A LAX EQUIVALENCE THEOREM FOR STOCHASTIC DIFFERENTIAL EQUATIONS

ANNIKA LANG

Abstract. In this paper, a stochastic mean square version of Lax's equivalence theorem for Hilbert space valued stochastic differential equations with additive and multiplicative noise is proved. Definitions for consistency, stability, and convergence in mean square of an approximation of a stochastic differential equation are given and it is shown that these notions imply similar results as those known for approximations of deterministic partial differential equations. Examples show that the made assumptions are met by standard approximations.

1. Introduction

A classical result in the theory of numerical methods for partial differential equations (PDEs) is Lax's equivalence theorem [21] which states that a consistent approximation of a linear PDE is convergent if and only if it is stable. Within the last years the extension of PDEs to stochastic partial differential equations (SPDEs) has become more and more important in applications especially in engineering such as image analysis, surface analysis, filtering [15, 18, 22, 23, 31]. On the other hand side, in finance, people extend finite dimensional systems of stochastic differential equations (SDEs) to infinite dimensional ones [5], i.e. to SPDEs. Explicit solutions to most of the problems do not exist. Therefore it is natural to simulate these SPDEs. In this paper we look at SPDEs of Itô type as Hilbert space valued SDEs and approximate their mild solutions. This approach has been done in recent works, see e.g. [1, 14, 18] and references therein. The main result of this paper is that we extend Lax's equivalence theorem for approximations of PDEs, which can be found in slightly different versions as for finite differences and in a Hilbert space framework in [7, 9, 10, 17, 26], in a mean square sense to these SDEs and their approximations. In order to make things compatible with our chosen Hilbert space framework, we apply Theorem XX.3.1 in [9] as classical Lax equivalence theorem. First approaches for a stochastic version of this theorem can be found in [27, 28, 29]. Roth shows in [27, 28] for finite difference approximations of SPDEs driven by a one-dimensional Brownian motion that his definitions of consistency and stability imply weak convergence of a subsequence of approximation schemes. In [29], systems of real valued SDEs are approximated and mean square convergence is shown under consistency, stability, and some further assumptions.

For the used definitions in this paper, it is important to mention that stability in the sense of Lax and Richtmyer just depends on the approximation of the deterministic part of the equation. As big difference to PDE theory the approximation scheme of an SPDE needs a

Date: April 30, 2010.

2000 *Mathematics Subject Classification.* 60G60, 60H15, 60H35, 65C30, 65C05.

Key words and phrases. Stochastic partial differential equations, Lax equivalence theorem, numerical approximation, consistency, stability, convergence.

The author wishes to express many thanks to Andrea Barth, Pao-Liu Chow, and Jürgen Potthoff for all the fruitful discussions and hints.

weaker formulation of consistency, i.e. a decay with the square root of the step size suffices instead of a linear one. Furthermore, it has to be ensured that the properties of stochastic Itô integrals are preserved in the corresponding approximations.

For equations with additive and multiplicative noise, i.e. equations of the form

$$dX(t) = AX(t)\,dt + G(X(t))\,dM(t)$$

with initial condition $X(0) = X_0$, where A is assumed to be generator of a C_0–semigroup, G is of Lipschitz type and M is a càdlàg square integrable martingale, we define approximation schemes. Definitions of convergence in mean square, consistency in mean square, and stability are given. Extensions to more general integrators are subject to further work. These definitions introduced in Section 2 are related to each other in Section 3, where the main result of this paper — Lax's equivalence theorem holds for SPDEs under the transformed definitions of convergence and consistency — is stated and proved. Finally, a finite difference scheme for the heat equation with multiplicative noise is introduced in Section 4. This example emphasizes the definitions of Section 2 and shows that the made assumptions are met by standard and even very simple approximations. References to advanced examples are given in that section.

2. Convergence, Consistency, and Stability

Let $(H, (\cdot,\cdot)_H)$ be a separable Hilbert space with corresponding norm $\|\cdot\|_H$, e.g. $H = L^2(D)$, where $D \subset \mathbb{R}^d$ is a bounded or unbounded region in $\mathbb{R}^d$. Furthermore let $V_h \subset H$ be a finite dimensional subspace where, in general, $h > 0$ represents a discretization step in space, such that V_h converges in the following sense to H as $h \to 0$: Assume that there exists an orthogonal projection P_h from H into V_h such that

$$\lim_{h\to 0} \|P_h u - u\|_H = 0$$

for all $u \in H$, where we use the norm induced by H for the subspaces V_h. We introduce $\mathcal{M}^2(U)$ as the space of all càdlàg square integrable martingales with values in a separable Hilbert space $(U, (\cdot,\cdot)_U)$ on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ satisfying the "usual conditions". Similarly to Section 8.6 in [25], we assume that for given $M \in \mathcal{M}^2(U)$ there exists Q in the space of all nuclear symmetric positive-definite operators from U into itself $L_1^+(U)$ such that for all $r < t$

$$\int_r^t Q_s\, d\langle M, M\rangle_s \leq (t-r)Q,$$

where the $L_1^+(U)$–valued process $(Q_t, t \geq 0)$ is the martingale covariance of M and $\langle M, M\rangle$ denotes the predictable variation process of M given by the Doob–Meyer decomposition. Since $Q \in L_1^+(U)$, there exists an orthonormal basis $(e_n, n \in \mathbb{N})$ of U consisting of eigenvectors of Q. This implies the representation $Qe_n = \lambda_n e_n$, where $\lambda_n \geq 0$ is the eigenvalue corresponding to e_n. The square root of Q is defined by

$$Q^{1/2}x = \sum_n (x, e_n)_U\, \lambda_n^{1/2}\, e_n$$

for $x \in U$ and $Q^{-1/2}$ is the pseudo inverse of $Q^{1/2}$.

Let us denote by $(\mathcal{H}, (\cdot,\cdot)_{\mathcal{H}})$ the Hilbert space defined by $\mathcal{H} = Q^{1/2}(U)$ endowed with the inner product $(x,y)_{\mathcal{H}} = (Q^{-1/2}x, Q^{-1/2}y)_U$ for $x, y \in \mathcal{H}$. Typical processes satisfying these conditions are Hilbert space valued Lévy processes as introduced in [25]. In what follows we define an analog to the Itô isometry for processes in $\mathcal{M}^2(U)$ with bounded covariance, where

$L_{HS}(\mathcal{H}, H)$ refers to the space of all Hilbert–Schmidt operators from $\mathcal{H}$ to H and $\|\cdot\|_{L_{HS}(\mathcal{H},H)}$ denotes the corresponding norm.

