SIAM J. NUMER. ANAL.
Vol. 50, No. 3, pp. 1735–1752

WEAK BACKWARD ERROR ANALYSIS FOR SDEs*

ARNAUD DEBUSSCHE† AND ERWAN FAOU‡

Abstract. We consider long time numerical approximations of stochastic differential equations (SDEs) by the Euler method. In the case where the SDE is elliptic or hypoelliptic, we show a weak backward error analysis result in the sense that the generator associated with the numerical solution coincides with the solution of a modified Kolmogorov equation up to high order terms with respect to the stepsize. This implies that every invariant measure of the numerical scheme is close to a modified invariant measure obtained by asymptotic expansion. Moreover, we prove that, up to negligible terms, the dynamic associated with the Euler scheme is exponentially mixing.

Key words. backward error analysis, stochastic differential equations, exponential mixing, numerical scheme, Kolmogorov equation, weak error

AMS subject classifications. 65C30, 60H35, 37M25

DOI. 10.1137/110831544

1. Introduction.

In recent decades, *backward error analysis* has become one of the most powerful tools for analyzing the long time behavior of numerical schemes applied to evolution equations. The main idea can be described as follows: Let us consider an ordinary differential equation of the form

$$\dot{y}(t) = f(y(t)),$$

where $f : \mathbb{R}^n \to \mathbb{R}^n$ is a smooth vector field, and denote by $\varphi_t^f(y)$ the associated flow. By definition, a numerical method defines for a small time step τ an approximation Φ_τ of the exact flow φ_τ: We have, for bounded $y \in \mathbb{R}^n$, $\Phi_\tau(y) = \varphi_\tau^f(y) + \mathcal{O}(\tau^{r+1})$, where r is the order of the method.

The idea of backward error analysis is to show that Φ_τ can be interpreted as the exact flow $\varphi_\tau^{f_\tau}$ or a *modified vector field* defined as a series in powers of τ,

$$f_\tau = f + \tau^r f_r + \tau^{r+1} f_{r+1} + \cdots ,$$

where f_ℓ, $\ell \geq r$, are vector fields depending on the numerical method. In general, the series defining f_τ does not converge, but it can be shown that for bounded y, we have for arbitrary N

$$\Phi_\tau(y) = \varphi_\tau^{f_\tau^N}(y) + C_N \tau^{N+1},$$

where f_τ^N is the truncated series

$$f_\tau^N = f + \tau^r f_r + \cdots + \tau^N f_N.$$

Under some analyticity assumptions, the constant $C_N \tau^{N+1}$ can be optimized in N, so that the error term in the previous equation can be made exponentially small with respect to τ.

*Received by the editors April 20, 2011; accepted for publication (in revised form) March 12, 2012; published electronically June 28, 2012.
http://www.siam.org/journals/sinum/50-3/83154.html
†Ecole Normale Supérieure de Cachan Bretagne, IRMAR-UMR 6625, Avenue Robert Schumann, F-35170 Bruz, France (Arnaud.Debussche@bretagne.ens-cachan.fr).
‡INRIA & Ecole Normale Supérieure de Cachan Bretagne, IRMAR-UMR 6625, Avenue Robert Schumann, F-35170 Bruz, France (Erwan.Faou@inria.fr).

Such a result is very important and has many applications in the case where f has some strong geometric properties, such as Hamiltonian or reversible structure. In this situation, and under some compatibility conditions on the numerical method Φ_τ, the modified vector field f_τ inherits the structure of f. For example, if Φ_τ is symplectic and f Hamiltonian, then f_τ remains Hamiltonian. This has major consequences such as the preservation of a modified Hamiltonian over a very long time (of order τ^{-N}) for the numerical solution, from which we can deduce long time stability results, existence of numerical invariant tori in the integrable case, etc.

In the Hamiltonian case, this idea goes back to Moser [16], but was applied later to symplectic integrators by Benettin and Giorgilli [4], Hairer and Lubich [8], and Reich [19]. Such results now form the core of the modern *geometric numerical integration theory*; for details we refer the reader to the classical textbooks [9] and [20].

More recently, these ideas have been extended in some situations to Hamiltonian PDEs: first in the linear case [5], and then in the semilinear case (nonlinear Schrödinger or wave equations) [6, 7].

As far as stochastic differential equations (SDEs) are concerned, this approach has not been developed very much so far. Let us recall that given an SDE in $\mathbb{R}^d$ of the form

$$\mathrm{d}X = f(X)\mathrm{d}t + \sigma(X)\mathrm{d}W$$

discretized by the Euler scheme with time step τ providing a discrete sequence $(X_p)_{p\in N}$—see (2.7) below—then the error can be measured in the strong or weak sense. Strong error means that X_p is a pathwise approximation of $X(p\tau)$, and it is well known that the Euler scheme has strong order 1/2. Under standard smoothness assumptions on f and σ, we have (see, for instance, [11, 15, 25])

$$\mathbb{E}\left(\sup_{p=0,\dots,[T/\tau]} \|X_p - X(t_p)\|^k\right) \le c_1(k,T)\tau^{1/2}, \quad k \ge 1,\ T > 0.$$

In this work, we consider another error which is often more important. We investigate the weak error which concerns the law of the solution. The Euler scheme has weak order 1. Under suitable smoothness assumptions on f, σ, and $\varphi : \mathbb{R}^d \to \mathbb{R}$ (see, for instance, [11, 15, 25]), we have

$$|\mathbb{E}(\varphi(X_p)) - \mathbb{E}(\varphi(X(t_p)))| \le c_2(\varphi,T)\tau, \quad p = 0,\dots,[T/\tau], \quad T > 0.$$

An attempt was made by Shardlow [22] to extend the idea of backward error analysis to this context. He showed that the construction of a modified SDE associated with the Euler scheme can be performed, but only at the first step, i.e., for $N = 2$, and only for additive noise, i.e., when $\sigma(X)$ does not depend on X. In this case, he was able to write down a modified SDE:

$$\mathrm{d}\tilde{X} = \tilde{f}(\tilde{X})\mathrm{d}t + \tilde{\sigma}(\tilde{X})\mathrm{d}W \tag{1.1}$$

such that

$$\left|\mathbb{E}(\varphi(X_p)) - \mathbb{E}\left(\varphi(\tilde{X}(t_p))\right)\right| \le c_3(\varphi,T)\tau^2, \quad p = 0,\dots,[T/\tau], \quad T > 0.$$

He explained that for multiplicative noise or higher order, there are too many conditions to be satisfied by the coefficients of the modified equations.

In this paper we take another approach and build a modified equation not at the level of the SDE, but at the level of the generator associated with the process solution of the SDE. It is well known that given $\varphi : \mathbb{R}^d \to \mathbb{R}$ and denoting by $X(t,x)$ the solution of the SDE satisfying $X(0) = x$, the function $u(t,x) = \mathbb{E}(\varphi(X(t,x)))$ satisfies the Kolmogorov equation [23, 24]

$$\partial_t u(t,x) = L(x,\partial_x)u(t,x),$$

where L is the order 2 Kolmogorov operator associated with the SDE; see (2.1) below.

In the case of the Euler method applied to an SDE, we show that with the numerical solution, we can associate the *modified Kolmogorov operator* of the form

$$L(\tau,x,\partial_x) = L(x,\partial_x) + \tau L_1(x,\partial_x) + \tau^2 L_2(x,\partial_x) + \cdots ,$$

where L_ℓ, $\ell \geq 1$, are some modified operator of order $2\ell + 2$. Again, the series does not converge, but the truncated series

$$L^{(N)}(\tau,x,\partial_x) = L(x,\partial_x) + \tau L_1(x,\partial_x) + \cdots + \tau^N L_N(x,\partial_x)$$

are considered. Note that this operator is a singular perturbation of the Kolmogorov operator $L(x,\partial_x)$ and is no longer of order 2 apart from specific situations (see [1, 28]). Hence in contrast with the classical case, we cannot straightforwardly define a solution to the *modified equation*

$$\partial_t v^N(t,x) = L^{(N)}(\tau,x,\partial_x)v^N(t,x), \tag{1.2}$$

and we cannot associated with the operator $L^{(N)}$ a modified SDE of the form (1.1) in general and for arbitrary value of N.

