# KERNEL CONVERGENCE ESTIMATES FOR DIFFUSIONS WITH CONTINUOUS COEFFICIENTS

CLAUDIO ALBANESE

Abstract. Bidirectional valuation models are based on numerical methods to obtain kernels of parabolic equations. Here we address the problem of robustness of kernel calculations vis a vis floating point errors from a theoretical standpoint.

We are interested in kernels of one-dimensional diffusion equations with continuous coefficients as evaluated by means of explicit discretization schemes of uniform step $h > 0$ in the limit as $h \to 0$. We consider both semidiscrete triangulations with continuous time and explicit Euler schemes with time step so small that the Courant condition is satisfied. We find uniform bounds for the convergence rate as a function of the degree of smoothness. We conjecture these bounds are indeed sharp. The bounds also apply to the time derivatives of the kernel and its first two space derivatives. The proof is constructive and is based on a new technique of path conditioning for Markov chains and a renormalization group argument. We make the simplifying assumption of time-independence and use longitudinal Fourier transforms in the time direction. Convergence rates depend on the degree of smoothness and Hölder differentiability of the coefficients. We find that the fastest convergence rate is of order $O(h^2)$ and is achieved if the coefficients have a bounded second derivative. Otherwise, explicit schemes still converge for any degree of Hölder differentiability except that the convergence rate is slower. Hölder continuity itself is not strictly necessary and can be relaxed by an hypothesis of uniform continuity.

Contents



## 1. Introduction and Notations

The numerical calculation of kernels for parabolic equations is essential for the construction of flexible bi-directional valuation models which can be used equally well to price European or callable options one the one hand and for Monte Carlo scenario generation on the other. However, computing kernels is notoriously a very difficult task numerically because of the interplay between high frequency noise due to floating point errors and the hard delta function singularity of the initial condition.

The simplest differentiation schemes are based on explicit Euler methods and have traditionally been discarded as they are stable only if the time step length is shorter than the Courant bound, (Courant *et al.* 1928). In Finance application, the Courant bound is quite short as it ranges from an hour to a day. If one insists on storing matrices in sparse format and insists

*Date*: First version October 24th, 2007, last revision August 25, 2010.

on basing a solution method upon sparse matrix-vector routines, then one is confined to methods involving a number of iterations inversely proportional to the time step and respecting the Courant bound is less than practical. There are thus two alternatives.

A first possibility is to use a semi-implicit Crank-Nicholson method which is unconditionally stable and thus lends itself to longer time steps. However, Crank-Nicholson (Crank and Nicolson 1947) schemes are only weakly stable: they are suitable for bounded initial conditions and thus not for kernels, (Mingyou and Thomee 1982). Even for non-smooth payoffs such as that of a call option one notices decaying spurious oscillations if the ratio of time step to the Courant bond for the stability of explicit Euler methods is larger than 3-4. An alternative within the realm of unconditionally stable methods is given by fully implicit schemes which are strongly stable and suitable for kernels but they are imprecise. A fully implicit scheme can in fact be interpreted as an explicit scheme for a time changed process with jumps characterized by a long stability time step due to an explicit regularization of the diffusion term.

In this paper, we explore a different venue and indeed choose to respect the Courant condition and to work with a conditionally stable explicit differentiation scheme. This is possible only if one is willing to use third level BLAS methods, i.e. methods based on matrix multiplication and ultimately fast exponentiation as these methods are insensitive to the length of the time step. Current computing technology not only gives access to high performance accelerators for matrix-matrix multiplication but even favors these methods as they are the only ones which are truly compute bound: second level sparse BLAS methods instead tend to be memory bound and are unable to scale well with the computing power of multi-core architectures.

This paper was motivated by the empirical observation of a great degree of robustness and stability of explicit methods when used for the purpose of calculating kernels of stochastic differential equations. Extensive empirical evidence accumulated over several years of model building for derivatives across nearly all asset classes show without any doubt that third level BLAS methods can be confidently used to implement truly bi-directional models even when using single precision floating point arithmetics. In addition to that, even derivatives of kernels are accessible via numerical differentiation.

The extraordinary difference in degree of robustness between explicit methods and semi-implicit one needs to be understood theoretically. Unlike a fully implicit scheme, an explicit scheme is not introducing an approximation and a regularization of the original process, but instead represents it quite faithfully.

The observed signal-to-noise level for explicit methods is much lower than what any naive back-of-the-envelope estimate would arrive at by means of worst case type of estimates whereby one resums errors in absolute value. One is led to conclude that systematic cancelations of errors of different sign is responsible for this behavior, a subtle phenomenon that we think is important to understand mathematically. In this paper, we attempt to demonstrate that the dampening of the noise caused by floating point errors is due in fact to the natural smoothing properties of parabolic differential equations.

The reality of floating point errors does not lend itself easily to faithful mathematical modeling, especially when one aims at deriving rigorous results. We thus have to deform the problem into a toy model. We consider a pair of backward and forward one-dimensional diffusion equations of the form

$$\frac{\partial}{\partial t} f(x;t) + \mathcal{L}^0_x f(x;t) = 0, \qquad \frac{\partial}{\partial t} g(y;t) = \mathcal{L}^{0*}_y g(y;t) \tag{1.1}$$

where

$$\mathcal{L}^0_x = \frac{1}{2}\sigma(x)^2 \frac{\partial^2}{\partial x^2} + \mu(x)\frac{\partial}{\partial x}. \tag{1.2}$$

and its adjoint formally acts as follows:

$$(\mathcal{L}^{0*}_y \phi)(y) = \frac{1}{2}\frac{\partial^2}{\partial y^2}\big(\sigma(y)^2\phi(y)\big) - \frac{\partial}{\partial y}\big(\mu(y)\phi(y)\big). \tag{1.3}$$

on a test function $\phi$. These equations are defined on the bounded interval $A = [-L, L] \subset \mathbb{R}$ where $0 < L < \infty$. For simplicity, we assume periodic boundary conditions and identify the two boundary points $\pm L$ with each other.

In order to model the effect of floating point errors, in most of the paper the coefficients $\sigma(x)$ and $\mu(x)$ are assumed to be just Hölder differentiable. More precisely, if $\alpha \in (0, 1]$, $k \in \mathbb{N}$ where $\mathbb{N} = \{0, 1, ...\}$ and the function $\phi(x) \in \mathcal{C}^k(A)$ has $k \geq 1$ continuous derivatives, then one says that $\phi$ is Hölder differentiable of order $(k, \alpha)$ if there exists a constant $c > 0$ such that

$$\left|\phi^{(k)}(x) - \phi^{(k)}(y)\right| \leq cd(x, y)^\alpha \tag{1.4}$$

uniformly for all $x, y \in A$. In case $k = 0$ the function is called Hölder continuous. The distance is defined consistently with the assumed periodic boundary conditions and is given by

$$d(x, y) = \min_n |x - y - 2Ln|. \tag{1.5}$$

The linear space of Hölder continuous or Hölder differentiable periodic functions of order $(k, \alpha)$ on $A$ is denoted with $\mathcal{C}^{k,\alpha}(A)$. We are interested in the case where $\sigma^2 \in \mathcal{C}^{k,\alpha}(A)$ and $\mu \in \mathcal{C}^{j,\beta}(A)$ with $k + \alpha > 0$ and $j + \beta > 0$.

The hypothesis of Hölder continuity can be relaxed slightly by assuming uniform continuity instead, i.e. that both $\sigma(x)^2$ and $\mu(x)$ satisfy a bound of the form

$$\left|\phi(x) - \phi(y)\right| \leq c\rho\big(d(x, y)\big) \tag{1.6}$$

where $\rho(d)$ is a non-decreasing function such that $\lim_{d\downarrow 0} \rho(d) = 0$.

Let $u_0(x, y; t)$ denote a kernel of equation (1.17), i.e. a weak solution of the forward equation

$$\frac{\partial}{\partial t} u_0(x, y; t) = \mathcal{L}_y^{0*} u_0(x, y; t) \tag{1.7}$$

where the operator $\mathcal{L}_y^{0*}$ acts on the $y$ coordinate and the following initial time condition is satisfied:

$$\lim_{t\downarrow 0} u_0(x, y; t) = \delta(x - y). \tag{1.8}$$

The kernel $u_0(x, y; t)$ formally satisfies also the backward equation

$$\frac{\partial}{\partial t} u_0(x, y; t) + \mathcal{L}_x^0 u_0(x, y; t) = 0 \tag{1.9}$$

where the operator $\mathcal{L}_x^0$ acts on the $x$ coordinate.

We are interested in existence, uniqueness, smoothness and approximation schemes for the kernel $u_0(x, y; t)$, its first two space derivatives with respect to the $x$ variable and its first time derivative $\partial_t u_0(x, y; t)$. As a byproduct of this analysis, we also find conclusions about the convergence of $\mathcal{L}_x^0 u_0(x, y; t)$ and $\mathcal{L}_y^{0*} u_0(x, y; t)$, as both expressions equal the first time derivative.