Proposition 1 ([25])**.** *Let $\mathbb{L}^2_{\mathcal{H},T}(H) := L^2(\Omega \times [0,T], \mathcal{P}_{[0,T]}, \mathbb{P} \otimes d\lambda; L_{HS}(\mathcal{H}, H))$ be the space of integrands, where $\mathcal{P}_{[0,T]}$ denotes the σ–field of predictable sets in $\Omega \times [0,T]$ and $d\lambda$ is the Lebesgue measure, then for every $X \in \mathbb{L}^2_{\mathcal{H},T}(H)$*

$$\mathbb{E}\big(\| \int_0^t X(s)\, dM(s)\|_H^2\big) \leq \mathbb{E}\big(\int_0^t \|X(s)\|^2_{L_{HS}(\mathcal{H},H)}\, ds\big).$$

We consider the following SPDE on the finite interval $[0,T]$, which is actually a Hilbert space valued SDE,

$$dX(t) = AX(t)\, dt + G(X(t))\, dM(t), \qquad X(0) = X_0, \tag{1}$$

with values in H, where A generates a C_0–semigroup S, $M \in \mathcal{M}^2(U)$ with bounded covariance process $(Q_t, t \geq 0)$, $Q_t \in L_1^+(U)$ for $t \geq 0$, and G is a mapping from H into the linear operators $L(\mathcal{H}, H)$. Furthermore G satisfies that there exists a constant $C \in \mathbb{R}_+$ such that for all $u, v \in H$

$$\begin{aligned} \|G(u)\|_{L_{HS}(\mathcal{H},H)} &\leq C(1 + \|u\|_H), \\ \|G(u) - G(v)\|_{L_{HS}(\mathcal{H},H)} &\leq C\|u - v\|_H. \end{aligned} \tag{2}$$

Then by results in Chapter 9 of [25], Equation (1) has a unique mild solution for an $\mathcal{F}_0$–measurable initial condition X_0, i.e. $\sup_{t\in[0,T]} \mathbb{E}(\|X(t)\|_H^2) < +\infty$ and $X(t)$ can be written as

$$X(t) = S(t)X_0 + \int_0^t S(t-s)G(X(s))\, dM(s).$$

Furthermore these assumptions imply that the corresponding PDE

$$\frac{\partial}{\partial t} u(t) = Au(t) \tag{3}$$

is well-posed, see Chapter 4 in [24].

We introduce a semi-discrete problem on V_h

$$dX_h(t) = A_h X_h(t)\, dt + G_h(X_h(t))\, dM(t), \qquad X_h(0) = X_{0,h} = P_h X_0,$$

where G_h also includes the projection of M into a finite dimensional space. The operator A_h can be obtained for example by finite difference methods (cf. [11],[26],[30]) or finite element methods (cf. [11],[30],[33]). Let $(t_j, j = 0, \ldots, n)$ be a partition of $[0,T]$ with $t_0 = 0$ and $t_n = T$. For the sake of simplicity we assume an equidistant partition of the interval with $\triangle t = T/n$ but the results also hold for arbitrary time discretizations, where the maximal step size converges to zero. In the following $\triangle t$ and n will be coupled by this relation. We define an *approximation method* or *approximation scheme* which allows to calculate $X_h^n \in V_h$, an approximation to $X_h(t_n)$ starting from X_h^{n-p} for $p = 1, \ldots, P$. In this paper we limit ourselves to $P = 1$ which is called a *two-level scheme*. This can be written as

$$X_h^{j+1} = D_h(\triangle t, j)X_h^j = D_h^d(\triangle t)X_h^j + D_h^s(\triangle t, j)X_h^j, \qquad X_h^0 = X_{0,h}, \tag{4}$$

where $D_h^d(\triangle t) \in L(V_h)$ is the linear operator approximating Equation (3) and $D_h^s(\triangle t, j)$ approximates the stochastic integral from t_j to t_{j+1} and does not have to be a linear operator.

A fundamental question for an approximation scheme is that of convergence when h and $\triangle t$ tend to zero. We choose a definition that involves convergence in mean square. The question of almost sure convergence will be addressed in a later paper. The corresponding deterministic definition as well as those for the successive terms can be found in Chapter XX, §1, 2 in [9].

Definition 2. The *error* $e(h, \triangle t) = (e_j(h, \triangle t), j = 0, \ldots, n)$ of an approximation scheme given by Equation (4) is defined by

$$e_j(h, \triangle t) = X(t_j) - X_h^j.$$

A discretization scheme given by Equation (4) is *convergent in mean square* to the solution of Equation (1), if for all $\epsilon > 0$ there exist $\eta, \delta > 0$ such that for all $0 < h < \eta$, $0 < \triangle t < \delta$, and $j \in \{0, \ldots, n\}$ it holds that

$$\mathbb{E}(\|e_j(h, \triangle t)\|_H^2) < \epsilon.$$

Examples of convergent approximation schemes are given in [18] and in Section 4 of this paper.

Two properties that are fundamental for convergence are those of consistency and stability what will be shown in the main result. In order to give stochastic analogs to the known deterministic definitions we need two more definitions. First we define some properties of the stochastic approximation that are not necessary for the approximation of PDEs.

Definition 3. The family of operators approximating the stochastic integral $(D_h^s(\triangle t, j), j \in \{0, \ldots, n-1\})$ in Equation (4) is *$\mathcal{F}$–compatible* with Equation (1) for given h and $\triangle t$, if $D_h^s(\triangle t, j)$ is $\mathcal{F}_{t_{j+1}}$–measurable and $\mathbb{E}(D_h^s(\triangle t, j)|\mathcal{F}_{t_j}) = 0$ for all $j = 0, \ldots, n-1$.

An immediate consequence of $\mathcal{F}$–compatibility is that $\mathbb{E}(D_h^s(\triangle t, j)) = 0$ for all j due to the properties of the conditional expectation. $\mathcal{F}$–compatiblity can already been found in [6]. In Remark 2.5, Buckwar and Winkler suggest an $\mathcal{F}$–compatible representation of extra perturbations of multilevel SDE approximations and use it in the proof of Theorem 3.3 for the approximation of the stochastic integral.