However, in the case where the SDE is elliptic or hypoelliptic, we can build an approximated solution $v^{(N)}$ such that

$$\left|\mathbb{E}\left(\varphi(X_p)\right) - v^{(N)}(p\tau,x)\right| \leq c_4(\varphi,T,N)\tau^{N+1}, \quad p = 0,\ldots,[T/\tau], \quad T > 0. \tag{1.3}$$

This solution is constructed not as an exact solution of the parabolic equation (1.2), but as an asymptotic expansion of this formal modified equation. Furthermore, using the exponential convergence to equilibrium, we prove that in fact the constant c_4 does not depend on T so that we have an approximation result valid on very long times. Note that $v^{(N)}$ is in fact constructed as a truncated series, $v^{(N)} = \sum_{n=0}^{N}\tau^n v_n$, and that $v_0 = u$ is the solution of the Kolmogorov equation. Therefore, our result provides an expansion of the error as in [27] (see also [2, 3]). However, the expansion is different here.

As a consequence, we prove that the numerical solution X_p, $p \geq 1$, obtained by the Euler scheme is exponentially mixing up to some very small error, and for all times, with respect to a modified invariant measure that turns out to be a modified invariant measure for the operator $L^{(N)}(\tau,x,\partial_x)$. This last result—see Theorem 2.1 below—is typically a geometric numerical integration result in the sense that we prove the persistence of a qualitative property of the exact flow (exponential mixing) to the numerical approximation, over long times.

Error estimates on long times for elliptic and hypoelliptic SDEs have already been proved. In [13, 21, 25, 26], it is shown that for a sufficiently small time step the Euler scheme defines an ergodic process and that the invariant measure of the Euler scheme is close to the invariant measure of the SDE. In [27], the first term of an expansion of the invariant measure of the Euler scheme with respect to τ is also given. In our work, we provide the expansion at any order.

We emphasize that in our result there is no particular smallness assumption on the stepsize τ used to define the numerical solution. In particular, the discrete process is not supposed to have a unique invariant measure, as in [13] or [21, 25, 26].

This is also the case in the recent work [14]. There, it is shown that given an elliptic or hypoelliptic SDE, the ergodic averages provided by the Euler scheme are asymptotically close to the average of the invariant measure of the SDE. Higher order schemes are also considered. The main tool in [14] is the ellipticity or hypoellipticity of the Poisson equation, i.e., the equation $L(x, \partial_x)u = g$.

As in [14], we consider the case where the SDE is set on the torus $\mathbb{T}^d$. This simplifies the presentation, and the main ideas are not hidden by technical difficulties. In the same spirit, we study only the Euler scheme. The extension to other situations such as SDEs set on $\mathbb{R}^n$ with polynomial growth coefficients or hypoelliptic equations of the Langevin form as treated in [26] will be studied in forthcoming works.

2. Main result. We consider the SDE

$$\mathrm{d}X(t) = f(X(t))\mathrm{d}t + \sigma(X(t))\mathrm{d}W,$$

where the unknown $X = (X^i)_{i=1,\dots,d}$ lives in the d-dimensional torus $\mathbb{T}^d$. Also, $f = (f^i)_{i=1}^d$ and $\sigma = (\sigma_\ell^i(x))_{i=1,\dots,d,\ell=1,\dots,m}$ are smooth vector fields periodic in $x \in \mathbb{T}^d$. The process $(W^1(t), \dots, W^m(t))$ is an m-dimensional standard Wiener process over a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ endowed with a filtration $(\mathcal{F}_t)_{t\geq 0}$. Using these notations, we rewrite the equation as

$$\mathrm{d}X^i(t) = f^i(X(t))\mathrm{d}t + \sum_{\ell=1}^m \sigma_\ell^i(X(t))\mathrm{d}W^\ell(t), \quad i = 1, \dots, d.$$

Throughout the paper, "smooth functions" means $\mathcal{C}^\infty$ functions. Given a smooth function ψ defined on $\mathbb{T}^d$, we denote $\|\psi\|_{\mathcal{C}^k}$ its norm in $\mathcal{C}^k(\mathbb{T}^d, \mathbb{R})$. We also denote $\|\psi\|_\infty = \sup_{x\in\mathbb{T}^d} |\psi(x)|$. For a multi-index $\boldsymbol{k} = (k_1, \dots, k_d) \in \mathbb{N}^d$, we set $|\boldsymbol{k}| = k_1 + \cdots + k_d$ and

$$\partial_{\boldsymbol{k}}\psi(x) = \frac{\partial^{|\boldsymbol{k}|}\psi(x)}{\partial x_1^{k_1} \cdots \partial x_d^{k_d}}, \quad x \in \mathbb{T}^d.$$

Therefore

$$\|\psi\|_{\mathcal{C}^k} := \sup_{\substack{\boldsymbol{j}=(j_1,\dots,j_d)\\ |\boldsymbol{j}|\leq k}} |\partial_{\boldsymbol{j}}\psi(x)|.$$

We also define the seminorm

$$|\psi|_{\mathcal{C}^k} := \sup_{\substack{\boldsymbol{j}=(j_1,\dots,j_d)\\ 1\leq|\boldsymbol{j}|\leq k}} |\partial_{\boldsymbol{j}}\psi(x)|.$$

In the following, we assume that f and σ are smooth, and since we are working with an SDE on the torus, standard theorems give existence and uniqueness of a solution for any initial data $X(0) = x \in \mathbb{T}^d$ (see, for instance, [18]). We denote this solution by $X(t,x),\ t \geq 0$. Also, since we chose to work on the torus, we do not have any problem of possible unbounded moments, and this solution has clearly all moments finite.

We denote by $L(x,\partial_x)$ the Kolmogorov generator associated with the SDE:

$$L(x,\partial_x)v(x) = f^i(x)\partial_i v(x) + a^{ij}(x)\partial_{ij}v(x), \tag{2.1}$$

where we use the summation convention for repeated indices and $\partial_i = \partial_{x_i}$ and

$$a^{ij}(x) := \frac{1}{2}\sum_{\ell=1}^{m} \sigma_\ell^i(x)\sigma_\ell^j(x).$$

With this SDE is associated the Kolmogorov equation

$$\frac{\mathrm{d}u}{\mathrm{d}t} = L(x,\partial_x)u, \quad x \in \mathbb{T}^d,\ t > 0, \qquad u(0,x) = \varphi(x), \quad x \in \mathbb{T}^d. \tag{2.2}$$

We wish to investigate the approximation properties of the Euler scheme for long times. We need assumptions on the long time behavior of the law of the solutions of (2.2), i.e., of the law of the Markov process. We assume the following mixing properties:

[H1] There exists a $\mathcal{C}^\infty(\mathbb{T}^d,\mathbb{R})$ function $\rho \geq 0$ such that

$$L(x,\partial_x)^*\rho(x) = 0 \quad \text{and} \quad \int_{\mathbb{T}^d} \rho(x)\mathrm{d}x = 1, \tag{2.3}$$

where $L(x,\partial_x)^*$ is the L^2-adjoint operator of $L(x,\partial_x)$. In other words, the measure $\rho(x)\mathrm{d}x$ is invariant by $X(t,x)$.

[H2] Let $g \in \mathcal{C}^\infty(\mathbb{T}^d,\mathbb{R})$, and assume that $\int_{\mathbb{T}^d} g(x)\mathrm{d}x = 0$. Then there exists a unique function $\mu(x) \in \mathcal{C}^\infty(\mathbb{T}^d,\mathbb{R})$ such that

$$L(x,\partial_x)^*\mu(x) = g(x) \quad \text{and} \quad \int_{\mathbb{T}^d} \mu(x)\rho(x)\mathrm{d}x = 0. \tag{2.4}$$

[H3] Let $u(t,x)$ be the solution of (2.2). Assume that $\int_{\mathbb{T}^d} \varphi(x)\rho(x)\mathrm{d}x = 0$; then there exists a constant λ and, for each $k \in \mathbb{N}$, a polynomial $p_k(t)$, such that if $\varphi(x) \in \mathcal{C}^\infty(\mathbb{T}^d,\mathbb{R})$, we have the estimates

$$\forall\, t \geq 0, \quad \|u(t,\cdot)\|_{\mathcal{C}^k} \leq p_k(t)e^{-\lambda t}\|\varphi\|_{\mathcal{C}^k}. \tag{2.5}$$

These hypotheses are usually satisfied under elliptic or hypoelliptic assumptions on the operator $L(x,\partial_x)$. The reader may find in [17] conditions to ensure [H1]. We also refer the reader to [2, 14] for a general definition of hypoellipticity and applications to numerical schemes. Combining kernel estimates for hypoelliptic diffusion [2, 12] and exponential convergence to equilibrium, [H3] can be proved. Note that similar estimates are used in [13, 21, 25, 26], where specific examples are considered. Finally, we mention that these hypotheses can be proved to be fulfilled using PDE techniques (see [10]).