Diffusion equations are one of the single most studied topics in the literature. Existence and uniqueness questions for the kernel were address in (Kolmogorov 1931), (Feller 1936), (Hille 1948), (Yosida 1951) and (Ito 1957). A classification of all the possible boundary conditions is in (Feller 1952). The case of Hölder continuous coefficients was resolved in (Philips 1961) based on methods in (Friedrichs 1958) and (Lax and Phillips 1960). The hypothesis of Hölder continuity was relaxed to uniform continuity in (Fabes and Riviere 1966) and (Stroock and Varadhan 1969). Strook and Varadhan also introduce a new probabilistic framework where existence is proved by reduction to the so-called martingale problem and a compactness argument, thus shifting the attention from the kernel itself to the underlying measure space.

The existence of a weak limit of continuous time Markov chains as $h_m \downarrow 0$ was established in (Sova 1967) and (Kurtz 1969) by using operator semigroup methods, see also the book (Ethier and Kurtz 1986) for a review. Convergence in the semigroup sense takes place if the limit

$$(T_t\phi)(x) = \lim_{m\to\infty, m\geq n} \sum_{y\in h_m Z\cap A} u_{h_m}(x, y; t)\phi(y) \tag{1.10}$$

exists for all test functions $\phi \in \mathcal{C}^\infty(A)$, uniformly for all $x \in A_n, n \geq 0$. A key result is that a necessary and sufficient condition for this limit to exist and define a semigroup $T_t$ is that generators converge also in the same Banach space, i.e. that also the limit

$$\lim_{h_m \downarrow 0} \mathcal{L}_x^{h_m} \phi = \mathcal{L}_x^0 \phi \tag{1.11}$$

exists in the uniform norm for all test functions $\phi \in \mathcal{C}^\infty(A)$. See (Ethier and Kurtz 1986) for a precise statement with all the technical conditions and a proof. In (Stroock and Varadhan 1979) convergence is reconsidered again by reduction to the martingale problem.

The problem has also been studied extensively in the numerical analysis literature. Explicit and implicit Euler schemes where coefficients are smooth and the data is rough in the sense that it belongs to a $L^2$ space have been considered by several authors. In the case that the Markov generator is symmetric and time independent, one can make use of a spectral representation as in (Baker *et al.* 1977) and with greater effort such methods may also be used for more general situations, see (Suzuki 1978). In (Luskin and Rannacher 1978), a parabolic duality argument is used to show convergence for the standard Galerkin method. (Mingyou and Thomee 1982) use a simpler argument based on energy estimates. In (Palencia 1996) one finds convergence bounds in maximum norm assuming the initial condition is uniformly bounded and coefficients are constant.

In this article, we revisit this classic theme by considering the problem of obtaining the kernel constructively as a limit of increasingly fine triangulations schemes and in assessing the rate of convergence with pointwise bounds on the kernel itself. More precisely, let $h_m = L2^{-m}, m \in \mathbb{N}$ and let $A_m h_m \mathbb{Z} \cap A$. Consider the sequence of operators

$$\mathcal{L}_x^m = \frac{\sigma(x)^2}{2} \Delta_x^m + \mu(x) \nabla_x^m. \tag{1.12}$$

defined on the $2^{m+1}$-dimensional space of all periodic functions $f_m : A_m \to \mathbb{R}$, where

$$\nabla_x^m f(x) = \frac{f(x+h_m) - f(x-h_m)}{2h_m}. \tag{1.13}$$

and

$$\Delta_x^m f(x) = \frac{f(x+h_m) + f(x-h_m) - 2f(x)}{{h_m}^2} \tag{1.14}$$

These definitions also apply to the boundary points by periodicity. We assume that $m \geq m_0$ where $m_0$ is the least integer such that

$$\frac{\sigma^2(x)}{2h_m^2} > \frac{|\mu(x)|}{2h_m} \tag{1.15}$$

for all $m \geq m_0$ and all $x \in A_m$.

Let $u_m(x, y; t)$ denote the kernel of equation (1.17), i.e. the solution of the (forward) equation

$$\frac{\partial}{\partial t} u_m(x, y; t) = \mathcal{L}_y^{m*} u_m(x, y; t) \tag{1.16}$$

where the operator $\mathcal{L}_y^{m*}$ acts on the $y$ coordinate and the following initial time condition is satisfied:

$$\lim_{t \downarrow 0} u_m(x, y; t) = h_m^{-1} \delta_m(x-y). \tag{1.17}$$

Here,

$$\delta_m(x-y) = \begin{cases} 1 & \text{if } \ x = y \mod 2L \\ 0 & \text{otherwise.} \end{cases} \tag{1.18}$$

Since (1.17) is a finite system of linear ordinary differential equations, the solution exists and is unique for all times. The kernel $u_m(x, y; t)$ satisfies also the backward equation

$$\frac{\partial}{\partial t} u_m(x, y; t) + \mathcal{L}_x^m u_m(x, y; t) = 0 \tag{1.19}$$

where the operator $\mathcal{L}^m_x$ acts on the $x$ coordinate. Using functional calculus notations for the exponential of a matrix, we also have that

$$u_m(x, y; t) = \frac{1}{h_m} \exp\big(t\mathcal{L}^m\big)(x, y). \tag{1.20}$$

Our main result can be stated as follows:

**Theorem 1.** *Suppose that $\sigma^2 \in \mathcal{C}^{k,\alpha}$ and $\mu \in \mathcal{C}^{j,\beta}$ and let*

$$\gamma = \min\{2, k + \alpha, j + \beta\}. \tag{1.21}$$

*Assume that $\gamma > 0$ and that $\sigma(x) \geq \Sigma_0$ for some constant $\Sigma_0 > 0$. Then there is a constant $c > 0$ such that for all $m' \geq m \geq m_0$ and all $x, y \in A_m$ the following inequalities hold:*

(i)

$$|u_m(x, y; t) - u_{m'}(x, y; t)| \leq c h_m^\gamma \tag{1.22}$$

(ii)

$$\begin{aligned} &|\partial_t u_m(x, y; t) - \partial_t u_{m'}(x, y; t)| \\ &\quad = |\mathcal{L}^{m*}_y u_m(x, y; t) - \mathcal{L}^{m'*}_y u_{m'}(x, y; t)| \\ &\quad = |\mathcal{L}^m_x u_m(x, y; t) - \mathcal{L}^{m'}_x u_{m'}(x, y; t)| \leq c h_m^\gamma. \end{aligned} \tag{1.23}$$

A version of this theorem under slightly weaker conditions can be formulated as follows:

**Theorem 2.** *Suppose that $\sigma(x)^2$ and $\mu(x)$ are uniformly continuous functions in $A$. Let the function $\rho(d)$ be non-decreasing and be such that $\lim_{d\downarrow 0} \rho(d) = 0$ and equation (1.6) holds. Assume also that $\sigma(x) \geq \Sigma_0$ for some constant $\Sigma_0 > 0$. Then there is a constant $c > 0$ such that for all $m' \geq m \geq m_0$ and all $x, y \in A_m$ the following inequalities hold:*

(i)

$$|u_m(x, y; t) - u_{m'}(x, y; t)| \leq c\rho(h_m) \tag{1.24}$$

(ii)

$$\begin{aligned} &|\partial_t u_m(x, y; t) - \partial_t u_{m'}(x, y; t)| \\ &\quad = |\mathcal{L}^{m*}_y u_m(x, y; t) - \mathcal{L}^{m'*}_y u_{m'}(x, y; t)| \\ &\quad = |\mathcal{L}^m_x u_m(x, y; t) - \mathcal{L}^{m'}_x u_{m'}(x, y; t)| \leq c\rho(h_m). \end{aligned} \tag{1.25}$$

Next, we consider the case where also time is discretized and prove the following result:

**Theorem 3.** *Suppose that $\sigma^2$ and $\mu$ satisfy equations of the form (1.6) with a non-decreasing function $\rho(d)$ such that $\lim_{d\downarrow 0} \rho(d) = 0$. Assume also that $\sigma(x) \geq \Sigma_0$ for some constant $\Sigma_0 > 0$. Consider the discretized kernel*

$$u^{\delta t}_m(x, y; t) = h_m^{-1} \left(1 + \delta t \mathcal{L}^m\right)^{\left[\frac{t}{\delta t_m}\right]}(x, y; t). \tag{1.26}$$

*where $\mathcal{L}^m$ is the operator in (1.12) and $\delta t_m$ is so small that*

$$\min_{x \in A_m} 1 + \delta t_m \mathcal{L}^m(x, x) > 0 \tag{1.27}$$

*Assume that boundary conditions are periodic and that the ratio $\frac{t}{\delta t} = N$ is an integer. Then here is a constant $c > 0$ such that the following bounds hold for all $m \geq m_0$ and all $x, y \in A_m$:*

(i)

$$|u_m(x, y; t) - u^{\delta t}_m(x, y; t)| \leq c h_m^2 \tag{1.28}$$

(ii)

$$\left|\partial_t u_m(x,y;t) - \frac{u_m^{\delta t}(x,y;t+\delta t) - u_m^{\delta t}(x,y;t)}{\delta t}\right|$$
$$= |\mathcal{L}_y^{m*} u_m(x,y;t) - \mathcal{L}_y^{m*} u_m^{\delta t}(x,y;t)|$$
$$= |\mathcal{L}_x^{m*} u_m(x,y;t) - \mathcal{L}_x^{m*} u_m^{\delta t}(x,y;t)| \leq ch^2 \tag{1.29}$$

The conditions we choose to work under are fairly restrictive but can presumably be relaxed. The assumption that boundary conditions are periodic is for instance a substantial limitation. Dirichlet, i.e. absorbing, boundary conditions would cause an accumulation of probability at the boundary translating in a delta function singularity for the kernel and thus spoil our regularity estimates. However, one would expect the uniform estimates to extend to this case if one had to restrict bounds to apply strictly away from the boundary.