The following simple example shows that $\mathcal{F}$–compatibility is a natural condition that is satisfied for known approximations.

Example 4 (Geometric Brownian Motion)**.** Let $H = \mathbb{R}$ and consider the geometric Brownian motion given by the SDE

$$dX_t = aX_t\,dt + bX_t\,dB_t$$

with $X_0 = x_0$, $a, b \in \mathbb{R}$, and B is a Brownian motion. The Euler–Maruyama scheme for this equation is given in [16] as

$$X_{\Delta t}^{j+1} = (1 + a\Delta t + b \triangle B_j) X_{\Delta t}^j,$$

where we set $\triangle B_j = B_{t_{j+1}} - B_{t_j}$, and the corresponding Milstein scheme as

$$X_{\Delta t}^{j+1} = \big(1 + a\Delta t + b \triangle B_j + \tfrac{1}{2} b^2 ((\triangle B_j)^2 - \triangle t)\big) X_{\Delta t}^j.$$

These two schemes are $\mathcal{F}$–compatible with the geometric Brownian motion which can be seen by easy calculations and with properties of the conditional expectation as presented in [13].

The truncation error is introduced for PDEs in [9], [26], and [30] for example. Note that this definitions vary by a factor of $\triangle t$ from the following definition because SPDEs are integral equations and not differential equations in the classical sense and therefore we do not divide by $\triangle t$.

Definition 5. The *truncation error* $T(h, \triangle t) = (T_j(h, \triangle t),\, j \in \{1, \ldots, n\})$ of a discretization scheme given by Equation (4) is defined by

$$T_j(h, \triangle t) = X(t_j) - D_h(\triangle t, j-1) P_h X(t_{j-1}).$$

The corresponding *deterministic truncation error* $T^d(h, \triangle t) = (T^d_j(h, \triangle t),\, j \in \{1, \ldots, n\})$ with respect to Equation (3) is defined by

$$T^d_j(h, \triangle t) = u(t_j) - D^d_h(\triangle t) P_h u(t_{j-1}).$$

With the previous definitions we are now able to define consistency which consists of three parts. In order to make the definition compatible with the deterministic one such that this definition extends the known ones, we ask for consistency of the corresponding deterministic problem. Furthermore we need a weaker condition for the SPDE due to the properties of Itô integrals. Finally compatibility is necessary to preserve the properties of stochastic Itô integrals in the approximations. Note that the missing $\triangle t$ in the deterministic truncation error is included in the consistency condition and therefore similar to consistency as defined in [9, 30].

Definition 6. A discretization scheme given by Equation (4) is *consistent in mean square* with Equation (1), if for all $\epsilon > 0$ there exist $\eta, \delta > 0$ such that for all $0 < h < \eta$, $0 < \triangle t < \delta$, and $j \in \{1, \ldots, n\}$

$$\mathbb{E}(\|T_j(h, \triangle t)\|_H^2) < \epsilon \, \triangle t \quad \text{and} \quad \|T^d_j(h, \triangle t)\|_H < \epsilon \, \triangle t,$$

as well as $(D^s_h(\triangle t, j), j \in \{0, \ldots, n-1\})$ is $\mathcal{F}$–compatible.

Remark 7. We remark that the definition of consistency requires for the SPDE convergence of the truncation error of order $\epsilon\sqrt{\triangle t}$ while convergence of order $\epsilon \triangle t$ is necessary for the corresponding deterministic problem. The calculations in the proof of Theorem 11 will show that this order of convergence is sufficient.

A direct consequence of the definition of consistency are the following lemmas that show properties of the approximation of the stochastic integral.

Lemma 8. *The approximation of the stochastic integral satisfies that for all* $\epsilon > 0$ *there exist* $\eta, \delta > 0$ *such that for all* $0 < h < \eta$, $0 < \triangle t < \delta$, *and* $j \in \{1, \ldots, n\}$

$$\mathbb{E}(\|T^s_j(h, \triangle t)\|_H^2) = \mathbb{E}\Big(\Big\| \int_{t_j}^{t_{j+1}} S(t_{j+1} - s) \, G(X(s)) \, dM(s) - D^s_h(\triangle t, j) P_h X(t_j) \Big\|_H^2 \Big) < \epsilon \triangle t.$$

Proof. To prove the lemma we use that the scheme is consistent and that the deterministic part satisfies a consistency condition separately. We estimate in the following way

$$\begin{aligned} \mathbb{E}(\|T^s_j(h, \triangle t)\|_H^2) &\leq 2 \left(\mathbb{E}(\|T_j(h, \triangle t)\|_H^2) + \mathbb{E}(\|(S(\triangle t) - D^d_h(\triangle t) P_h) X(t_j)\|_H^2) \right) \\ &< 2 \left(\epsilon \, \triangle t + \epsilon^2 \, (\triangle t)^2 \right), \end{aligned}$$

where we used that $\mathbb{E}(\|X(t)\|_H^2)$ is bounded for all $t \in [0, T]$, and the lemma is proved. $\square$

This lemma implies a second lemma on the properties of the operator $D^s_h(\triangle t, \cdot)$ that will be needed for later estimates.

Lemma 9. *The approximation of the stochastic integral satisfies for all square-integrable H–valued, $\mathcal{F}_{t_j}$–measurable random variables X, Y, that for all $\epsilon > 0$ there exist $\eta, \delta, C > 0$, such that for all $0 < h < \eta$, $0 < \triangle t < \delta$, and $j \in \{1, \ldots, n\}$*

$$\mathbb{E}(\|D_h^s(\triangle t, j)P_h X - D_h^s(\triangle t, j)P_h Y\|_H^2) \leq \epsilon \triangle t + C\triangle t \mathbb{E}(\|X - Y\|_H^2).$$