Under these assumptions, (2.2) has a unique smooth solution for a given smooth function φ, and for all $x \in \mathbb{T}^d$ we have (see, for instance, [23, 24])

$$u(t,x) = \mathbb{E}(\varphi(X(t,x))).$$

Moreover, this solution is smooth. In the following, we write $u(t) = P_t\varphi$ so that $(P_t)_{t\geq 0}$ is the transition semigroup associated with the Markov process $(X(t,x))_{t\geq 0,\, x\in\mathbb{T}^d}$. Note that we use the standard identification $u(t) = u(t,\cdot)$.

For a smooth function ψ, we set

$$\langle\psi\rangle = \int_{\mathbb{T}^d} \psi(x)\rho(x)\mathrm{d}x.$$

Note that by [H3], we have for any solution of (2.2) and $k \in \mathbb{N}$

$$\forall\, t \geq 0, \quad \|u(t,\cdot) - \langle\varphi\rangle\|_{\mathcal{C}^k} \leq p_k(t)e^{-\lambda t}\|\varphi - \langle\varphi\rangle\|_{\mathcal{C}^k}. \tag{2.6}$$

Now for a small time step $\tau > 0$ and $x \in \mathbb{T}^d$, we consider the Euler method defined, for $i = 1,\dots,d$, by $X_0 = x$ and the formula

$$X^i_{n+1} = X^i_n + \tau f^i(X_n) + \sigma^i_\ell(X_n)(W^\ell((n+1)\tau) - W^\ell(n\tau)) \tag{2.7}$$

for $n \geq 0$. Our main result can be stated as follows.

THEOREM 2.1. *Let N and $\tau_0 > 0$ be fixed. Then there exist a modified smooth density*

$$\mu^N(x) = \rho(x) + \tau\mu_1(x) + \cdots + \tau^N\mu_N(x)$$

such that $\int_{\mathbb{T}^d}\mu^N(x)\mathrm{d}x = 1$, a constant C_N, and a polynomial $P_N(t)$ such that the following holds: For all smooth functions $\varphi(x)$ on $\mathbb{T}^d$, we have

$$\forall\, p \in \mathbb{N}, \quad \left\|\mathbb{E}\varphi(X_p) - \int_{\mathbb{T}^d}\varphi\mathrm{d}\mu^N\right\|_\infty \leq \left(P_N(t_p)e^{-\lambda t_p} + C_N\tau^{N+1}\right)\|\varphi\|_{\mathcal{C}^{8N+10}}, \tag{2.8}$$

where for all p, $t_p = p\tau$ and $\mathrm{d}\mu^N(x) = \mu^N(x)\mathrm{d}x$.

This result can be viewed as a discrete version of (2.6). Note that it implies that all the invariant measures of the numerical process X_p are close to $\mathrm{d}\mu^N$ up to a very small error term $C_N\tau^N$.

Using this result, we can also recover the weak convergence result

$$\left\|\mathbb{E}\varphi(X_p) - \int_{\mathbb{T}^d}\varphi\mathrm{d}\rho\right\|_\infty \leq C(\varphi)\tau$$

for p large enough, and for some constant $C(\varphi)$ depending on φ, and where we set $\mathrm{d}\rho(x) = \rho(x)\mathrm{d}x$. This can be compared with [13, 14, 21, 25]. As in [14], the only assumption made on τ is that $\tau \leq \tau_0$, where τ_0 is any fixed number. The influence of τ_0 is only reflected in the constants on the right-hand side—we can, for example, take $\tau_0 = 1$. In particular, we do not assume that X_p has a unique invariant measure—something that would be guaranteed only if τ is small enough. We also recover an expansion of the invariant measure as in [27].

In the following sections, the constants appearing in the estimate depend in general on bounds on derivatives of f and g defining the SDE. They will also depend in general on τ_0 and N, but not on φ.

3. Asymptotic expansion of the weak error.

We have the formal expansion for small t:

$$u(t,x) = \varphi(x) + tL(x,\partial_x)\varphi(x) + \frac{t^2}{2}L(x,\partial_x)^2\varphi(x) + \cdots + \frac{t^n}{n!}L(x,\partial_x)^n\varphi(x) + \cdots.$$

This is just obtained by Taylor expansion in time since by (2.2) $\frac{\mathrm{d}^n}{\mathrm{d}t^n}u(t,x) = L(x,\partial_x)^n u(t,x)$ and in particular $\frac{\mathrm{d}^n}{\mathrm{d}t^n}u(0,x) = L(x,\partial_x)^n\varphi(x)$.

Since the solution u of the Kolmogorov equation is smooth and has its derivatives bounded in terms of the initial data φ, the above formal expansion can be justified, and we have the following proposition whose proof is easy and left to the reader.

PROPOSITION 3.1. *Assume that $\varphi \in \mathcal{C}^\infty(\mathbb{T}^d,\mathbb{R})$, and let $\tau_0 > 0$. Then for all N there exists a constant C_N such that for all $\tau < \tau_0$,*

$$\left\| u(\tau,x) - \sum_{n=0}^{N} \frac{\tau^n}{n!} L(x,\partial_x)^n \varphi(x) \right\|_\infty \leq C_N \tau^{N+1} |\varphi|_{\mathcal{C}^{2N+2}}. \tag{3.1}$$

With the Euler scheme defined in (2.7), we associate the continuous process

$$\tilde{X}_x^i(t) = X_n^i + (t-n\tau) f^i(X_n) + \sigma_\ell^i(X_n)(W^\ell(t) - W^\ell(n\tau)), \quad t \in [n\tau,(n+1)\tau], \tag{3.2}$$

and $\tilde{X}(n\tau) = X_n$. We thus have $X_{n+1} = \tilde{X}_x(t_{n+1})$. The process (3.2) satisfies the equation

$$\mathrm{d}\tilde{X}_x^i(t) = f^i(X_n)\mathrm{d}t + \sigma_\ell^i(X_n)\mathrm{d}W^\ell(t), \quad t \in [n\tau,(n+1)\tau]. \tag{3.3}$$

Clearly, (X_n) defines a discrete-in-time homogeneous Markov process but $\tilde{X}_x$ is not Markov.

In this work, we are only interested in the distributions of the solutions and of their approximation. We now examine in detail the first time step and its approximation properties in terms of the law. By the Markov property, it is sufficient to then obtain information at all steps. The next result gives an expansion similar to Proposition 3.1 for the Euler process (see also [28] for a similar analysis).

THEOREM 3.2. *For all $n \geq 1$, there exist operators $A_n(x,\partial_x)$ of order $2n$, such that for all $N \geq 1$, there exists a constant C_N satisfying*

$$\left\| \mathbb{E}\,\varphi(\tilde{X}_x(\tau)) - \sum_{n=0}^{N} \tau^n A_n(x,\partial_x)\varphi(x) \right\|_\infty \leq C_N \tau^{N+1} |\varphi|_{\mathcal{C}^{2N+2}} \tag{3.4}$$

for all $\varphi \in \mathcal{C}^\infty(\mathbb{T}^d,\mathbb{R})$.

