{j-1})^2 \ge \frac{1}{2}\frac{T^2}{N}. \tag{6.3}$$

Thus the local truncation error is bounded from below by a term of order $\mathcal{O}(\left(\frac{T}{N}\right)^{\frac{1}{2}})$.

In case of an equidistant step size $h=\frac{T}{N}$ we have equality in (6.3). The two-sided error estimate (3.13) then reads

$$C_1\sqrt{\frac{1}{2}Th} \le \|X-X_h\|_0 \le C_2\sqrt{\frac{1}{2}Th},$$

which shows that the STM converges with the exact order $\gamma=\frac{1}{2}$.

7. Conclusions

The bistability of a numerical discretization method is formulated in terms of two norms that allow to estimate differences of grid functions by differences of residuals and vice versa. This property plays an important role in deriving two-sided estimates for the convergence error. In this paper we have set up a pair of norms that guarantee bistability of the stochastic theta method for an SODE satisfying Lipschitz and Hölder conditions. One of the norms is the maximum of the strong norm at grid points while the second norm is a stochastic generalization of Spijker's norm. We have also shown that upper and lower estimates of type $\sqrt{h}$ follow in a natural way from the corresponding two-sided error estimates. As a by-product of our approach we succeeded in embedding the theory of convergence for the stochastic theta method into the standard framework of consistency, stability and convergence as developed e.g. by Stummel.

Several natural questions follow from our approach. First, it is natural to ask whether the two-sided error estimates extend to higher methods (e.g. those in [10]). A positive answer will be given in the forthcoming paper [8]. Second, one may ask for other pairs of norms, either both stronger or both weaker than the ones

considered here, that allow to prove bistability. Finally, it seems natural to ask for bistability of discretization methods applied to infinite dimensional stochastic equations, such as delay equations and partial differential equations.

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] 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.

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

[5] D.J. Higham. Mean-square and asymptotic stability of the stochastic theta method. *SIAM J. Numer. Anal.*, 38(3):753–769, 2000.

[6] P.E. Kloeden and E. Platen. *Numerical solution of stochastic differential equations*. Springer, Berlin, third edition, 1999.

[7] R. Kruse. Diskrete Approximationen gewöhnlicher und partieller stochastischer Differentialgleichungen. Diploma Thesis, Bielefeld University, 2008.

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

[9] X. Mao. *Stochastic differential equations and applications*. Horwood, UK, 1997.

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

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

[12] M.N. Spijker. *Stability and convergence of finite-difference methods*. PhD thesis, University Leiden, 1968.

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

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

[15] F. Stummel. Diskrete Konvergenz linearer Operatoren I. *Math. Ann.*, 190:45–92, 1971.

[16] F. Stummel. Diskrete Konvergenz linearer Operatoren II. *Math. Z.*, 120:231–264, 1971.

[17] F. Stummel. Diskrete Konvergenz linearer Operatoren III. In *Linear operators and approximation (Proc. Conf., Oberwolfach, 1971)*, volume 20 of *Internat. Ser. Numer. Math.*, pages 196–216, Basel, 1972. Birkhäuser.

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

E-mail address: beyn@math.uni-bielefeld.de

E-mail address: rkruse@math.uni-bielefeld.de