Proof. First we remark that for $t \geq t_j$ the SPDE is given by

$$X(t) = S(t - t_j)X(t_j) + \int_{t_j}^{t} S(t-s)\, G(X(s))\, dM(s).$$

If we set $X(t_j) = X, Y$ we have that

$$\begin{aligned}
&\mathbb{E}(\|D_h^s(\triangle t, j)P_h X - D_h^s(\triangle t, j)P_h Y\|_H^2) \\
&\quad \leq 5\Big(2\,\mathbb{E}(\|T_j^s(h, \triangle t)\|_H^2) + \mathbb{E}\big(\|\int_{t_j}^{t_{j+1}} S(t_{j+1} - s)\,(G(X(s)) - G(X))\, dM(s)\|_H^2\big) \\
&\qquad + \mathbb{E}\big(\|\int_{t_j}^{t_{j+1}} S(t_{j+1} - s)\,(G(Y(s)) - G(Y))\, dM(s)\|_H^2\big) \\
&\qquad + \mathbb{E}\big(\|\int_{t_j}^{t_{j+1}} S(t_{j+1} - s)\,(G(X) - G(Y))\, dM(s)\|_H^2\big)\Big).
\end{aligned}$$

The first expression is bounded by $\epsilon\, \triangle t$ by Lemma 8. The last expression satisfies by Proposition 1 that

$$\begin{aligned}
&\mathbb{E}\big(\|\int_{t_j}^{t_{j+1}} S(t_{j+1} - s)\,(G(X) - G(Y))\, dM(s)\|_H^2\big) \\
&\qquad \leq \mathbb{E}\big(\int_{t_j}^{t_{j+1}} \|S(t_{j+1} - s)\,(G(X) - G(Y))\|_{L_{HS}(\mathcal{H},H)}^2\, ds\big).
\end{aligned}$$

The boundedness of the semigroup [24] and Equation (2) imply

$$\begin{aligned}
&\mathbb{E}\big(\int_{t_j}^{t_{j+1}} \|S(t_{j+1} - s)\,(G(X) - G(Y))\|_{L_{HS}(\mathcal{H},H)}^2\, ds\big) \\
&\qquad \leq C\, \mathbb{E}\big(\int_{t_j}^{t_{j+1}} \|X - Y\|_H^2\, ds\big) = C\, \triangle t\, \mathbb{E}(\|X - Y\|_H^2),
\end{aligned}$$

where C denotes a generic constant that changes. It remains to show that the two expressions in the middle go faster to zero than $\triangle t$. As the estimates are the same for X and Y, we just give those for X. First we observe that we have similarly to the previous estimates

$$\mathbb{E}\big(\|\int_{t_j}^{t_{j+1}} S(t_{j+1} - s)\,(G(X(s)) - G(X))\, dM(s)\|_H^2\big) \leq C\, \mathbb{E}\big(\int_{t_j}^{t_{j+1}} \|X(s) - X\|_H^2\, ds\big).$$

This can be bounded by $\epsilon\, \triangle t$, if $\mathbb{E}(\|X(s) - X\|_H^2)$ goes to zero for $s \to t_j$. But this is true due to the properties of the solution, i.e.

$$\mathbb{E}(\|X(s) - X\|_H^2) \leq 2\Big(\mathbb{E}(\|(S(s - t_j) - \mathbb{1})X\|_H^2) + \mathbb{E}\big(\|\int_{t_j}^{s} S(s-r)G(X(r))\, dM(r)\|_H^2\big)\Big),$$

where the first expression tends to zero because S is a C_0–semigroup (see e.g. [24]) and the second expression is bounded by

$$\mathbb{E}\Big(\int_{t_j}^{s} \|S(s-r)G(X(r))\|^2_{L_{HS}(\mathcal{H},H)}\,dr\Big)$$

which tends to zero due to the boundedness of the integrand. So we conclude that

$$\mathbb{E}(\|D_h^s(\triangle t,j)P_hX - D_h^s(\triangle t,j)P_hY\|_H^2) \le \epsilon\triangle t + C\triangle t\mathbb{E}(\|X-Y\|_H^2),$$

which proves the lemma. □

Finally we define stability in the sense of Lax and Richtmyer which could also be called numerical stability in order to avoid confusions with other concepts of stability like Lyapunov stability or asymptotic stability. It turns out that the extension of a PDE to an SPDE and a deterministic approximation scheme to a stochastic one does not affect the stability of the scheme. In the proof of a stochastic version of Lax's equivalence theorem it turns out that just stability of the corresponding deterministic scheme is necessary for this type of SPDEs.

Definition 10. A discretization scheme defined by Equation (4) is *stable*, if there exists $K \ge 1$ such that for all h, $\triangle t > 0$ and all $j \in \{0,\ldots,n\}$ it holds that

$$\|(D_h^d(\triangle t))^j P_h\|_{L(H)} \le K.$$

where $L(H)$ denotes the space of all linear mappings from H into itself.

Examples of stable discretization schemes for a given SPDE are all approximations, where the approximation of the corresponding PDE is stable.

3. Lax Equivalence Theorem

In this section we state and prove the main result of this paper.

Theorem 11 (Stochastic Mean Square Lax Equivalence Theorem)**.** *Assume that a consistent approximation scheme defined by Equation* (4) *with respect to an SPDE of type* (1) *is given. Then it is convergent in mean square if and only if it is stable.*

The following lemma will be essential in the proof of the main result.

Lemma 12. *Let $(\Omega,\mathcal{A},P)$ be a probability space and $\mathcal{B} \subset \mathcal{A}$ a σ–algebra. Furthermore assume that X, Y are $(H,\mathbb{H})$–valued random variables with $\mathbb{E}(\|X\|_H^2), \mathbb{E}(\|Y\|_H^2) < +\infty$. If Y is also $\mathcal{B}/\mathbb{H}$–measurable, then*

$$\mathbb{E}((X,Y)_H|\mathcal{B}) = (\mathbb{E}(X|\mathcal{B}),Y)_H.$$

Proof. To prove the lemma, we first use the separability of the Hilbert space and the existence of an orthonormal basis. This transforms the problem into a real valued one. As X and Y are in L^2, the dominated convergence theorem for conditional expectations of real valued random variables (see e.g. [4]) can be applied. Parseval's relation, the continuity of the inner product and Equation (3.7.5) in [12] conclude the proof. □

Proof of Theorem 11. We first assume that the approximation scheme is stable and consistent and show that it converges in mean square, i.e. for n large enough and all $j \in \{0,\ldots,n\}$

$$\mathbb{E}(\|X(t_j) - X_h^j\|_H^2) < \epsilon.$$

We observe that X_h^j can be rewritten as

$$X_h^j = (D_h^d(\triangle t))^j X_{h,0} + \sum_{i=0}^{j-1} (D_h^d(\triangle t))^{j-(i+1)} D_h^s(\triangle t, i) X_h^i.$$