Proof. Using (3.3) and the Itô formula, we get for $t \leq \tau$

$$\mathrm{d}\varphi(\tilde{X}_x(t)) = L(x,\partial_x)\varphi(\tilde{X}_x(t))\mathrm{d}t + \sigma_\ell^i(x)\partial_i\varphi(\tilde{X}_x(t))\mathrm{d}W^\ell(t),$$

or, equivalently,

$$\varphi(\tilde{X}_x(t)) = \varphi(x) + \int_0^t L(x,\partial_x)\varphi(\tilde{X}_x(s))\mathrm{d}s + \int_0^t \sigma_\ell^i(x)\partial_i\varphi(\tilde{X}_x(s))\mathrm{d}W^\ell(s). \tag{3.5}$$

Note that the last term is a martingale. We define the operator

$$R_{0,\ell}(x,\partial_x) = \sigma_\ell^i(x)\partial_i.$$

We have for all s and all x

$$L(x,\partial_x)\varphi(\tilde{X}_x(s)) = (f^i(x)\partial_i\varphi)(\tilde{X}_x(s)) + (a^{ij}(x)\partial_{ij}\varphi)(\tilde{X}_x(s)).$$

Hence applying (3.5) to $\partial_i\varphi$ and $\partial_{ij}\varphi$, we obtain

$$\begin{aligned}
L(x,\partial_x)\varphi(\tilde X_x(s)) = L(x,\partial_x)\varphi(x)&\\
&+\int_0^s f^i(x)f^j(x)\partial_{ij}\varphi(\tilde X_x(\theta)) + f^i(x)a^{n\ell}(x)\partial_{n\ell i}\varphi(\tilde X_x(\theta))\mathrm{d}\theta\\
&+\int_0^s a^{ij}(x)f^n(x)\partial_{ijn}\varphi(\tilde X_x(\theta)) + a^{ij}(x)a^{n\ell}(x)\partial_{ijn\ell}\varphi(\tilde X_x(\theta))\mathrm{d}\theta\\
&+\int_0^s f^i(x)\sigma^j_\ell(x)\partial_{ij}\varphi(\tilde X_x(\theta))\mathrm{d}W^\ell(\theta)\\
&+\int_0^s a^{in}(x)\sigma^j_\ell(x)\partial_{inj}\varphi(\tilde X_x(\theta))\mathrm{d}W^\ell(\theta).
\end{aligned}$$

We set $A_0 = I$, $A_1 = L$ and plug this into (3.5) to obtain

$$\begin{aligned}
\varphi(\tilde X_x(t)) = \varphi(x) + tA_1(x,\partial_x)\varphi(x) + \int_0^t\int_0^s A_2(x,\partial_x)\varphi(\tilde X_x(\theta))\mathrm{d}\theta\mathrm{d}s&\\
+\int_0^t\int_0^s R_{1,\ell}(x,\partial_x)\varphi(\tilde X_x(\theta))\mathrm{d}W^\ell(\theta)\mathrm{d}s + \int_0^t R_{0,\ell}(x,\partial_x)\varphi(\tilde X_x(\theta))\mathrm{d}W^\ell(\theta),&
\end{aligned}$$

where

$$A_2(x,\partial_x) = f^i(x)f^j(x)\partial_{ij} + f^i(x)a^{n\ell}(x)\partial_{n\ell i} + a^{ij}(x)f^n(x)\partial_{ijn} + a^{ij}(x)a^{n\ell}(x)\partial_{ijn\ell}$$

is an operator of order 4, and

$$R_{1,\ell}(x,\partial_x) = f^i(x)\sigma^j_\ell(x)\partial_{ij} + a^{in}(x)\sigma^j_\ell(x)\partial_{inj}$$

are operators of order 3. Taking the expectation so that the last two terms disappear, we easily deduce the result for $N=1$.

Let us now prove recursively that there exist operators $A_n(x,\partial_x)$ of order $2n$ and $R_{n,\ell}(x,\partial_x)$ of order $2n+1$, such that

(3.6)

$$\begin{aligned}
\varphi(\tilde X_x(t)) = \varphi(x) + tL(x,\partial_x)\varphi(x) + \sum_{n=2}^N t^n A_n(x,\partial_x)\varphi(x)&\\
+\int_0^t\cdots\int_0^{s_N} A_{N+1}(x,\partial_x)\varphi(\tilde X_x(s_{N+1}))\mathrm{d}s_1\cdots\mathrm{d}s_{N+1}&\\
+\sum_{n=1}^N\int_0^t\cdots\int_0^{s_n} R_{n,\ell}(x,\partial_x)\varphi(\tilde X_x(s_{n+1}))\mathrm{d}s_1\cdots\mathrm{d}s_n\mathrm{d}W^\ell(s_{n+1}).&
\end{aligned}$$

Note that the expectation of the last term vanishes so that (3.6) easily implies (3.4).

To prove (3.6), assume that A_{N+1} and $R_{N,\ell}$ are known, and let us decompose A_{N+1} as $A_{N+1}(x,\partial_x) = A^{\boldsymbol{j}}_{N+1}(x)\partial_{\boldsymbol{j}}$, where $\boldsymbol{j} = (j_1,\dots,j_m)$ are multi-indices (with the summation convention) and $A^{\boldsymbol{j}}_{N+1}$ smooth functions of x. Such a decomposition is easy to write for $N=1$ or 2. We apply (3.5) to $\partial_{\boldsymbol{j}}\varphi(\tilde X_x(s_{N+1}))$ for a given multi-index $\boldsymbol{j}$, and obtain

$$\begin{aligned}A^{\boldsymbol{j}}_{N+1}(x)\partial_{\boldsymbol{j}}\varphi(\tilde{X}_x(s_{N+1})) = {} & A^{\boldsymbol{j}}_{N+1}(x)\partial_{\boldsymbol{j}}\varphi(x)\\ & + \int_0^{s_{N+1}} A^{\boldsymbol{j}}_{N+1}(x)f^n(x)\partial_n\partial_{\boldsymbol{j}}\varphi(\tilde{X}_x(s_{N+2}))\mathrm{d}s_{N+2}\\ & + \int_0^{s_{N+1}} A^{\boldsymbol{j}}_{N+1}(x)a^{n\ell}(x)\partial_{n\ell}\partial_{\boldsymbol{j}}\varphi(\tilde{X}_x(s_{N+2}))\mathrm{d}s_{N+2}\\ & + \int_0^{s_{N+1}} A^{\boldsymbol{j}}_{N+1}(x)\sigma^n_\ell(x)\partial_n\partial_{\boldsymbol{j}}\varphi(\tilde{X}_x(s_{N+2}))\mathrm{d}W^\ell(s_{N+2}).\end{aligned}$$

We thus choose

$$A_{N+2}(x) = A^{\boldsymbol{j}}_{N+1}(x)f^n(x)\partial_n\partial_{\boldsymbol{j}} + A^{\boldsymbol{j}}_{N+1}(x)a^{ij}(x)\partial_{ij}\partial_{\boldsymbol{j}}$$

and

$$R_{N+1,\ell}(x) = A^{\boldsymbol{j}}_{N+1}(x)\sigma^n_\ell(x)\partial_n\partial_{\boldsymbol{j}},$$

and obtain (3.6) with $N+1$ replaced by $N+2$. ∎

4. Modified generator.

4.1. Formal series analysis. Let us now consider τ as fixed. We want to construct a formal series

$$L(\tau;x,\partial_x) = L(x,\partial_x) + \tau L_1(x,\partial_x) + \cdots \tau^n L_n(x,\partial_x) + \cdots \tag{4.1}$$

with operator coefficients $L_n(x,\partial_x)$ smooth on $\mathbb{T}^d$, and such that formally the solution $v(t,x)$ at time $t=\tau$ of the equation

$$\partial_t v(t,x) = L(\tau;x,\partial_x)v(t,x), \quad v(0,x) = \varphi(x)$$

coincides in the sense of asymptotic expansion with the approximation of the transition semigroup $\mathbb{E}(\varphi(\tilde{X}_x(\tau)))$ studied in the previous section. In other words, we want to have the equality in the sense of asymptotic expansion in powers of τ

$$\exp(\tau L(\tau;x,\partial_x))\varphi(x) = \varphi(x) + \sum_{n\geq 1}\tau^n A_n(x,\partial_x)\varphi(x),$$

where the operators $A_n(x,\partial_x)$ are defined in Theorem 3.2.