The strong ellipticity condition in (1.15) can also presumably be relaxed. In (Albanese 2006) we extend the results in this article to the hypo-elliptic case of the joint process between a stochastic differential equation and a stochastic integral over it. However, the estimates we obtain involve a Hardy-type norm in the analytic continuation of a partial Fourier transform of the joint kernel. This is needed to regularize the singularities of the joint kernel itself.

Finally, a question that is interesting but will remain unanswered is whether the results in this article extend to higher dimensional problems. We expect they do under the Hörmander conditions, but the methods of proof in this paper do not seem to extend immediately to the more general multi-dimensional case.

The paper is organized as follows. In Section 2 we consider the case of Brownian motion and review a result in (Albanese and Mijatovic 2006) which establishes the theorems above in this simple particular case where Fourier analysis in the space direction can be used to carry out a precise calculation. In Section 3, we consider the case of a diffusion where both the volatility and the drift have two bounded derivatives. In this case, we make use of time-homogeneity and carry out a Fourier transform in the time direction after path conditioning. In Section 4, we extend the derivation to the case of non-smooth coefficients. Section 5 is dedicated to the case where time is discretized and we prove Theorem 3.

## 2. Constant Coefficients

In this Section, we prove Theorem 1 in the special case where the volatility and the drift coefficients are constant, i.e.

$$\mathcal{L}^m{}_x = \mu \nabla_x^m + \frac{1}{2}\sigma^2 \Delta_x^m. \tag{2.1}$$

It suffices to consider the case $m' = m+1$. Let $B_m$ be the Brillouin zone defined as follows:

$$B_m = \left\{ -\frac{2^{m-1}\pi}{L} + \frac{k\pi}{L}, k = 0, ..2^m - 1 \right\} \tag{2.2}$$

Let $\mathcal{F}_m : \ell^2(A_m) \to \ell^2(B_m)$ be the Fourier transform operator defined so that:

$$\hat{f}(p) \equiv \mathcal{F}_m(f)(p) = h_m \sum_{x \in A_m} f(x) e^{-ipx} \tag{2.3}$$

for all $p \in B_m$. The inverse Fourier transform is given by

$$\mathcal{F}_m^{-1}(\hat{f})(x) = \frac{1}{2L} \sum_{p \in B_m} \hat{f}(p) e^{ipx}. \tag{2.4}$$

The Fourier transformed generator is diagonal and is given by the operator of multiplication by

$$\hat{\ell}^m(p) = \mathcal{F}_m \mathcal{L}^m \mathcal{F}_m^{-1}(p,p) = -i\mu \frac{\sin h_m p}{h_m} + \sigma^2 \frac{\cos h_m p - 1}{h_m^2}. \tag{2.5}$$

We have

$$u_m(x,y;t) = \frac{1}{2L}\sum_{p\in B_m} e^{t\hat{\ell^m}(p)} e^{ip(y-x)}. \tag{2.6}$$

Using this Fourier series representation, we find

$$\begin{aligned}&\left|u_m(x,y;t)-u_{m+1}(x,y;t)\right|\\ &\le \frac{1}{2L}\left|\sum_{p\in B_m}\left(e^{t\hat{\ell}^m(p)}-e^{t\hat{\ell}^{m+1}(p)}\right)e^{ip(y-x)}\right| + \frac{1}{2L}\left|\sum_{p\in B_{m+1}\setminus B_m} e^{t\hat{\ell}^{m+1}(p)}e^{ip(y-x)}\right|.\end{aligned} \tag{2.7}$$

Let

$$K_m = \sqrt{\frac{|\log h_{m+1}|}{\sigma^2 t}}. \tag{2.8}$$

If $h_m$ is small enough, i.e. if $m_0$ is sufficiently large, we have that

$$\frac{1}{2L}\left|\sum_{p\in B_m,|p|\ge K_m} e^{t\hat{\ell}^m(p)}e^{ip(y-x)}\right| \le \frac{1}{2L}\sum_{p\in B_m,|p|\ge K_m} e^{t\Re(\hat{\ell}^m(p))} \le c\exp\left(t\sigma^2\frac{\cos h_m K_m - 1}{h_m^2}\right) \le ch_m^2. \tag{2.9}$$

where $\Re(a)$ denotes the real part of $a\in\mathbb{C}$ and $c$ denotes a generic constant. Similarly

$$\begin{aligned}\frac{1}{2L}\left|\sum_{p\in B_{m+1},|p|\ge K_m} e^{t\hat{\ell}^{m+1}(p)}e^{ip(y-x)}\right| &\le \frac{1}{2L}\sum_{p\in B_m,|p|\ge K_m} e^{t\Re(\hat{\ell}^{m+1}(p))}\\ &\le c\exp\left(t\sigma^2\frac{\cos h_{m+1}K - 1}{h_{m+1}^2}\right) \le ch_{m+1}^2\end{aligned} \tag{2.10}$$

Since

$$\frac{1}{2}h^2p^3 - \frac{1}{8}h^4p^5 \le \frac{\sin hp}{h} - \frac{\sin 2hp}{2h} \le \frac{1}{2}h^2p^3 \tag{2.11}$$

and

$$-\frac{1}{8}h^2p^4 \le \frac{\cos hp - 1}{h^2} - \frac{\cos 2hp - 1}{(2h)^2} \le -\frac{1}{8}h^2p^4 + \frac{1}{48}h^4p^6. \tag{2.12}$$

we find that if $|p|\le \frac{\sqrt{2}}{h}$ then

$$|\hat{\ell}^m(p) - \hat{\ell}^{m+1}(p)| \le \frac{\mu}{4}h^2|p|^3 + \frac{\sigma^2}{16}h^2p^4. \tag{2.13}$$

Moreover, since

$$-\frac{1}{2}p^2 \le \frac{\cos hp - 1}{h} \le -\frac{1}{2}p^2 + \frac{1}{24}h^2p^4 \tag{2.14}$$

we conclude that in case $|p|\le h^{-1}\sqrt{\frac{2}{3}}$, the following inequality holds:

$$\frac{\cos hp - 1}{h} \le -\frac{1}{4}p^2 \tag{2.15}$$

Hence, if $m_0$ is large enough, we find

$$\frac{1}{2L}\Bigg|\sum_{p\in B_m,|p|\le K}\left(e^{t\hat\ell^m(p)}-e^{t\hat\ell^{m+1}(p)}\right)e^{ip(y-x)}\Bigg|$$

$$\le\frac{1}{2L}\sum_{p\in B_m,|p|\le K}e^{-\frac14 p^2}\left(e^{\frac{\mu t}{4}h_m^2|p|^3+\frac{\sigma^2 t}{16}h_m^2p^4}-1\right)$$

$$\le\frac{1}{2L}\sum_{p\in B_m,|p|\le K}e^{-\frac14 p^2}\left(\frac{\mu t}{4}h_m^2|p|^3+\frac{\sigma^2 t}{16}h_m^2p^4\right)\le ch_m^2 \tag{2.16}$$

for some constant $c>0$ independent of $m$. This concludes the proof of convergence for the kernel in the special case of constant coefficients.

To estimate the first derivative, notice that

$$\nabla u_m(x,y;t)=\frac{1}{L}\sum_{p\in B_m}e^{t\hat\ell^m(p)}\frac{\sin ph_m}{h_m}e^{ip(y-x)}. \tag{2.17}$$

and

$$\big|u_m(x,y;t)-u_{m+1}(x,y;t)\big|$$

$$\le\frac{1}{2L}\Bigg|\sum_{p\in B_m}\left(e^{t\hat\ell^m(p)}\frac{\sin ph_m}{h_m}-e^{t\hat\ell^{m+1}(p)}\frac{\sin ph_{m+1}}{h_{m+1}}\right)e^{ip(y-x)}\Bigg|$$

$$+\frac{1}{2L}\Bigg|\sum_{p\in B_{m+1}\setminus B_m}e^{t\hat\ell^{m+1}(p)}e^{ip(y-x)}\Bigg|. \tag{2.18}$$