This implies for the difference of the mild solution and the approximation that

$$\begin{aligned}
X(t_j) - X_h^j = {} & \big(S(t_j) - (D_h^d(\triangle t))^j P_h\big) X_0 \\
& + \sum_{i=0}^{j-1} \big(S(t_j - t_{i+1}) - (D_h^d(\triangle t))^{j-(i+1)} P_h\big) \int_{t_i}^{t_{i+1}} S(t_{i+1} - s) G(X(s))\, dM(s) \\
& + \sum_{i=0}^{j-1} (D_h^d(\triangle t))^{j-(i+1)} P_h \Big(\int_{t_i}^{t_{i+1}} S(t_{i+1} - s) G(X(s))\, dM(s) - D_h^s(\triangle t, i) P_h X(t_i) \Big) \\
& + \sum_{i=0}^{j-1} (D_h^d(\triangle t))^{j-(i+1)} (D_h^s(\triangle t, i) P_h X(t_i) - D_h^s(\triangle t, i) X_h^i)
\end{aligned}$$

and for the expression to be estimated by Hölder's inequality

$$\begin{aligned}
& \mathbb{E}(\|X(t_j) - X_h^j\|_H^2) \\
& \quad \le 4 \Big(\mathbb{E}\big(\| \big(S(t_j) - (D_h^d(\triangle t))^j P_h\big) X_0 \|_H^2\big) \\
& \qquad + \mathbb{E}\big(\| \sum_{i=0}^{j-1} \big(S(t_j - t_{i+1}) - (D_h^d(\triangle t))^{j-(i+1)} P_h\big) \int_{t_i}^{t_{i+1}} S(t_{i+1} - s) G(X(s))\, dM(s) \|_H^2\big) \\
& \qquad + \mathbb{E}\big(\| \sum_{i=0}^{j-1} (D_h^d(\triangle t))^{j-(i+1)} P_h \big(\int_{t_i}^{t_{i+1}} S(t_{i+1} - s) G(X(s))\, dM(s) - D_h^s(\triangle t, i) P_h X(t_i) \big) \|_H^2\big) \\
& \qquad + \mathbb{E}\big(\| \sum_{i=0}^{j-1} (D_h^d(\triangle t))^{j-(i+1)} (D_h^s(\triangle t, i) P_h X(t_i) - D_h^s(\triangle t, i) X_h^i) \|_H^2\big) \Big).
\end{aligned}$$

Next, we give estimates on each of the four expressions before finishing the first implication. To the first term we apply the classical Lax equivalence theorem [9] as the approximation scheme of the corresponding PDE is consistent and stable. Therefore the first term is smaller than any ϵ for h and $\triangle t$ small enough. For the second term we set

$$R_i = S(t_j - t_{i+1}) - (D_h^d(\triangle t))^{j-(i+1)} P_h$$

and have for $i < k$ by Lemma 12 and the properties of the conditional expectation as well as of the stochastic integral

$$\begin{aligned}
& \mathbb{E}\big((R_i \int_{t_i}^{t_{i+1}} S(t_{i+1} - s) G(X(s))\, dM(s), R_k \int_{t_k}^{t_{k+1}} S(t_{k+1} - s) G(X(s))\, dM(s))_H \big) \\
& \quad = \mathbb{E}\big((R_i \int_{t_i}^{t_{i+1}} S(t_{i+1} - s) G(X(s))\, dM(s), R_k\, \mathbb{E}\big(\int_{t_k}^{t_{k+1}} S(t_{k+1} - s) G(X(s))\, dM(s) | \mathcal{F}_{t_k} \big))_H \big) \\
& \quad = 0.
\end{aligned}$$

This implies that

$$
\begin{aligned}
&\mathbb{E}\big(\|\sum_{i=0}^{j-1}\big(S(t_j-t_{i+1})-(D_h^d(\triangle t))^{j-(i+1)}P_h\big)\int_{t_i}^{t_{i+1}}S(t_{i+1}-s)G(X(s))\,dM(s)\|_H^2\big)\\
&\qquad=\mathbb{E}\big(\sum_{i=0}^{j-1}\|\big(S(t_j-t_{i+1})-(D_h^d(\triangle t))^{j-(i+1)}P_h\big)\int_{t_i}^{t_{i+1}}S(t_{i+1}-s)G(X(s))\,dM(s)\|_H^2\big)\\
&\qquad\le\epsilon^2\sum_{i=0}^{j-1}\mathbb{E}\big(\|\int_{t_i}^{t_{i+1}}S(t_{i+1}-s)G(X(s))\,dM(s)\|_H^2\big),
\end{aligned}
$$

where the last inequality follows from the convergence of the corresponding deterministic problem. Finally we apply the properties of the semigroup, Proposition 1, and the assumptions made in (2) to get

$$
\begin{aligned}
&\mathbb{E}\big(\|\sum_{i=0}^{j-1}\big(S(t_j-t_{i+1})-(D_h^d(\triangle t))^{j-(i+1)}P_h\big)\int_{t_i}^{t_{i+1}}S(t_{i+1}-s)G(X(s))\,dM(s)\|_H^2\big)\\
&\qquad\le\epsilon^2\,C\sum_{i=0}^{j-1}\int_{t_i}^{t_{i+1}}\big(1+\mathbb{E}(\|X(s)\|_H^2)\big)\,ds.
\end{aligned}
$$

The claim follows by the boundedness of the solution on $[0,T]$.