Formally, this equation can be written

$$\exp(\tau L(\tau;x,\partial_x)) - \mathrm{Id} = \tau\widetilde{A}(\tau), \tag{4.2}$$

where $\widetilde{A}(\tau) = \sum_{n\geq 1}\tau^{n-1}A_n$.

We have

$$\exp(\tau L(\tau;x,\partial_x)) - \mathrm{Id} = \tau L(\tau;x,\partial_x)\left(\sum_{n\geq 0}\frac{\tau^n}{(n+1)!}L(\tau;x,\partial_x)^n\right).$$

Note that the (formal) inverse of the series is given by

$$\left(\sum_{n\geq 0}\frac{\tau^n}{(n+1)!}L(\tau;x,\partial_x)^n\right)^{-1} = \sum_{n\geq 0}\frac{B_n}{n!}\tau^n L(\tau;x,\partial_x)^n,$$

where the B_n are the Bernoulli numbers; see, for instance, [5, 6, 9] for a similar analysis involving operators. Hence (4.1), (4.2) are equivalent in the sense of formal series to

(4.3)
$$\begin{aligned} L(\tau;x,\partial_x) &= \sum_{\ell\geq 0}\frac{B_\ell}{\ell!}\tau^\ell L(\tau;x,\partial_x)^\ell \widetilde{A}(\tau) \\ &= \sum_{n\geq 0}\tau^n\Big(A_{n+1}+\sum_{\ell=1}^{n}\frac{B_\ell}{\ell!}\sum_{n_1+\cdots+n_\ell+n_{\ell+1}=n-\ell}L_{n_1}\cdots L_{n_\ell}A_{n_{\ell+1}+1}\Big). \end{aligned}$$

By identifying the right-hand sides of (4.1) and (4.3), we get the following recursion formula:

$$L_n = A_{n+1}+\sum_{\ell=1}^{n}\frac{B_\ell}{\ell!}\sum_{n_1+\cdots+n_\ell+n_{\ell+1}=n-\ell}L_{n_1}\cdots L_{n_\ell}A_{n_{\ell+1}+1}. \tag{4.4}$$

Each of the terms of the above sum is an operator of order $2n+2$ with smooth coefficients, and therefore L_n is also an operator of order $2n+2$ with smooth coefficients.

Note that (4.2) immediately gives the inverse relation of this formal series equation:

$$A_n = \sum_{\ell=1}^{n}\frac{1}{\ell!}\sum_{n_1+\cdots+n_\ell=n-\ell}L_{n_1}\cdots L_{n_\ell}. \tag{4.5}$$

Moreover, we clearly have

$$L_n(x,\partial_x)\mathbb{1}=0,$$

where $\mathbb{1}$ denotes the constant function equal to 1.

4.2. Approximate solution of the modified flow. For a given N, we have constructed in the previous section an operator

$$L^{(N)}(\tau;x,\partial_x) = L(x,\partial_x)+\sum_{n=1}^{N}\tau^n L_n(x,\partial_x). \tag{4.6}$$

In order to perform weak backward error analysis and estimate recursively the modified invariant law of the numerical process, we should be able to define a solution $v^N(t,x)$ of the modified flow

$$\partial_t v^N(t,x) = L^{(N)}(\tau;x,\partial_x)v^N(t,x), \quad v^N(0,x)=\varphi(x). \tag{4.7}$$

However, in our situation we do not know whether this equation has a solution in general. This is in contrast with standard backward error analysis where the modified flow can always be defined.

The goal of the next proposition is to give a proper definition of the modified equation (4.7).

THEOREM 4.1. *Let φ be a smooth function on $\mathbb{T}^d$. For all $n\in\mathbb{N}$, there exist smooth functions $v_\ell(t,x)$, defined for all times $t\geq 0$, and such that for all $t\geq 0$ and $n\in\mathbb{N}$,*

$$\partial_t v_n(t,x) - Lv_n(t,x) = \sum_{\ell=1}^{n}L_\ell v_{n-\ell}(t,x), \tag{4.8}$$

with initial conditions $v_0(0,x) = \varphi(x)$ *and* $v_n(0,x) = 0$ *for* $n \geq 1$. *For all* $N \geq 0$, *setting*

$$v^{(N)}(t,x) = \sum_{k=0}^{N} \tau^k v_k(t,x), \tag{4.9}$$

then there exists a constant C_N *such that for all times* $t \geq 0$ *and all* $\tau \geq 0$,

$$\|\mathbb{E}v^{(N)}(t,\tilde{X}_x(\tau)) - v^{(N)}(t+\tau,x)\|_\infty \leq C_N \tau^{N+1} \sup_{\substack{s\in(0,\tau)\\ n=0,\dots,N}} |v_n(t+s)|_{\mathcal{C}^{4N+2}}. \tag{4.10}$$

Proof. For $n = 0$, (4.8) implies $v_0(t,x) = u(t,x)$, the solution of (2.2). Let $n \in \mathbb{N}$ and assume that $v_j(t,x)$ are constructed for $j = 1,\dots,n-1$. Let

$$F_n(t,x) := \sum_{\ell=1}^{n} L_\ell v_{n-\ell}(t,x) \tag{4.11}$$

be the right-hand side in (4.8). Then v_n is uniquely defined and given by the formula

$$v_n(t) = \int_0^t P_{t-s} F_n(s)\mathrm{d}s, \quad t \geq 0. \tag{4.12}$$

By [H3], it is not difficult to check that v_n, $n \in \mathbb{N}$, are smooth and that for all $t \geq 0$

$$\|v_n(t)\|_{\mathcal{C}^k} \leq C\|\varphi\|_{\mathcal{C}^{k+4n}}, \quad k \in \mathbb{N}, \tag{4.13}$$

where the constant C depends on k, n and on the coefficient of the equation. This proves the first part of the theorem.

To prove (4.10), we consider a fixed time t and define the functions $w_n(x,s) := v_n(t+s)$ for $s \geq 0$ and $n \in \mathbb{N}$. By definition, these functions satisfy the relation

$$\partial_s w_n(s,x) - Lw_n(s,x) = \sum_{\ell=1}^{n} L_\ell w_{n-\ell}(s,x), \quad w_n(0,x) = v_n(t,x).$$

Let us consider the successive time derivatives of the functions $w_n(s,x)$. We have, using the definition of w_n, for all $s \geq 0$,

$$\partial_s^2 w_n(s,x) = \sum_{\ell=0}^{n} L_\ell \partial_s w_{n-\ell}(s,x) = \sum_{k=0}^{n} \sum_{\ell_1+\ell_2=k} L_{\ell_1} L_{\ell_2} w_{n-k}(s,x),$$

and we see by induction that for all $m \geq 1$ and $s \geq 0$

$$\partial_s^m w_n(s,x) = \sum_{\ell_1+\dots+\ell_{m+1}=n} L_{\ell_1}\cdots L_{\ell_m} w_{\ell_{m+1}}(s,x). \tag{4.14}$$

Using the fact that the operators L_ℓ are of order $2\ell+2$ with no terms of order zero, we see that there exists a constant C depending on n and m, such that

$$\|\partial_s^m w_n(s,x)\|_\infty \leq C \sup_{k=0,\dots,n} |w_k(s)|_{\mathcal{C}^{2k+2m}}.$$

Now let us consider the Taylor expansion of $w_n(\tau)$ for $\tau \leq \tau_0$ and $n = 0, \dots, N$,

$$\begin{aligned} w_n(\tau, x) &= \sum_{m=0}^{N-n} \frac{\tau^m}{m!} \partial_t^m w_n(0,x) + \int_0^\tau \frac{s^{N-n}}{(N-n)!} \partial_t^{N-n+1} w_n(s,x) \mathrm{d}s \\ &= \sum_{m=0}^{N-n} \frac{\tau^m}{m!} \sum_{\ell_1 + \dots + \ell_{m+1} = n} L_{\ell_1} \cdots L_{\ell_m} w_{\ell_{m+1}}(0,x) + R_{N,n}(\tau,x). \end{aligned}$$