Let

$$K_m=2\sqrt{\frac{|\log h_{m+1}|}{\sigma^2 t}}. \tag{2.19}$$

If $h_m$ is small enough, we have that

$$\frac{1}{2L}\Bigg|\sum_{p\in B_m,|p|\ge K_m}e^{t\hat\ell^m(p)}\frac{\sin ph_m}{h_m}e^{ip(y-x)}\Bigg|\le\frac{1}{2L}\sum_{p\in B_m,|p|\ge K_m}e^{t\Re(\hat\ell^m(p))}\frac{\sin ph_m}{h_m}$$

$$\le c\left|\frac{\sin Kh_m}{h_m}\right|\exp\left(t\sigma^2\frac{\cos Kh_m-1}{h_m^2}\right)\le ch_m^2. \tag{2.20}$$

where $c$ denotes a generic constant. Similarly

$$\frac{1}{2L}\Bigg|\sum_{p\in B_{m+1},|p|\ge K_m}e^{t\hat\ell^{m+1}(p)}e^{ip(y-x)}\Bigg|\le ch^2. \tag{2.21}$$

If $m$ is large enough, we also find

$$\frac{1}{2L}\Bigg|\sum_{p\in B_m,|p|\le K_m}\left(\frac{\sin ph_m}{h_m}e^{t\hat\ell^m(p)}-\frac{\sin ph_{m+1}}{h_{m+1}}e^{t\hat\ell^{m+1}(p)}\right)e^{ip(y-x)}\Bigg|$$

$$\le\frac{1}{2L}\sum_{p\in B_m,|p|\le K_m}\left|\frac{\sin ph_m}{h_m}\right|e^{-\frac14 p^2}\left(e^{\frac{\mu t}{4}h_m^2|p|^3+\frac{\sigma^2 t}{16}h_m^2p^4}-1\right)$$

$$+e^{-\frac14 p^2}\left|\frac{\sin ph_{m+1}}{h_{m+1}}-\frac{\sin ph_m}{h_m}\right|\le ch_m^2 \tag{2.22}$$

for some constant $c>0$ independent of $m$. This concludes the proof of the bound of the first derivative. The second derivative can be derived in a similar way.

Finally, consider the following Fourier representation for the discretized kernel

$$u^{\delta t}_m(x,y;t) = \frac{1}{L}\sum_{p\in B_m}\left(1+\delta t\hat{\ell^m}(p)\right)^{\frac{t}{\delta t}} e^{ip(y-x)}. \tag{2.23}$$

Consider the formula

$$\left(1+\delta t\hat{\ell^m}(p)\right)^{\frac{t}{\delta t}} = \exp\left(t\log\left(1+\hat{\ell^m}(p)\right)\right). \tag{2.24}$$

and let's represent the difference between the discrete and continuous time kernels as follows:

$$\begin{aligned}
&|u_m(x,y;t) - u^{\delta t}_m(x,y;t)|\\
&\leq \frac{1}{2L}\left|\sum_{p\in B_m}\left(\exp\left(t\hat{\ell}^m(p)\right) - \exp\left(\frac{t}{\delta t}\log\left(1+\delta t\hat{\ell}^m(p)\right)\right)e^{ip(y-x)}\right|\\
&\leq \frac{1}{2L}\sum_{p\in B_m, p\leq K_m} e^{-\frac{1}{4}p^2}\left|\exp\left(\frac{t}{\delta t}\log\left(1+\delta t\hat{\ell}^m(p)\right) - t\hat{\ell}^m(p)\right)-1\right|\\
&\quad \frac{1}{2L}\sum_{p\in B_m, p\geq K_m}\left|\exp\left(t\hat{\ell}^m(p)\right)\right| + \frac{1}{2L}\sum_{p\in B_m, p\geq K_m}\left|\exp\left(\frac{t}{\delta t}\log\left(1+\delta t\hat{\ell}^m(p)\right)\right)\right|
\end{aligned} \tag{2.25}$$

where $K_m$ is chosen as in (2.8). The very same bounds above lead to the conclusion that this difference is $\leq ch_m^2$.

## 3. Smooth Coefficients

In this section, we prove Theorem 1 in the case where the drift and volatility are both of class $\mathcal{C}^{(3,0)}$, i.e. they depend smoothly on the space coordinate but not on the time coordinate.

Let us introduce the following two constants characterizing the volatility function:

$$\Sigma_0 = \inf_{x\in A_m}\sigma(x), \qquad \Sigma_1 = \sup_{x\in A_m}\sqrt{\sigma(x)^2 + h_m|\mu(x)|}. \tag{3.1}$$

and let

$$M = \sup_{x\in A_m}|\mu(x)|. \tag{3.2}$$

Due to our assumptions, we have that $\Sigma_0 > 0$ and $\Sigma_1, M < \infty$.

A *symbolic path* $\gamma = \{\gamma_0,\gamma_1,\gamma_2,....\}$ is an infinite sequence of sites in $A_m$ such that $\gamma_j \neq \gamma_{j-1}$ for all $j = 1,....$ Let $\Gamma_m$ be the set of all symbolic paths in $A_m$. The kernel of the diffusion process admits the following representation in terms of a summation over symbolic paths

$$u_m(x,y;t) = \frac{1}{h_m}\sum_{q=1}^{\infty}2^{-q}\sum_{\substack{\gamma\in\Gamma_m:\gamma_0=x,\gamma_q=y\\ |\gamma_j-\gamma_{j-1}|=1\ \forall j\geq 1}} W_m(\gamma,q,t) \tag{3.3}$$

where

$$W_m(\gamma,q,t) = \frac{1}{h_m}\int_0^t ds_1\int_{s_1}^t ds_2...\int_{s_{q-1}}^t ds_q e^{(t-s_q)\mathcal{L}^m(\gamma_q,\gamma_q)}\prod_{j=0}^{q-1}\left(e^{(s_{j+1}-s_j)\mathcal{L}^m(\gamma_j,\gamma_j)}2\mathcal{L}^m(\gamma_j,\gamma_{j+1})\right) \tag{3.4}$$

with $s_0 = 0$.

Let us introduce the following Green's function:

$$G_m(x,y;\omega) = \int_0^\infty u_m(x,y;t)e^{-i\omega t}dt = h_m^{-1}\frac{1}{\mathcal{L}^m + i\omega}(x,y). \tag{3.5}$$

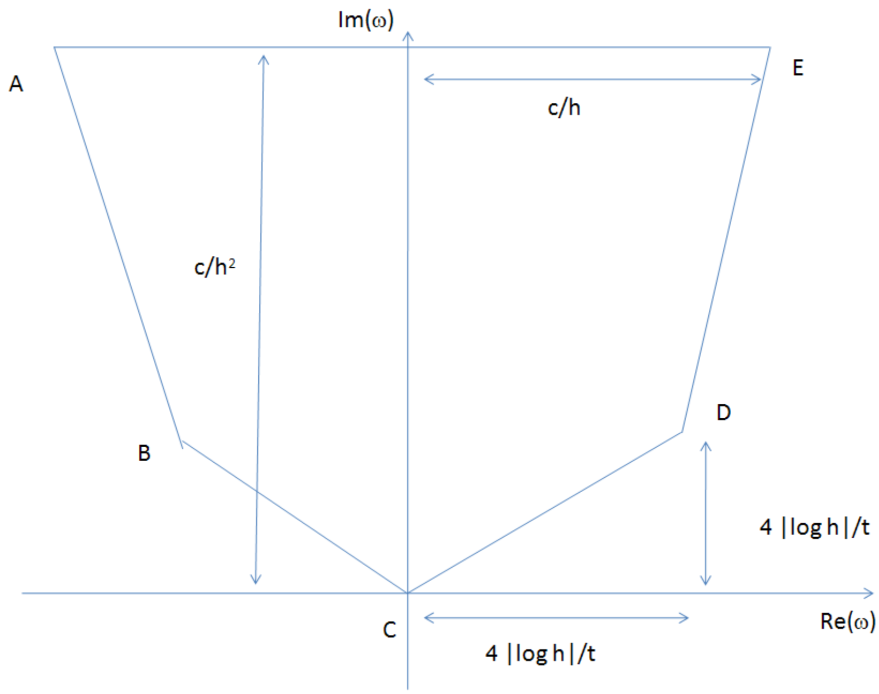


FIGURE 1. Contour of integration for the integral in (3.52). $\mathcal{C}_+$ is the countour joining the point $D$ to the points $E, A, B$. $\mathcal{C}_-$ is the countour joining the point $B$ to $C$ to $D$.

The propagator can be expressed as the following contour integral

$$u_m(x, y; t) = \int_{\mathcal{C}_-} \frac{d\omega}{2\pi} G_m(x, y; \omega) e^{i\omega t} + \int_{\mathcal{C}_+} \frac{d\omega}{2\pi} G_m(x, y; \omega) e^{i\omega t}. \tag{3.6}$$

Here, $\mathcal{C}_+$ is the contour joining the point $D$ to the points $E, A, B$ in Fig. 1, while $\mathcal{C}_-$ is the contour joining the point $B$ to $C$ to $D$. By design, each point $\omega$ on the upper path $\mathcal{C}_+$ is separated from the spectrum of $\mathcal{L}$.