The compatibility of the approximation implies in a similar calculation as for the second term that the mixed expressions in the fourth term are zero. The stability of the approximation and Lemma 9 lead for this term to

$$
\begin{aligned}
&\mathbb{E}\big(\|\sum_{i=0}^{j-1}(D_h^d(\triangle t))^{j-(i+1)}(D_h^s(\triangle t,i)P_hX(t_i)-D_h^s(\triangle t,i)X_h^i)\|_H^2\big)\\
&\qquad=\sum_{i=0}^{j-1}\mathbb{E}\big(\|(D_h^d(\triangle t))^{j-(i+1)}(D_h^s(\triangle t,i)P_hX(t_i)-D_h^s(\triangle t,i)X_h^i)\|_H^2\big)\\
&\qquad\le K^2\,(T\epsilon+C\sum_{i=0}^{j-1}\triangle t\,\mathbb{E}(\|X(t_i)-X_h^i\|_H^2)).
\end{aligned}
$$

The mixed expressions of the third term are split into four terms and satisfy by the properties of the stochastic integral, the compatibility of the approximation and Lemma 12 for $i<k$, if

we set $D_i = (D_h^d(\triangle t))^{j-(i+1)} P_h$,

$$\mathbb{E}\big(\big(D_i\big(\int_{t_i}^{t_{i+1}} S(t_{i+1}-s)G(X(s))\,dM(s) - D_h^s(\triangle t, i)P_h X(t_i)\big),$$
$$D_k\big(\int_{t_k}^{t_{k+1}} S(t_{k+1}-s)G(X(s))\,dM(s) - D_h^s(\triangle t, k)P_h X(t_k)\big)\big)_H\big)$$
$$\begin{aligned}
&= \mathbb{E}\big(\big(D_i \int_{t_i}^{t_{i+1}} S(t_{i+1}-s)G(X(s))\,dM(s), D_k\,\mathbb{E}\big(\int_{t_k}^{t_{k+1}} S(t_{k+1}-s)G(X(s))\,dM(s)|\mathcal{F}_{t_k}\big)\big)_H\big)\\
&\quad + \mathbb{E}\Big(\big(D_i D_h^s(\triangle t, i)P_h X(t_i), D_k\,\mathbb{E}(D_h^s(\triangle t, k)|\mathcal{F}_{t_k})P_h X(t_k)\big)_H\Big)\\
&\quad - \mathbb{E}\Big(\big(D_i \int_{t_i}^{t_{i+1}} S(t_{i+1}-s)G(X(s))\,dM(s), D_k\,\mathbb{E}(D_h^s(\triangle t, k)|\mathcal{F}_{t_k})P_h X(t_k)\big)_H\Big)\\
&\quad - \mathbb{E}\Big(\big(D_i D_h^s(\triangle t, i)P_h X(t_i), D_k\,\mathbb{E}\big(\int_{t_k}^{t_{k+1}} S(t_{k+1}-s)G(X(s))\,dM(s)|\mathcal{F}_{t_k}\big)\big)_H\Big)
\end{aligned}$$

which is equal to zero as seen in the previous estimates. This implies for the third term combined with the stability of the approximation and its consistency with Lemma 8

$$\begin{aligned}
\mathbb{E}\big(\|\sum_{i=0}^{j-1}(D_h^d(\triangle t))^{j-(i+1)}P_h\big(\int_{t_i}^{t_{i+1}} S(t_{i+1}-s)G(X(s))\,dM(s) - D_h^s(\triangle t, i)P_h X(t_i)\big)\|_H^2\big)\\
\leq \sum_{i=0}^{j-1} K^2\epsilon\triangle t \leq K^2 T\epsilon.
\end{aligned}$$

So overall we have by a discrete version of Gronwall's inequality [32]

$$\mathbb{E}(\|X(t_j) - X_h^j\|_H^2) \leq \epsilon\, C_1 + C_2\triangle t\sum_{i=0}^{j-1}\mathbb{E}(\|X(t_i) - X_h^i\|_H^2) \leq \epsilon\, C$$

which proves convergence in mean square.

Next we prove that convergence in mean square implies the stability of the approximation scheme by contradiction. Therefore we assume that the approximation scheme is not stable, i.e. for any $K > 0$ there exist $\triangle t$, h, j, and $X_0 \in H$ such that

$$\|(D_h^d(\triangle t))^j P_h X_0\|_H > K.$$

This implies by the deterministic Lax equivalence theorem [9] that the deterministic scheme does not converge to the corresponding PDE, i.e. there exists $R > 0$ such that for all $\eta, \delta > 0$ there exist $0 < h < \eta$, $0 < \triangle t < \delta$, $j \in \{0, \ldots, n\}$ with

$$\|(S(t_j) - (D_h^d(\triangle t))^j P_h)X_0\|_H > R.$$

The properties of the expectation and the integral as well as Cauchy–Schwartz's inequality imply

$$\begin{aligned}
R < \|(S(t_j) - (D_h^d(\triangle t))^j P_h)X_0\|_H = \|\,\mathbb{E}(X(t_j) - X_h^j)\|_H \leq \mathbb{E}(\|X(t_j) - X_h^j\|_H)\\
\leq \mathbb{E}(\|X(t_j) - X_h^j\|_H^2)^{1/2},
\end{aligned}$$

i.e. the scheme does not converge in mean square and the theorem is proved. □

Remark 13. In the case of additive noise, Theorem 11 also holds for approximations of equations where G is a mapping from the interval $[0,T]$ into the linear operators $L(U,H)$ that satisfies for all $t \in [0,T]$

$$\|G(t)\|_{L_{HS}(\mathcal{H},H)} < C$$

for some constant $C \in \mathbb{R}_+$. The proof is similar to the one of the given theorem and therefore omitted.

4. Examples

In this section an example of the heat equation is given to emphasize how the definitions and the main result of this paper are related to practical problems. We will look at the heat equation with multiplicative noise approximated by an Euler–Maruyama as well as a Milstein scheme. More examples can be found in [18].

Consider the following heat equation on $H = L^2([0,2\pi))$ and on the finite time interval $[0,T]$

$$dX(t) = \tfrac{1}{2}\Delta X(t)\,dt + G(X(t))\,dW(t) \tag{5}$$

with initial condition $X(0) = X_0 \in H$, W is a Q–Wiener process on H as introduced in [8] with $\operatorname{Tr} Q < +\infty$. The operator G, which is a linear mapping from H to $L(H,H)$, is given by

$$G(\phi)\psi(x) = g(x)\,\phi(x)\,\psi(x)$$

for $g, \phi, \psi : [0,2\pi) \to \mathbb{R}$. Furthermore we assume the Laplace operator on $[0,2\pi)$ with periodic boundary conditions. Then it generates a semigroup of contractions which we denote by $S = (S(t), t \in [0,T])$. Let W be given by

$$W(t) = \sum_{k=0}^{\infty} \sqrt{a_k}\,\beta_k(t)\,e_k,$$

where $(e_k, k \in \mathbb{N}_0)$ is the $L^2([0,2\pi))$ orthonormal basis consisting of sine and cosine functions, the elements $\beta_k(t)$ are real valued, independent Brownian motions, and the coefficients a_k are given by