Using the bounds on the time derivatives of $w_n(s,x)$, we obtain that for all $\tau \geq 0$ and all $n = 0, \dots, N$

$$\|R_{N,n}(\tau,x)\|_\infty \leq C \tau^{N-n+1} \sup_{\substack{s \in (0,\tau) \\ k=0,\dots,N}} |w_k(s,x)|_{\mathcal{C}^{4N+2}}$$

for some constant depending on N, m. After summation in n, we get

$$v^{(N)}(t+\tau, x) = \sum_{n=0}^{N} \sum_{m=0}^{N-n} \frac{\tau^{m+n}}{m!} \sum_{\ell_1 + \dots + \ell_{m+1} = n} L_{\ell_1} \cdots L_{\ell_m} w_{\ell_{m+1}}(0,x) + R_N(t,\tau,x),$$

where

$$\|R_N(t,\tau,x)\|_\infty \leq C_N \tau^{N+1} \sup_{\substack{s \in (0,\tau) \\ n=0,\dots,N}} |v_n(t+s,x)|_{\mathcal{C}^{4N+2}}.$$

But we have

$$\begin{aligned} &\sum_{n=0}^{N} \sum_{m=0}^{N-n} \frac{\tau^{m+n}}{m!} \sum_{\ell_1 + \dots + \ell_{m+1} = n} L_{\ell_1} \cdots L_{\ell_m} w_{\ell_{m+1}}(0,x) \\ &\qquad = \sum_{p=0}^{N} \tau^p \sum_{m=0}^{p} \frac{1}{m!} \sum_{\ell_1 + \dots + \ell_m = p - \ell_{m+1} - m} L_{\ell_1} \cdots L_{\ell_m} w_{\ell_{m+1}}(0,x) \\ &\qquad = \sum_{p=0}^{N} \tau^p \sum_{q=0}^{p} \left(\sum_{m=0}^{p-q} \frac{1}{m!} \sum_{\ell_1 + \dots + \ell_m = p-q-m} L_{\ell_1} \cdots L_{\ell_m} \right) w_q(0,x) \\ &\qquad = \sum_{p=0}^{N} \tau^p \sum_{q=0}^{p} A_{p-q} v_q(t,x), \end{aligned}$$

using the expression (4.5) defining the operators A_n and the definition of w_n.

To conclude, we use (3.4) applied to $\varphi = v^{(N)}(t,x)$, and we easily verify that $\mathbb{E} v^{(N)}(\tau, \tilde{X}_x(\tau))$ satisfy the same asymptotic expansion. ∎

Note that in the previous theorem, we have constructed a function $v^{(N)}(t,x)$ which is an approximate solution of (4.7). More precisely, we can easily show that we have for all time $t \geq 0$

$$\partial_t v^{(N)}(t,x) = L^{(N)}(\tau; x, \partial_x) v^{(N)}(t,x) + R^{(N)}(t,x), \quad v^{(N)}(0,x) = \varphi(x),$$

where

$$R^{(N)}(t,x) = - \sum_{\substack{\ell_1, \ell_2 = 0, \dots, N \\ \ell_1 + \ell_2 > N}} \tau^{\ell_1 + \ell_2} L_{\ell_1} v_{\ell_2}(t,x)$$

is of order $\mathcal{O}(\tau^{N+1})$.

5. Asymptotic expansion of the invariant measure and long time behavior. We now analyze the long time behavior of the solution of the modified equation (4.7). In the following, for a given operator $B(x,\partial_x)$, we denote by $B(x,\partial_x)^*$ its adjoint with respect to the L^2 product. We start by an asymptotic expansion of a formal invariant measure for the numerical scheme.

PROPOSITION 5.1. *Let $(L_n)_{n\geq 0}$ be the collection of operators defined recursively by (4.4). There exists a collection of functions $(\mu_n(x))_{n\geq 0}$ such that $\mu_0(x)=\rho(x)$, $\int_{\mathbb{T}^d}\mu_n(x)\mathrm{d}x=0$ for $n\geq 1$, and for all $n\geq 0$*

$$L_0^*\mu_n = -\sum_{\ell=1}^{n}(L_\ell)^*\mu_{n-\ell}. \tag{5.1}$$

Let $N\geq 0$, let τ_0 be fixed, and for $\tau\in[0,\tau_0]$ let $L^{(N)}(\tau;x,\partial_x)$ be the operator defined by (4.6). Then for $\tau\in[0,\tau_0]$ the function

$$\mu^{(N)}(\tau;x)=\rho(x)+\sum_{n=1}^{N}\tau^n\mu_n(x)\in\mathcal{C}^\infty(\mathbb{T}^d,\mathbb{R})$$

satisfies

$$\int_{\mathbb{T}^d}\mu^{(N)}(\tau;x)\mathrm{d}x=1$$

and

$$L^{(N)}(\tau;x,\partial_x)^*\mu^{(N)}(\tau;x)=G^{(N)}(\tau;x),$$

with, for all k and all $\tau\in[0,\tau_0]$,

$$\left\|G^{(N)}(\tau;x)\right\|_{\mathcal{C}^k}\leq C_{N,k}\tau^{N+1}\quad \textit{and}\quad \int_{\mathbb{T}^d}G^{(N)}(\tau;x)\mathrm{d}x=0,$$

where $C_{N,k}$ depends on N, k, and τ_0.

Proof. Assume that $\mu_0=\rho$ and $\mu_j(x)$ are known for $j=0,\dots,n-1$ with $n\geq 1$. Consider (5.1) given by

$$L_0^*\mu_n=-\sum_{\ell=1}^{n}(L_\ell)^*\mu_{n-\ell}=G_n.$$

Note that the right-hand side $G_n(x)$ is a smooth function satisfying

$$\int_{\mathbb{T}^d}G_n(x)\mathrm{d}x=-\sum_{\ell=1}^{n}\int_{\mathbb{T}^d}(L_\ell)^*\mu_{n-\ell}\mathrm{d}x=-\sum_{\ell=1}^{n}\int_{\mathbb{T}^d}\mu_{n-\ell}L_\ell\mathbb{1}\mathrm{d}x=0,$$

where $\mathbb{1}$ denotes the constant function equal to 1, which as already seen is in the Kernel of all the L_ℓ.

Using hypothesis [H2], we easily obtain the existence of a $\mathcal{C}^\infty$ function $\tilde{\mu}_n$ satisfying the equation $L_0^*\tilde{\mu}_n=G_n$ and $\int_{\mathbb{T}^d}\mu_n(x)\rho(x)\mathrm{d}x=0$. As $L_0^*\rho=0$, the function $\mu_n(x)=\tilde{\mu}_n(x)-\rho(x)(\int_{\mathbb{T}^d}\tilde{\mu}_n(y)\mathrm{d}y)$ satisfies (5.1) and the condition $\int_{\mathbb{T}^d}\mu_n(x)\mathrm{d}x=0$. This shows the first part of the proposition.

We then write

$$\begin{aligned} L^{(N)}(\tau; x, \partial_x)^* \mu^{(N)}(\tau; x) &= \sum_{n=0}^{2N} \tau^n \sum_{\substack{\ell_1+\ell_2=n \\ \ell_i \leq N}} (L_{\ell_1})^* \mu_{\ell_2} \\ &= \sum_{n=N+1}^{2N} \tau^n \sum_{\substack{\ell_1+\ell_2=n \\ \ell_i \leq N}} (L_{\ell_1})^* \mu_{\ell_2} \\ &=: G^{(N)}(\tau; x), \end{aligned}$$

and we easily verify that $G^{(N)}$ satisfies the hypothesis of the proposition, owing to the fact that $L_\ell \mathbb{1} = 0$ for all $\ell \geq 0$. □

PROPOSITION 5.2. *For all n and k there exists a polynomial $P_{k,n}(t)$ such that for all $t \geq 0$*

$$\left\| v_n(t,x) - \int_{\mathbb{T}^d} \varphi \mathrm{d}\mu_n \right\|_{\mathcal{C}^k} \leq P_{k,n}(t) e^{-\lambda t} \|\varphi - \langle \varphi \rangle\|_{\mathcal{C}^{k+4n}} . \tag{5.2}$$

Proof. Using the fact that $\mu_0 = \rho$ and $v_0 = u$, we see that estimate (5.2) is satisfied for $n = 0$ (see (2.6)). Let $n \geq 1$ and assume that v_j, $j = 0, \ldots, n-1$, satisfy for $k \in \mathbb{N}$, $t \geq 0$

$$\left\| v_j(t,x) - \int_{\mathbb{T}^d} \varphi(x) \mu_j(x) \mathrm{d}x \right\|_{\mathcal{C}^k} \leq P_{k,j}(t) e^{-t\lambda} \|\varphi(x) - \langle \varphi \rangle\|_{\mathcal{C}^{k+4j}}$$

for some polynomial $P_{k,j}$.