**Lemma 1.** *There is a constant $c$ such that, if $m \geq m_0$ then*

$$\left| \int_{\mathcal{C}_+} \frac{d\omega}{2\pi} G_m(x, y; \omega) e^{i\omega t} \right| \leq ch^2. \tag{3.7}$$

*Proof.* The proof is based on the geometric series expansion

$$G_m(\omega) = h_m^{-1} \frac{1}{\mathcal{L}^m + i\omega} = h_m^{-1} \sum_{j=0}^{\infty} \frac{1}{\frac{1}{2}\sigma^2 \Delta^m + i\omega} \left[ \mu \nabla^m \frac{1}{\frac{1}{2}\sigma^2 \Delta^m + i\omega} \right]^j \tag{3.8}$$

whose convergence for $\omega \in \mathcal{C}_+$ can be established by means of a Kato-Rellich relative bound, see (Kato 1966). More precisely, for any $\alpha > 0$, one can find a $\beta > 0$ such that the operators $\nabla^m$ and $\Delta^m$ satisfy the following relative bound estimate:

$$||\nabla^m f||_2 \leq \alpha ||\Delta^m f||_2 + \beta ||f||_2. \tag{3.9}$$

for all periodic functions $f$ and all $m \geq m_0$. This bound can be derived by observing that $\nabla^m$ and $\Delta^m$ can be diagonalized simultaneously by a Fourier transform, as done in the previous section, and by observing that for any $\alpha > 0$, one can find a $\beta > 0$ such that

$$\left| \frac{\sin h_m p}{h_m} \right| \leq \alpha \left| \frac{\cos h_m p - 1}{h_m^2} \right| + \beta \tag{3.10}$$

for all $m \geq m_0$ and all $p \in B_m$.

Under the same conditions, we also have that

$$||\mu\nabla^m f||_2 \leq \frac{2M\alpha}{\Sigma_0^2}\left|\left|\frac{1}{2}\sigma^2\Delta^m f\right|\right|_2 + \beta||f||_2. \tag{3.11}$$

Hence

$$\left|\left|\mu\nabla^m \frac{1}{\frac{1}{2}\sigma^2\Delta^m + i\omega}f\right|\right|_2 \leq \frac{2M\alpha}{\Sigma_0^2}\left|\left|\frac{1}{2}\sigma^2\Delta^m \frac{1}{\frac{1}{2}\sigma^2\Delta^m + i\omega}f\right|\right|_2 + \beta\left|\left|\frac{1}{\frac{1}{2}\sigma^2\Delta^m + i\omega}f\right|\right|_2 < 1 \tag{3.12}$$

where the last inequality holds if $\omega \in \mathcal{C}_+$, if $\alpha$ is chosen sufficiently small and if $m$ is large enough. In this case, the geometric series expansion converges in (3.8) converges in $L^2$ operator norm. The uniform norm of the kernel $|G_m(x, y; \omega)|$ is pointwise bounded from above by $h_m^{-1}$.

Since the points $B$ and $D$ have imaginary part equal at height $4\frac{|\log h_m|}{t}$, the integral over the contour $\mathcal{C}_+$ converges also and is bounded from above by $ch_m^2$ in uniform norm.

□

**Lemma 2.** *If* $q \geq \frac{e^2\Sigma_1^2 t}{2h_m^2}$ *we have that*

$$W_m(\gamma, q; t) \leq \sqrt{\frac{q}{2\pi}} \exp\left(-\frac{\Sigma_0^2 t}{2} - q\right). \tag{3.13}$$

*Proof.* Let us define the function

$$\phi(t) = \frac{\Sigma_1^2}{2h_m^2}\, e^{-\frac{\Sigma_0^2 t}{2h_m^2}}\, 1(t \geq 0) \tag{3.14}$$

where $1(t \geq 0)$ is the characteristic function of $\mathbb{R}_+$. We have that

$$W_m(\gamma, q; t) \leq \phi^{\star q}(t) \tag{3.15}$$

where $\phi^{\star q}$ is the $q-$th convolution power, i.e. the $q-$fold convolution product of the function $\phi$ by itself. The Fourier transform of $\phi(t)$ is given by

$$\hat{\phi}(\omega) = \frac{\Sigma_1^2}{2h_m^2}\int_0^\infty e^{-i\omega t - \frac{\Sigma_0^2 t}{2h_m^2}}\, dt = \frac{\Sigma_1^2}{2i\omega h_m^2 + \Sigma_0^2}. \tag{3.16}$$

The convolution power is given by the following inverse Fourier transform:

$$\phi^{\star q}(t) = \int_{-\infty}^\infty \hat{\phi}(\omega)^q e^{i\omega t}\frac{d\omega}{2\pi} = \left(\frac{\Sigma_1}{\Sigma_0}\right)^{2q}\int_{-\infty}^\infty \left(1 + \frac{2i\omega h_m^2}{\Sigma_0^2}\right)^{-q} e^{i\omega t}\frac{d\omega}{2\pi}. \tag{3.17}$$

Introducing the new variable $z = 1 + \frac{2i\omega h_m^2}{\Sigma_0^2}$, the integral can be recast as follows

$$\phi^{\star q}(t) = \frac{\Sigma_0^{2-2q}\Sigma_1^{2q}}{4\pi i h_m^2}\lim_{R\to\infty}\int_{\mathcal{C}_R} z^{-q}\exp\left(\frac{\Sigma_0^2 t}{2h_m^2}(z-1)\right)dz \tag{3.18}$$

where $\mathcal{C}_R$ is the contour in Fig. 2. Using the residue theorem and noticing that the only pole of the integrand is at $z = 0$, we find

$$\phi^{\star q}(t) = \frac{1}{(q-1)!}\left(\frac{\Sigma_1^2 t}{2h_m^2}\right)^q \exp\left(\frac{-\Sigma_0^2 t}{2h_m^2}\right). \tag{3.19}$$

Making use of Stirling's formula $q! \approx \sqrt{2\pi}q^{q+\frac{1}{2}}e^{-q}$, we find

$$\phi^{\star q}(t) \approx \sqrt{\frac{q}{2\pi}}\exp\left(-\frac{\Sigma_0^2 t}{2h_m^2} + q\log\frac{\Sigma_1^2 t}{2h_m^2} + q(1 - \log q)\right). \tag{3.20}$$

If $\log q \geq \log\frac{\Sigma_1^2 t}{2h_m^2} + 2$, then we arrive at the bound in (3.13).

□

It suffices to consider the case $m' = m + 1$ for all values of $m$ above a fixed threshold. In fact, given this particular case, the general statement can be derived with an iterative argument. To this end, we introduce a renormalization group transformation based on the notion of decorating path.

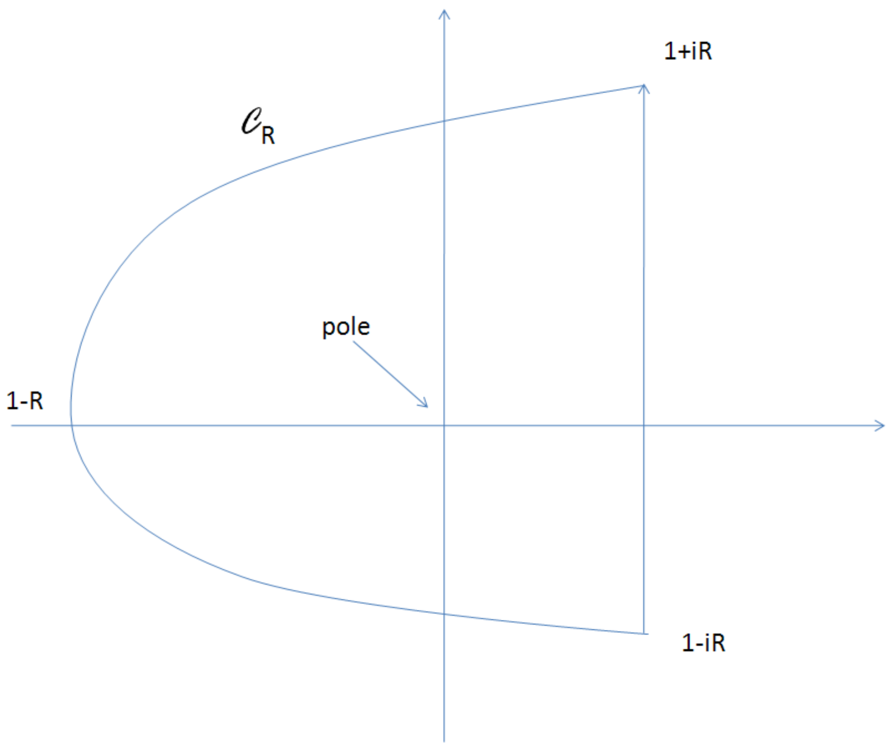


FIGURE 2. Contour of integration $\mathcal{C}_R$ for the integral in (3.18).