$$a_k = (m^l + k^l)^{-n}$$

for $m \in \mathbb{R}_+$ and $l/2, n \in \mathbb{N}$. This equation has for $g \in C_B^4([0,2\pi))$ and $l \cdot n > 10$ a mild solution that satisfies for $A = 1/2\,\Delta$

$$\mathbb{E}(\|A^2 X(t)\|_H^2) < K$$

for all $t \in [0,T]$ which can be shown by meeting the prerequisites of Theorem 6.7 in [8]. This condition implies that a finite difference approximation of A converges of order $\triangle t + (\triangle x)^2$ [30], where $\triangle t$ denotes the equidistant step size in time and $\triangle x$ the one in space. Let the Euler–Maruyama and Milstein approximations of Equation (5) be given by

$$X^{j+1} = (1 + \triangle t A_h + g\,\eta_j) X^j, \tag{E}$$

$$X^{j+1} = (1 + \triangle t A_h + g\,\eta_j + \tfrac{1}{2}\,g^2\,(\eta_j^2 - \triangle t)) X^j \tag{M}$$

with

$$(g\,\eta_j X^j)(x) = G(X^j)\eta_j(x) = g(x) \sum_{k=0}^{\lfloor 2\pi/\triangle x \rfloor - 1} \sqrt{a_k}\,(\beta_k(t_{j+1}) - \beta_k(t_j))\,e_k(x)\,X^j(x),$$

i.e. all multiplications of functions are pointwise in x. The properties of the elements η_j and how to simulate them efficiently can be found in [20]. The operator A_h denotes a finite difference approximation and is given by

$$A_h f(x) = \frac{f(x+\triangle x) - 2f(x) + f(x-\triangle x)}{2(\triangle x)^2},$$

where calculations in x are done modulo 2π.

The corresponding PDE

$$\frac{\partial}{\partial t} u(t) = \tfrac{1}{2}\Delta u(t)$$

with the approximation scheme given by

$$u^{j+1} = (1 + \triangle t A_h) u^j$$

is known to be consistent, stable and convergent of order $\triangle t + (\triangle x)^2$ in the deterministic sense for $\triangle t \leq (\triangle x)^2$ (see e.g. [11],[26],[30]). This implies stability of the approximation schemes given by Equation (E) and (M) for $\triangle t \leq (\triangle x)^2$. The approximations of the stochastic integral are compatible with Equation (5) by the properties of the Brownian motion in a similar way as in Example 4. For consistency in mean square it remains to show that the truncation error converges in mean square faster than $\triangle t$. We first look at the Euler–Maruyama scheme. For $j \in \{1, \ldots, n\}$ we have that

$$\begin{aligned}\mathbb{E}(\|T_j(h,\triangle t)\|_H^2) \leq 2 \big(&\mathbb{E}(\|(S(\triangle t) - (1+\triangle t A_h))X(t_{j-1})\|_H^2) \\ &+ \mathbb{E}(\| \int_{t_{j-1}}^{t_j} S(t_j - s) G(X(s))\, dW(s) - G(X(t_{j-1}))\eta_{j-1}\|_H^2)\big)\end{aligned}$$

by Hölder's inequality. The first term is by the properties of the corresponding deterministic problem of order $O((\triangle t)^4 + (\triangle t)^2(\triangle x)^4)$ as $\mathbb{E}(\|A^2 X(t_j)\|_H^2)$ is bounded. The second term is split into

$$\begin{aligned}&\int_{t_{j-1}}^{t_j} S(t_j - s) G(X(s))\, dW(s) - G(X(t_{j-1}))\,\eta_{j-1} \\ &\qquad = \int_{t_{j-1}}^{t_j} (S(t_j - s) - \mathbb{1}) G(X(s))\, dW(s) + \int_{t_{j-1}}^{t_j} G(X(s) - X_{t_{j-1}})\, dW(s) \\ &\qquad\qquad + G(X_{t_{j-1}})((W(t_j) - W(t_{j-1})) - \eta_{j-1}).\end{aligned}$$

The first of these three terms is by the properties of the semigroup [24] and of the stochastic integral of order $O((\triangle t)^2)$. For the second term the regularity of the solution is needed which is calculated in Lemma 3.3 in [19]. Overall by the properties of the stochastic integral the term is also of order $O((\triangle t)^2)$. The last term can be estimated with the property that Q is nuclear. This implies that it is of order $O(\triangle t(\triangle x)^{ln-1})$ and consistency in mean square of order $O((\triangle t)^2(\triangle x)^4 + (\triangle t)^2 + \triangle t(\triangle x)^{ln-1})$ follows. Similar calculations for the Milstein approximation given by Equation (M) lead to consistency in mean square of order $O((\triangle t)^2(\triangle x)^4 + (\triangle t)^3 + \triangle t(\triangle x)^{ln-1})$. To prove convergence in mean square we do similar estimates as for convergence in the proof of Lax's equivalence theorem. These lead for the Euler scheme with the properties of the approximation of the corresponding PDE to

$$\mathbb{E}(\|e_j(h,\triangle t)\|_H^2) = O(\triangle t + (\triangle x)^4),$$

i.e. convergence of order $O(\sqrt{\triangle t} + (\triangle x)^2)$ and for the Milstein scheme to convergence of order $O(\triangle t + (\triangle x)^2)$ always under the assumption that $\triangle t \leq (\triangle x)^2$.

Finally we remark that the conditions on the discretization step size for an explicit approximation of a stochastic heat equation are the same as those for the deterministic heat equation. The only difference that occurs is the worse order of convergence in time for the simplest approximation of the Itô integral.

Examples with additive noise can be found in [18] and with Lévy noise and finite element methods in [2] and [3].