Let us set

$$c_n(t) = \sum_{m=0}^{n} \int_{\mathbb{T}^d} v_{n-m}(t,x) \mu_m(x) \mathrm{d}x.$$

We claim that $c_n(t)$ does not depend on time. Indeed,

$$\begin{aligned} \sum_{m=0}^{n} \partial_t \int_{\mathbb{T}^d} v_{n-m}(t,x) \mu_m(x) \mathrm{d}x &= \sum_{m=0}^{n} \partial_t \int_{\mathbb{T}^d} v_m(t,x) \mu_{n-m}(x) \mathrm{d}x \\ &= \sum_{m=0}^{n} \sum_{\ell=0}^{m} \int_{\mathbb{T}^d} L_{m-\ell} v_\ell(t,x) \mu_{n-m}(x) \mathrm{d}x \\ &= \sum_{\ell=0}^{n} \sum_{m=\ell}^{n} \int_{\mathbb{T}^d} v_\ell(t,x) L_{m-\ell}^* \mu_{n-m}(x) \mathrm{d}x \\ &= \sum_{\ell=0}^{n} \int_{\mathbb{T}^d} v_\ell(t,x) \sum_{m=0}^{n-\ell} L_m^* \mu_{n-\ell-m}(x) \mathrm{d}x = 0, \end{aligned}$$

by definition of the coefficients μ_n; see (5.1). Note that, thanks to the smoothness properties of all the functions, the computation above is easily justified.

We deduce

$$\int_{\mathbb{T}^d} \partial_t v_n(t,x) \rho(x) \mathrm{d}x = -\sum_{m=1}^{n} \int_{\mathbb{T}^d} \partial_t v_{n-m}(t,x) \mu_m(x) \mathrm{d}x. \tag{5.3}$$

Next, we compute the average of F_n. By (4.8), (4.11), and (5.3), we have

$$\begin{aligned}\langle F_n(t)\rangle = \int_{\mathbb{T}^d} F_n(t,x)\rho(x)\mathrm{d}x &= \int_{\mathbb{T}^d} \partial_t v_n(t,x)\rho(x)\mathrm{d}x - \int_{\mathbb{T}^d} L v_n(t,x)\rho(x)\mathrm{d}x \\ &= \int_{\mathbb{T}^d} \partial_t v_n(t,x)\rho(x)\mathrm{d}x \\ &= -\sum_{m=1}^{n}\int_{\mathbb{T}^d} \partial_t v_{n-m}(t,x)\mu_m(x)\mathrm{d}x.\end{aligned}$$

We rewrite (4.12) as follows:

$$v_n(t,x) = \int_0^t \langle F_n(s)\rangle \mathrm{d}s + \int_0^t P_{t-s}(F_n(s,x) - \langle F_n(s)\rangle)\mathrm{d}s.$$

Using the previous expression obtained for $\langle F_n(s)\rangle$ and recalling the initial data for v_n, we deduce that

$$\begin{aligned}v_n(t,x) = -\sum_{m=1}^{n}\int_{\mathbb{T}^d} v_{n-m}(t,x)\mu_m(x)\mathrm{d}x + \int_{\mathbb{T}^d}\varphi(x)\mu_n(x)\mathrm{d}x \\ + \int_0^t P_{t-s}(F_n(s,x) - \langle F_n(s)\rangle)\mathrm{d}s.\end{aligned}$$

Then, using $\int_{\mathbb{T}^d}\mu_m(x)\mathrm{d}x = 0$, $m\in\mathbb{N}$, we get

$$\begin{aligned}&v_n(t,x) - \int_{\mathbb{T}^d}\varphi(x)\mu_n(x)\mathrm{d}x \\ &\quad = \sum_{m=1}^{n}\int_{\mathbb{T}^d}\left(v_{n-m}(t,x) - \int_{\mathbb{T}^d}\varphi(x)\mu_{n-m}(x)\mathrm{d}x\right)\mu_m(x)\mathrm{d}x \\ &\qquad + \int_0^t P_{t-s}(F_n(s,x) - \langle F_n(s)\rangle)\mathrm{d}s.\end{aligned}$$

Note that, since L_ℓ, $\ell\in\mathbb{N}$, is a differential operator of order $2\ell+2$ with smooth coefficients and contains no zero order terms, we have

$$\|F_n(s) - \langle F_n(s)\rangle\|_{\mathcal{C}^k} \le c_{k,\ell}\sum_{\ell=0}^{n-1}\left\|v_\ell(s) - \int_{\mathbb{T}^d} v_\ell(s,x)\rho(x)dx\right\|_{\mathcal{C}^{k+2(n-\ell)+2}}.$$

Then by [H3]

$$\begin{aligned}&\left\|v_n(t,x) - \int_{\mathbb{T}^d}\varphi(x)\mu_n(x)\mathrm{d}x\right\|_{\mathcal{C}^k} \\ &\quad \le \sum_{m=1}^{n}\left\|v_{n-m}(t,x) - \int_{\mathbb{T}^d}\varphi(x)\mu_{n-m}(x)\mathrm{d}x\right\|_\infty \int_{\mathbb{T}^d}|\mu_m(x)|\mathrm{d}x \\ &\qquad + \int_0^t p_k(t-s)e^{-\lambda(t-s)}\|F_n(s,x) - \langle F_n(s)\rangle\|_{\mathcal{C}^k}\,\mathrm{d}s,\end{aligned}$$

and using the recursion assumption

$$\left\| v_n(t,x) - \int_{\mathbb{T}^d} \varphi(x)\mu_n(x)\mathrm{d}x \right\|_{\mathcal{C}^k} \leq \sum_{m=1}^{n} c_{n,m} P_{0,n-m}(t) e^{-t\lambda} \|\varphi(x) - \langle\varphi\rangle\|_{\infty}$$
$$+ \sum_{\ell=0}^{n-1} \int_0^t p_k(t-s) P_{k+2(n-\ell)+2,\ell}(s) e^{-\lambda t} \mathrm{d}s \|\varphi(x) - \langle\varphi\rangle\|_{k+4n}.$$

The conclusion follows. ∎

We now give our main result concerning the long time behavior of the numerical solution. It implies Theorem 2.1 and estimate (1.3) mentioned in the introduction.

THEOREM 5.3. *Let τ_0 and N be fixed. Then there exist C_N and a polynomial $P_N(t)$ such that the following holds: Let X_p be the discrete process defined by (2.7); then we have for $p \geq 0$, $\tau \leq \tau_0$ and smooth function $\varphi(x)$*

$$\forall\, p \in \mathbb{N}, \quad \|\mathbb{E}\,\varphi(X_p) - v^{(N)}(t_p,x)\|_{\infty} \leq C_N \tau^{N+1} \|\varphi\|_{\mathcal{C}^{8N+10}}, \tag{5.4}$$

where for all p, $t_p = p\tau$. Moreover, we have

$$\forall\, p \in \mathbb{N}, \quad \left\| \mathbb{E}\varphi(X_p) - \int_{\mathbb{T}^d} \varphi \mathrm{d}\mu^{(N)} \right\|_{\infty} \leq \left(P_N(t_p) e^{-\lambda t_p} + C_N \tau^{N+1} \right) \|\varphi\|_{\mathcal{C}^{8N+10}}, \tag{5.5}$$

where $\mathrm{d}\mu^{(N)}(x) = \mu^{(N)}(x)\mathrm{d}x$.