**Definition 1. (Decorating Paths.)** *Let $m \geq m_0$ and let $\gamma = \{y_0, y_1, y_2, ....\}$ be a symbolic sequence in $\Gamma_m$. A decorating path around $\gamma$ is defined as a symbolic sequence $\gamma' = \{y_0, y_1', y_2', ....\}$ with $y_i' \in h_{m+1}\mathbb{Z}$ containing the sequence $\gamma$ as a subset and such that if $y_j' = y_i$ and $y_k' = y_{i+1}$, then all elements $y_n'$ with $j < n < k$ are such that $|y_n' - y_j'| \leq h_{m+1}$. Let $\mathcal{D}_{m+1}(\gamma)$ be the set of all decorating sequences around $\gamma$. The decorated weights are defined as follows:*

$$\tilde{W}_m(\gamma, q; t) = \sum_{q'=q}^{\infty} \sum_{\substack{\gamma' \in \mathcal{D}_{m+1}(\gamma) \\ \gamma'_{q'} = \gamma_q}} W_{m+1}(\gamma', q'; t). \tag{3.21}$$

*Finally, let us introduce also the following Fourier transform:*

$$\hat{W}_m(\gamma, q; \omega) = \int_0^\infty W_m(\gamma, q; t) e^{i\omega t} dt, \quad \hat{\tilde{W}}_m(\gamma, q; \omega) = \int_0^\infty \tilde{W}_m(\gamma, q; t) e^{i\omega t} dt. \tag{3.22}$$

**Notation 1.** *In the following, we set $h = h_{m+1}$ so that $h_m = 2h$. We also use the Landau notation $O(h^n)$ to indicate a function $f(h)$ such that $h^{-n} f(h)$ is bounded in a neighborhood of 0.*

**Lemma 3.** *Let $x, y \in A_m$ and let $\mathcal{C}_-$ be an integration contour as in Fig. 1. Then*

$$\left| \left( \int_{\mathcal{C}_-} 2G_{m+1}(x, y; \omega) - G_m(x, y; \omega) \right) e^{i\omega t} \frac{d\omega}{2\pi} \right| = O(h^2). \tag{3.23}$$

*Proof.* We have that

(3.24)
$$2G_{m+1}(x,y;\omega) - G_m(x,y;\omega) = \frac{1}{h}\sum_{q=1}^{\infty} 2^{-q} \sum_{\substack{\gamma\in\Gamma_m : \gamma_0 = x, \gamma_q = y \\ |\gamma_j-\gamma_{j-1}|=1 \forall j\geq 1}} \left(2\hat{\hat{W}}_m(\gamma,q;\omega) - \hat{W}_m(\gamma,q;\omega)\right).$$

The number of paths over which the summation is extended is

$$N(\gamma,q;x,y) \equiv \sharp\{\gamma\in\Gamma_m : \gamma_0 = x, \gamma_q = y, |\gamma_j-\gamma_{j-1}|=1 \forall j\geq 1\} = \binom{q}{\frac{q}{2}+k} \tag{3.25}$$

where $k = \frac{|y-x|}{h_m}$. Applying Stirling's formula we find

$$N_\gamma \lesssim 2^q\sqrt{\frac{2}{\pi q}}. \tag{3.26}$$

Hence

$$\begin{aligned}&\left|\int_{\mathcal{C}_-}\left(2G_{m+1}(x,y;\omega) - G_m(x,y;\omega)\right)e^{i\omega t}\frac{d\omega}{2\pi}\right| \\ &\leq \frac{c}{h}\sum_{q=1}^{\infty}\sqrt{\frac{1}{q}} \max_{\substack{\gamma\in\Gamma_m : \gamma_0 = x, \gamma_q = y \\ |\gamma_j-\gamma_{j-1}|=1 \forall j\geq 1}} \left|\int_{\mathcal{C}_-}\left(2\hat{\hat{W}}_m(\gamma,q;\omega) - \hat{W}_m(\gamma,q;\omega)\right)e^{i\omega t}\frac{d\omega}{2\pi}\right|.\end{aligned} \tag{3.27}$$

for some constant $c \approx \sqrt{\frac{2}{\pi}} > 0$. It suffices to extend the summation over $q$ only up to

$$q_{\max} \equiv \frac{e^2\Sigma_1^2 t}{2h^2}. \tag{3.28}$$

To resum beyond this threshold, one can use the previous lemma. More precisely, we have that

$$\begin{aligned}&\left|\int_{\mathcal{C}_-}\left(2G_{m+1}(x,y;\omega) - G_m(x,y;\omega)\right)e^{i\omega t}\frac{d\omega}{2\pi}\right| \\ &\leq \frac{c\sqrt{q_{\max}}}{h} \max_{\substack{q,\gamma\in\Gamma_m : \gamma_0 = x, \gamma_q = y \\ |\gamma_j-\gamma_{j-1}|=1 \forall j\geq 1}} \left|\int_{\mathcal{C}_-}\left(2\hat{\hat{W}}_m(\gamma,q;\omega) - \hat{W}_m(\gamma,q;\omega)\right)e^{i\omega t}\frac{d\omega}{2\pi}\right|.\end{aligned} \tag{3.29}$$

Let $v(x) = \sigma(x)^2$. To evaluate the resummed weight function, let us form the matrix

$$\bar{\mathcal{L}}(x;h) = \begin{pmatrix} -\frac{v(x+h)}{h^2} & \frac{v(x+h)}{2h^2} - \frac{\mu(x+h)}{2h} & 0 \\ \frac{v(x)}{2h^2} + \frac{\mu(x)}{2h} & -\frac{v(x)}{h^2} & \frac{v(x)}{2h^2} - \frac{\mu(x)}{2h} \\ 0 & \frac{v(x-h)}{2h^2} + \frac{\mu(x-h)}{2h} & -\frac{v(x-h)}{h^2} \end{pmatrix} \tag{3.30}$$

and decompose it as follows:

$$\bar{\mathcal{L}}(x;h) = \frac{1}{h^2}\bar{\mathcal{L}}_0(x) + \frac{1}{h}\bar{\mathcal{L}}_1(x) + \bar{\mathcal{L}}_2(x) + h\bar{\mathcal{L}}_3(x) + O(h^2). \tag{3.31}$$

where

$$\bar{\mathcal{L}}_0(x) = \begin{pmatrix} -v(x) & \frac{1}{2}v(x) & 0 \\ \frac{1}{2}v(x) & -v(x) & \frac{1}{2}v(x) \\ 0 & \frac{1}{2}v(x) & -v(x) \end{pmatrix}, \tag{3.32}$$

$$\bar{\mathcal{L}}_1(x) = \begin{pmatrix} -v'(x) & \frac{1}{2}v'(x) - \frac{1}{2}\mu(x) & 0 \\ \frac{1}{2}\mu(x) & 0 & -\frac{1}{2}\mu(x) \\ 0 & -\frac{1}{2}v'(x) + \frac{1}{2}\mu(x) & v'(x) \end{pmatrix}, \tag{3.33}$$

$$\bar{\mathcal{L}}_2(x) = \begin{pmatrix} -\frac{1}{2}v''(x) & \frac{1}{4}v''(x) - \frac{1}{2}\mu'(x) & 0 \\ 0 & 0 & 0 \\ 0 & \frac{1}{4}v''(x) - \frac{1}{2}\mu'(x) & -\frac{1}{2}v''(x) \end{pmatrix}. \tag{3.34}$$

and

$$\bar{\mathcal{L}}_3(x) = \begin{pmatrix} -\frac{1}{6}v'''(x) & \frac{1}{12}v'''(x) - \frac{1}{4}\mu''(x) & 0 \\ 0 & 0 & 0 \\ 0 & -\frac{1}{12}v'''(x) + \frac{1}{4}\mu''(x) & \frac{1}{6}v'''(x) \end{pmatrix}. \tag{3.35}$$

Let us introduce the sign variable $\tau = \pm 1$, the functions

$$\phi_0(t, x, \tau) \equiv 2\mathcal{L}^m(x, x + 2\tau h) e^{t\mathcal{L}^m(x,x)} 1(t \geq 0) \tag{3.36}$$

$$\phi_1(t, x, \tau) \equiv 2\mathcal{L}^{m+1}(x + \tau h, x + 2\tau h) e^{t\bar{\mathcal{L}}(x;h)}(x, x + \tau h) 1(t \geq 0) \tag{3.37}$$

and their Fourier transforms

$$\hat{\phi}_0(\omega, x, \tau) = \left(\frac{v(x)}{4h^2} + \tau\frac{\mu(x)}{2h}\right)\left(\frac{v(x)}{4h^2} + i\omega\right)^{-1}$$

$$\hat{\phi}_1(\omega, x, \tau) = \left(\frac{v(x)}{h^2} + \tau\frac{\mu(x) + v'(x)}{h} + \frac{v''(x) + \mu'(x)}{2} + \left(\frac{v'''(x)}{6} + \frac{\mu''(x)}{2}\right)\tau h + O(h^2)\right)$$
$$< x|\left(-\bar{\mathcal{L}}(x; h) + i\omega\right)^{-1}|x + \tau h > . \tag{3.38}$$

where

$$|x >= \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \text{and} \quad |x + \tau h >= \begin{pmatrix} \delta_{\tau,1} \\ 0 \\ \delta_{\tau,-1} \end{pmatrix}. \tag{3.39}$$

We also require the functions

$$\psi_0(t, x) \equiv e^{t\mathcal{L}^m(x,x)} 1(t \geq 0), \qquad \psi_1(t, x) \equiv e^{t\bar{\mathcal{L}}(y;h)}(x, x) 1(t \geq 0) \tag{3.40}$$

and the corresponding Fourier transforms

$$\hat{\psi}_0(\omega, x) = \left(\frac{v(x)}{4h^2} + i\omega\right)^{-1}, \qquad \hat{\psi}_1(\omega, x) =< x|\left(-\bar{\mathcal{L}}(x; h) + i\omega\right)^{-1}|x > . \tag{3.41}$$