References

1. Andrea Barth, *Stochastic partial differential equations: Approximations and applications*, Ph.D. thesis, University of Oslo, CMA, September 2009.
2. Andrea Barth, *A finite element method for martingale-driven stochastic partial differential equations*, Comm. Stoch. Anal. **4** (2010), no. 3.
3. Andrea Barth and Annika Lang, *Almost sure convergence of a Galerkin–Milstein approximation for stochastic partial differential equations driven*, submitted, 2009.
4. Heinz Bauer, *Probability theory. (Wahrscheinlichkeitstheorie.) 4., völlig überarb. u. neugestaltete Aufl. des Werkes: Wahrscheinlichkeitstheorie und Grundzüge der Maßtheorie*, De Gruyter Lehrbuch. Berlin etc.: Walter de Gruyter, 1991.
5. Fred Espen Benth, Jurate Šaltyte Benth, and Steen Koekebakker, *Stochastic modelling of electricity and related markets*, World Scientific, Advanced Series on Statistical Science and Applied Probability - Vol. 11, 2008.
6. Evelyn Buckwar and Renate Winkler, *Multistep methods for SDEs and their application to problems with small noise*, SIAM J. Numer. Anal. **44** (2006), no. 2, 779–803.
7. Alexandre J. Chorin, Thomas J.R. Hughes, Marjorie F. McCracken, and Jerrold E. Marsden, *Product formulas and numerical algorithms*, Commun. Pure Appl. Math. **31** (1978), 205–256.
8. Giuseppe Da Prato and Jerzy Zabczyk, *Stochastic equations in infinite dimensions*, Encyclopedia of Mathematics and Its Applications. 44. Cambridge: Cambridge University Press, 1992.
9. Robert Dautray and Jacques-Louis Lions, *Mathematical analysis and numerical methods for science and technology. Vol. 6: Evolution problems ii. 2nd printing*, Berlin: Springer, 2000.
10. John Gary, *A generalization of the Lax–Richtmyer theorem on finite difference schemes*, SIAM J. Numer. Anal. **3** (1966), 467–473.
11. Christian Grossmann, Hans-Görg Roos, and Martin Stynes, *Numerical treatment of partial differential equations. revised translation of the 3rd german edition of 'numerische behandlung partieller differentialgleichungen' by martin stynes.*, Universitext. Berlin: Springer, 2007.
12. Einar Hille and Ralph S. Phillips, *Functional analysis and semi-groups*, Colloquium Publications, 31. Providence, R.I.: American Mathematical Society (AMS), 1981.
13. Nobuyuki Ikeda and Shinzo Watanabe, *Stochastic differential equations and diffusion processes. 2nd ed.*, North-Holland Mathematical Library, 24. Amsterdam etc.: North-Holland; Tokyo: Kodansha Ltd., 1989.
14. Arnulf Jentzen, *Taylor expansions for stochastic partial differential equations*, Ph.D. thesis, Johann Wolfgang Goethe Universität Frankfurt am Main, 2009.
15. Olivier Juan, Renaud Keriven, and Gheorghe Postelnicu, *Stochastic motion and the level set method in computer vision: Stochastic active contours*, International Journal of Computer Vision **69** (2006), no. 1, 7–25.
16. Peter E. Kloeden and Eckhard Platen, *Numerical solution of stochastic differential equations*, Applications of Mathematics. 23. Berlin: Springer, 1992.
17. Heinz-Otto Kreiss, *Numerical methods for solving time-dependent problems for partial differential equations*, Séminaire de Mathématiques Superieures, Département de Mathématiques et de Statistique, Université de Montréal, 14ème Session, 16 juin au 11 juillet 1975. Montréal: Les Presses de l'Université de Montréal, 1978.
18. Annika Lang, *Simulation of stochastic partial differential equations and stochastic active contours*, Ph.D. thesis, Universität Mannheim, 2007.

19. Annika Lang, Pao-Liu Chow, and Jürgen Potthoff, *Almost sure convergence of a semidiscrete Milstein scheme for SPDEs of Zakai type*, to appear in Stochastics, 2009.
20. Annika Lang and Jürgen Potthoff, *Fast simulation of Gaussian random fields*, Submitted, 2009.
21. Peter D. Lax and Robert D. Richtmyer, *Survey of the stability of linear finite difference equations*, Commun. Pure Appl. Math. **9** (1956), 267–293.
22. Yiming Lou and P.D. Christofides, *Nonlinear feedback control of surface roughness using a stochastic PDE*, Decision and Control, 2006 45th IEEE Conference on, Dec. 2006, pp. 936–943.
23. Dong Ni and Panagiotes D. Christofide, *Construction of stochastic PDEs for feedback control of surface roughness in thin film deposition*, Proceedings of the 2005 American Control Conference, vol. 4, June 2005, pp. 2540–2547.
24. Amnon Pazy, *Semigroups of linear operators and applications to partial differential equations*, Applied Mathematical Sciences, 44. New York – Berlin – Heidelberg – Tokyo: Springer, 1983.
25. Szymon Peszat and Jerzy Zabczyk, *Stochastic partial differential equations with Lévy noise. An evolution equation approach*, Encyclopedia of Mathematics and Its Applications 113. Cambridge: Cambridge University Press, 2007.
26. Robert D. Richtmyer and K.W. Morton, *Difference methods for initial-value problems. Repr. of the 2nd ed. 1967*, Melbourne, FL: Krieger Publishing Company, 1994.
27. Christian Roth, *Stochastische partielle differentialgleichungen 1.ordnung*, Ph.D. thesis, Martin-Luther-Universitt Halle-Wittenberg, July 2002.
28. Christian Roth, *Weak approximations of solutions of a first order hyperbolic stochastic partial differential equation*, Monte Carlo Methods Appl. **13** (2007), no. 2, 117–133.
29. Henri Schurz, *General theorems for numerical approximation of stochastic processes on the Hilbert space* $H_2([0,T],\mu,\mathbb{R}^d)$, ETNA, Electron. Trans. Numer. Anal. **16** (2003), 50–69.
30. Granville Sewell, *The numerical solution of ordinary and partial differential equations. 2nd ed.*, Pure and Applied Mathematics. A Wiley-Interscience Series of Texts, Monographs, and Tracts. Hoboken, NJ: John Wiley & Sons, 2005.
31. Tony Shardlow, *Numerical simulation of stochastic PDEs for excitable media*, J. Comput. Appl. Math. **175** (2005), no. 2, 429–446.
32. Lü Tao and Huang Yong, *A generalization of discrete Gronwall inequality and its application to weakly singular Volterra integral equation of the second kind*, J. Math. Anal. Appl. **282** (2003), no. 1, 56–62.
33. Vidar Thomée, *Galerkin finite element methods for parabolic problems. 2nd revised and expanded ed.*, Berlin: Springer, 2006.

Annika Lang
Fakultät für Mathematik und Informatik, Universität Mannheim
D–68131 Mannheim, Germany
E-mail address: `lang@math.uni-mannheim.de`