Proof. For all p, with $t_j = j\tau$, we have

$$\begin{aligned}
\mathbb{E}\,\varphi(X_p) - v^{(N+1)}(t_p,x) &= \mathbb{E}\, v^{(N+1)}(0,X_p) - v^{(N+1)}(t_p,x) \\
&= \mathbb{E} \sum_{j=0}^{p-1} \mathbb{E}^{X_{p-j-1}} \Big(v^{(N+1)}(t_j, X_{p-j}) - v^{(N+1)}(t_{j+1}, X_{p-j-1}) \Big).
\end{aligned}$$

Here we have used the notation $\mathbb{E}^{X_{p-j-1}}$ for the conditional expectation with respect to the filtration generated by X_{p-j-1}. By the Markov property of the Euler process at times t_j,

$$\begin{aligned}
&\mathbb{E}^{X_{p-j-1}} \Big(v^{(N+1)}(t_j, X_{p-j}) - v^{(N+1)}(t_{j+1}, X_{p-j-1}) \Big) \\
&\qquad = \mathbb{E}^{X_{p-j-1}} \Big(v^{(N+1)}(t_j, \tilde{X}_{X_{p-j-1}}(\tau)) - v^{(N+1)}(t_{j+1}, X_{p-j-1}) \Big).
\end{aligned}$$

Using (4.10) with $t = t_j$, and Proposition 5.2, we deduce that

$$\begin{aligned}
\|\mathbb{E}\,\varphi(X_p) - v^{(N+1)}(t_p,x)\|_{\infty} &\leq C_N \tau^{N+2} \sum_{j=0}^{p-1} \sup_{\substack{s\in(0,\tau) \\ =0,\dots,N+1}} |v_n(t_j+s,x)|_{4N+6} \\
&\leq C_N \tau^{N+2} \|\varphi\|_{\mathcal{C}^{8N+10}} \sum_{j=0}^{p-1} Q_N(t_j) e^{-\lambda t_j}
\end{aligned}$$

for some constant C_N and polynomial $Q_N(t)$. We have used $|v_n(t_j+s,x)|_{4N+6} = |v_n(t_j+s,x) - \int_{\mathbb{T}^d} \varphi d\mu_n|_{4N+6}$. We conclude by using the fact that for a fixed constant $\gamma > 0$, we have

$$\sum_{j=0}^{p-1} e^{-\gamma j \tau} \leq \frac{1}{1-e^{-\gamma\tau}} \leq \frac{C}{\tau},$$

where the constant C depends on γ and τ_0. This shows

$$\left\|\mathbb{E}\,\varphi(X_p) - v^{(N+1)}(t_p, x)\right\|_{\infty} \le C_N \tau^{N+1} \|\varphi\|_{\mathcal{C}^{8N+10}}.$$

To prove (5.4), we note that

$$v^{(N+1)}(t_p, x) = v^{(N)}(t_p, x) + \tau^{N+1} v_n(t_p, x)$$

with, for all p,

$$\left\|v_n(t_p, x)\right\|_{\infty} \le C_N \|\varphi\|_{\mathcal{C}^{4N+4}}$$

upon using (5.2).

The second estimate is then a consequence of Proposition 5.2 and the definition of $v^{(N)}$. ∎

REFERENCES

[1] A. Abdulle, D. Cohen, G. Vilmart, and K. Zygalakis, *High weak order methods for stochastic differential equations based on modified equations*, SIAM J. Sci. Comput., to appear.

[2] V. Bally and D. Talay, *The law of the Euler scheme for stochastic differential equations* (I): *Convergence rate of the distribution function*, Probab. Theory Related Fields, 104 (1995), pp. 43–60.

[3] V. Bally and D. Talay, *The law of the Euler scheme for stochastic differential equations* (II): *Convergence rate of the density*, Monte Carlo Methods Appl., 2 (1996), pp. 93–128.

[4] G. Benettin and A. Giorgilli, *On the Hamiltonian interpolation of near to the identity symplectic mappings with application to symplectic integration algorithms*, J. Statist. Phys., 74 (1994), pp. 1117–1143.

[5] A. Debussche and E. Faou, *Modified energy for split-step methods applied to the linear Schrödinger equation*, SIAM J. Numer. Anal., 47 (2009), pp. 3705–3719.

[6] E. Faou, *Geometric numerical integration and Schrödinger equations*, European. Math. Soc., to appear.

[7] E. Faou and B. Grébert, *Hamiltonian interpolation of splitting approximations for nonlinear PDEs*, Found. Comput. Math.,11 (2011), pp. 381–415.

[8] E. Hairer and C. Lubich, *The life-span of backward error analysis for numerical integrators*, Numer. Math., 76 (1997), pp. 441–462.

[9] E. Hairer, C. Lubich, and G. Wanner, *Geometric Numerical Integration. Structure-Preserving Algorithms for Ordinary Differential Equations*, 2nd ed., Springer-Verlag, Berlin, 2006.

[10] L. Hörmander, *The Analysis of Linear Partial Differential Operators.* III, Classics in Mathematics, Springer-Verlag,, Berlin, 2007.

[11] P. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, Appl. Math. (N.Y.) 23, Springer-Verlag, Berlin, 1992.

[12] S. Kusuoka and D. Stroock, *Applications of the Malliavin Calculus.* II, J. Fac. Sci. Univ. Tokyo Sect. IA Math., 32 (1985), pp 1–76.

[13] J. C. Mattingly, A. M. Stuart, and D. J. Higham, *Ergodicity for SDEs and approximations: Locally Lipschitz vector fields and degenerate noise*, Stochastic Process. Appl., 101 (2002), pp. 185–232.

[14] J. C. Mattingly, A. M. Stuart, and M. V. Tretyakov, *Convergence of numerical time-averaging and stationary measures via Poisson equations*, SIAM J. Numer. Anal., 48 (2010), pp. 552–577.

[15] G. Milstein and M. Tretyakov, *Stochastic Numerics for Mathematical Physics*, Springer-Verlag, Berlin, Heidelberg, New York, 2004.

[16] J. Moser, *Lectures on Hamiltonian systems*, Mem. Amer. Math. Soc., 81 (1968), pp. 1–60.

[17] D. Nualart, *Malliavin Calculus and Related Topics*, 2nd ed., Springer-Verlag, Berlin, 2006.

[18] B. Øksendal, *Stochastic Differential Equations: An Introduction with Applications*, 6th ed., Springer-Verlag, Berlin, 2003.

[19] S. Reich, *Backward error analysis for numerical integrators*, SIAM J. Numer. Anal., 36 (1999), pp. 1549–1570.

[20] B. Leimkuhler and S. Reich, *Simulating Hamiltonian Dynamics*, Cambridge Monogr. Appl. Comput. Math. 14, Cambridge University Press, Cambridge, UK, 2004.

[21] D. Talay, *Second order discretization schemes of stochastic differential systems for the computation of the invariant law*, Stoch. Stoch. Rep., 29 (1990), pp. 13–36.

[22] T. Shardlow, *Modified equations for stochastic differential equations*, BIT, 46 (2006), pp. 111–125.

[23] D. W. Stroock, *Partial Differential Equations for Probabilists*, Cambridge Stud. Adv. Math. 112, Cambridge University Press, Cambridge, UK, 2008.

[24] D. W. Stroock and S. R. S. Varadhan, *Multidimensional Diffusion Processes*, Springer-Verlag, Berlin, 1979.

[25] D. Talay. *Probabilistic numerical methods for partial differential equations: Elements of analysis*, in Probabilistic Models for Nonlinear Partial Differential Equations, D. Talay and L. Tubaro, eds., Lecture Notes in Math. 1627, Springer-Verlag, Berlin, 1996, pp. 48–196.

[26] D. Talay, *Stochastic Hamiltonian dissipative systems: Exponential convergence to the invariant measure, and discretization by the implicit Euler scheme*, Markov Process. Related Fields, 8 (2002), pp. 163–198.

[27] D. Talay and L. Tubaro, *Expansion of the global error for numerical schemes solving stochastic differential equations*, Stochastic Anal. Appl. 8 (1990), pp. 94–120.

[28] K. C. Zygalakis, *On the existence and the applications of modified equations for stochastic differential equations*, SIAM J. Sci. Comput., 33 (2011), pp. 102–130.