If $\gamma$ is a symbolic sequence, then

$$\hat{W}_m(\gamma, q; \omega) = \hat{\psi}_0(\omega, \gamma_q) \prod_{j=0}^{q-1} \hat{\phi}_0(\omega; \gamma_j, \mathrm{sgn}(\gamma_{j+1} - \gamma_j)) \tag{3.42}$$

$$\hat{\hat{W}}_m(\gamma, q; \omega) = \hat{\psi}_1(\omega, \gamma_q) \prod_{j=0}^{q-1} \hat{\phi}_1(\omega; \gamma_j, \mathrm{sgn}(\gamma_{j+1} - \gamma_j)). \tag{3.43}$$

Let us estimate the difference between the functions $\hat{\phi}_1(\omega, x, \tau)$ and $\hat{\phi}_2(\omega, x, \tau)$ assuming that $\omega$ is in the contour $\mathcal{C}_-$ in Fig. 2. Retaining only terms up to order up to $O(h^4)$, we find

$$\hat{\phi}_0(\omega, x, \tau) = 1 + \frac{2\mu(x)\tau h}{v(x)} - \frac{4i\omega h^2}{v(x)} - 8\mu(x)\frac{i\omega\tau h^3}{v(x)^2} - \frac{16\omega^2 h^4}{v(x)^2} + O(h^5). \tag{3.44}$$

A lengthy but straightforward calculation which is best carried out using a symbolic manipulation program, gives

$$\hat{\phi}_1(\omega, x, \tau) = 1 + \frac{2\mu(x)\tau h}{v(x)} - \frac{4i\omega h^2}{v(x)} - \left[8\mu(x) - v'(x)\right]\frac{i\omega\tau h^3}{v(x)^2}$$
$$+ r(x) \cdot h^3\tau + i\omega h^4 p(x) - \frac{14\omega^2 h^4}{v(x)^2} + O(h^5) \tag{3.45}$$

where

$$r(x) = \frac{1}{2v(x)^3}\big[\mu''(x)v(x)^2 - 4\mu(x)^3 + 2v'(x)\mu(x)^2 - 2v'(x)v(x)\mu'(x) - \big(2\mu(x)\mu'(x) + v''(x)v(x) - 2v'(x)^2\big)\mu(x)\big].$$

$$p(x) = \frac{1}{v(x)^3}\big[-4\mu(x)^2 + 2v'(x)\mu(x) - 4v(x)\mu'(x) - v''(x)v(x) + 2v'(x)^2\big].$$

(3.46)

We have that

$$\begin{aligned}
&\sum_{j=0}^{q-1}\Big(\log\hat\phi_0(\omega;\gamma_j,\mathrm{sgn}(\gamma_{j+1}-\gamma_j)) - \log\hat\phi_1(\omega;\gamma_j,\mathrm{sgn}(\gamma_{j+1}-\gamma_j))\Big)\\
&= \sum_{j=0}^{q-1}\Big(\frac{i\omega v'(\gamma_j)}{v(\gamma_j)^2} + r(\gamma_j)\Big)h^3\mathrm{sgn}(\gamma_{j+1}-\gamma_j) + \big(|\omega|||p||_\infty + 2|\omega|^2||v^{-2}||_\infty\big)O(h^4q)\\
&= i\omega h^2\Big(\frac{1}{v(\gamma_0)} - \frac{1}{v(\gamma_q)}\Big) + h^2\big(R(\gamma_q) - R(\gamma_0)\big) + \big(|\omega|||p||_\infty + 2|\omega|^2||v^{-2}||_\infty\big)O(h^4q)
\end{aligned}$$

(3.47)

where $R(x)$ is a primitive of $r(x)$, i.e.

$$R(x) = \int_{L_x}^x r(x')dx'. \tag{3.48}$$

We conclude that there is a constant $c > 0$ such that

$$\Big|\int_{\mathcal{C}_-}\Big(\prod_{j=0}^{q-1}\hat\phi_0(\omega;\gamma_j,\mathrm{sgn}(\gamma_{j+1}-\gamma_j)) - \prod_{j=0}^{q-1}\hat\phi_1(\omega;\gamma_j,\mathrm{sgn}(\gamma_{j+1}-\gamma_j))\Big)e^{i\omega t}\frac{d\omega}{2\pi}\Big| \le ch^2. \tag{3.49}$$

for all $q \le q_{\max}$. Here we use the decay of $e^{i\omega t}$ in the upper half of the complex $\omega$ plane to offset the $\omega$ dependencies in the integrand. Similar calculations lead to the following expansions:

$$\hat\psi_0(\omega, x) = \frac{4h^2}{v(x)} + O(\omega h^4), \qquad \hat\psi_1(\omega, x) = \frac{2h^2}{v(x)} + O(\omega h^4) = \frac{1}{2}\hat\psi_0(\omega, x) + O(\omega h^4). \tag{3.50}$$

Since $q < ch^{-2}$ and $\omega \le |\log h|$, we find

$$\Big|\int_{\mathcal{C}_-}\Big(2G_{m+1}(x,y;\omega) - G_m(x,y;\omega)\Big)e^{i\omega t}\frac{d\omega}{2\pi}\Big| \le c\frac{\sqrt{q_{\max}}}{h}h^4 \le ch^2. \tag{3.51}$$

By differentiating with respect to time in equation (3.52), we find that

$$\frac{\partial}{\partial t}u_m(x,y;t) = \int_{\mathcal{C}_-} i\omega G_m(x,y;\omega)e^{i\omega t}\frac{d\omega}{2\pi} + \int_{\mathcal{C}_+} i\omega G_m(x,y;\omega)e^{i\omega t}\frac{d\omega}{2\pi}. \tag{3.52}$$

All the derivations above carry through and we conclude that

$$\Big|\int_{\mathcal{C}_+}\frac{d\omega}{2\pi} i\omega G_m(x,y;\omega)e^{i\omega t}\Big| \le ch^2. \tag{3.53}$$

and also

$$\Big|\int_{\mathcal{C}_-} i\omega\Big(2G_{m+1}(x,y;\omega) - G_m(x,y;\omega)\Big)e^{i\omega t}\frac{d\omega}{2\pi}\Big| \le c\frac{\sqrt{q_{\max}}}{h}h^4 \le ch^2. \tag{3.54}$$

Hence the first time derivatives of the kernel satisfy a similar Cauchy convergence condition as the kernel itself.

□

## 4. The General Case of Holder Continuous Coefficients

In this section we assume coefficients are either Hölder continuous or obey the conditions in Theorem 2.

**Lemma 4.** *Let $f(x)$ be a continuous function in $[-L, L]$ satisfying periodic boundary conditions. Then, for all $h > 0$, we have that*

$$f(x+h_m) = f(x) + h_m \nabla_x^m f(x) + \frac{h^2}{2}\Delta_x^m f(x) \tag{4.1}$$

$$f(x-h_m) = f(x) - h_m \nabla_x^m f(x) + \frac{h^2}{2}\Delta_x^m f(x). \tag{4.2}$$

This is the result of a simple calculation, which is however useful as it allows one to extend the derivation in the previous section by making the following replacements:

$$v'(x) \to \nabla_x^m v(x), \quad v''(x) \to \Delta^m v(x), \quad v'''(x) \to 0 \tag{4.3}$$

$$\mu'(x) \to \nabla_x^m \mu(x), \quad \mu''(x) \to \Delta_x^m \mu(x). \tag{4.4}$$

In fact,

$$\bar{\mathcal{L}}(x;h) = \frac{1}{h^2}\bar{\mathcal{L}}_0(x) + \frac{1}{h}\bar{\mathcal{L}}_1(x) + \bar{\mathcal{L}}_2(x) + h\bar{\mathcal{L}}_3(x) \tag{4.5}$$

without any $O(h^3)$ corrections as long as one re-defines the matrices on the right hand side as follows:

$$\bar{\mathcal{L}}_0(x) = \begin{pmatrix} -v(x) & \frac{1}{2}v(x) & 0 \\ \frac{1}{2}v(x) & -v(x) & \frac{1}{2}v(x) \\ 0 & \frac{1}{2}v(x) & -v(x) \end{pmatrix}, \tag{4.6}$$

$$\bar{\mathcal{L}}_1(x) = \begin{pmatrix} -\nabla_x^m v(x) & \frac{1}{2}\nabla_x^m v(x) - \frac{1}{2}\mu(x) & 0 \\ \frac{1}{2}\mu(x) & 0 & -\frac{1}{2}\mu(x) \\ 0 & -\frac{1}{2}\nabla_x^m v(x) + \frac{1}{2}\mu(x) & \nabla_x^m v(x) \end{pmatrix}, \tag{4.7}$$

$$\bar{\mathcal{L}}_2(x) = \begin{pmatrix} -\frac{1}{2}\Delta_x^m v(x) & \frac{1}{4}\Delta_x^m v(x) - \frac{1}{2}\nabla_x^m \mu(x) & 0 \\ 0 & 0 & 0 \\ 0 & \frac{1}{4}\Delta_x^m v(x) - \frac{1}{2}\nabla_x^m \mu(x) & -\frac{1}{2}\Delta_x^m v(x) \end{pmatrix}. \tag{4.8}$$

and

$$\bar{\mathcal{L}}_3(x) = \begin{pmatrix} 0 & -\frac{1}{4}\Delta_x^m \mu(x) & 0 \\ 0 & 0 & 0 \\ 0 & \frac{1}{4}\Delta_x^m \mu(x) & 0 \end{pmatrix}. \tag{4.9}$$

All derivations in the previous section go through formally unchanged and one arrives at the following expressions

$$\hat{\phi}_0(\omega, x, \tau) = 1 + \frac{2\mu(x)\tau h}{v(x)} - \frac{4i\omega h^2}{v(x)} - 8\mu(x)\frac{i\omega\tau h^3}{v(x)^2} - \frac{16\omega^2 h^4}{v(x)^2} + O(h^5). \tag{4.10}$$

and

$$\begin{aligned}\hat{\phi}_1(\omega, x, \tau) = 1 + \frac{2\mu(x)\tau h}{v(x)} - \frac{4i\omega h^2}{v(x)} - \left[8\mu(x) - \nabla_x^m v(x)\right]\frac{i\omega\tau h^3}{v(x)^2} \\ + r(x)\cdot h^3\tau + i\omega h^4 p(x) - \frac{14\omega^2 h^4}{v(x)^2} + O(h^5)\end{aligned} \tag{4.11}$$

where

$$
\begin{aligned}
r(x) = \frac{1}{2v(x)^3}\Big[ & v(x)^2 \Delta_x^m m(x) - 4m(x)^3 + 2\nabla_x^m v(x) m(x)^2 - 2v(x)\nabla_x^m v(x) \nabla_x^m m(x) \\
& - \big(2m(x)\nabla_x^m m(x) + v(x)\Delta_x^m v(x) - 2\big(\nabla_x^m v(x)\big)^2\big) m(x)\Big]. \\
p(x) = \frac{1}{v(x)^3}\Big[ & -4m(x)^2 + 2m(x)\nabla_x^m v(x) - 4v(x)\nabla_x m(x) - v(x)\Delta_x^m v(x) + 2\big(\nabla_x^m v(x)\big)^2\Big].
\end{aligned} \tag{4.12}
$$

We have that

$$
\begin{aligned}
&\Big| \sum_{j=0}^{q-1} \Big( \log \hat\phi_0(\omega;\gamma_j, \mathrm{sgn}(\gamma_{j+1}-\gamma_j)) - \log \hat\phi_1(\omega;\gamma_j, \mathrm{sgn}(\gamma_{j+1}-\gamma_j)) \Big) \Big| \\
&\le \Big| \sum_{j=0}^{q-1} \Big( \frac{i\omega \nabla_{\gamma_j}^m v(\gamma_j)}{v(\gamma_j)^2} + r(\gamma_j) \Big) h^3 \mathrm{sgn}(\gamma_{j+1}-\gamma_j) \Big| + \big(|\omega| ||p||_\infty + 2|\omega|^2 ||v^{-2}||_\infty\big) O(h^4 q). \\
&\le 2Lh^2 \sup_{x\in A_m} \Big| \frac{i\omega\nabla_x^m v(x)}{v(x)^2} + r(x) \Big| + \big(|\omega| ||p||_\infty + 2|\omega|^2 ||v^{-2}||_\infty\big) O(h^4 q) \le c h^\gamma.
\end{aligned} \tag{4.13}
$$

where in the last step we made use of the Hölder continuity assumptions of Theorem 1. The other bounds staying the same, we arrive at

$$
\Big| \int_{\mathcal{C}_-} \Big( 2G_{m+1}(x,y;\omega) - G_m(x,y;\omega) \Big) e^{i\omega t} \frac{d\omega}{2\pi} \Big| \le c \frac{\sqrt{q_{\max}}}{h} h^{2+\gamma} \le c h^\gamma. \tag{4.14}
$$

Under the weaker assumption of Theorem 2, the bound that applies is instead

$$
\Big| \int_{\mathcal{C}_-} \Big( 2G_{m+1}(x,y;\omega) - G_m(x,y;\omega) \Big) e^{i\omega t} \frac{d\omega}{2\pi} \Big| \le c \frac{\sqrt{q_{\max}}}{h} h^{2}\rho(h) \le c \rho(h). \tag{4.15}
$$

Similar bounds also extend to the case of the first time derivative, since multiplication by a factor $i\omega$ inside of the contour integral is immaterial as far as establishing a bound of this sort is concerned. This completes the proof of Theorem 1 and Theorem 2.

## 5. Explicit Euler Scheme

In this section we prove Theorem 3. A Dyson expansion can also be obtained for the time-discretized kernel and has the form

$$
\begin{aligned}
u_m^{\delta t}(y_1, y_2; t) = & \frac{1}{h_m} \sum_{q=1}^{\infty} \sum_{\gamma\in\Gamma_m:\gamma_0=y_1,\gamma_q=y_2} \sum_{k_1=1}^{N} \sum_{k_2=k_1+1}^{N} \cdots \sum_{k_q=k_{q-1}+1}^{N} \\
& \Big(1+\delta t \mathcal{L}_m(\gamma_0,\gamma_0)\Big)^{k_1-1} (\delta t)^q \prod_{j=1}^{q} \mathcal{L}_m(\gamma_{j-1},\gamma_j) \Big(1+\delta t\mathcal{L}_m(\gamma_j,\gamma_j)\Big)^{k_{j+1}-k_j-1}
\end{aligned} \tag{5.1}
$$

where $t_{q+1} = t$ and $k_{q+1} = N$. In this case, the propagator can be expressed through a Fourier integral as follows:

$$
u_m^{\delta t}(y_1,y_2;t) = \int_{-\frac{\pi}{\delta t}}^{\frac{\pi}{\delta t}} G_m^{\delta t}(y_1,y_2;\omega) e^{i\omega t} \frac{d\omega}{2\pi} \tag{5.2}
$$

where

$$
G_m^{\delta t}(y_1,y_2;\omega) = \delta t \sum_{j=0}^{\infty} u_m^{\delta t}(y_1,y_2; j\delta t) e^{-i\omega j \delta t}. \tag{5.3}
$$

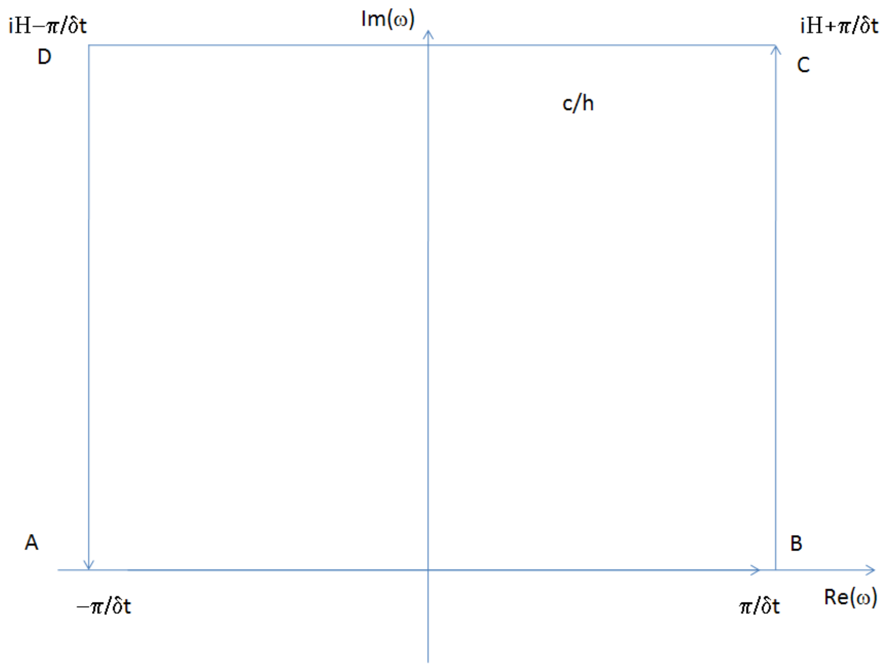


FIGURE 3. Contour of integration for the integral in (5.4).

The propagator can also be represented as the limit

$$u_m^{\delta t}(y_1, y_2; t) = \lim_{H\to\infty} \int_{\mathcal{C}_H} G_m^{\delta t}(y_1, y_2; \omega) e^{i\omega t} \frac{d\omega}{2\pi} \tag{5.4}$$

where $\mathcal{C}_H$ is the contour in Fig. 3. This is due to the fact that the integral along the segments $BC$ and $DA$ are the negative of each other, while the integral over $CD$ tends to zero exponentially fast as $\Im(\omega) \to \infty$, where $\Im(\omega)$ is the imaginary part of $\omega$. Using Cauchy's theorem, the contour in Fig. 3 can be deformed into the contour in Fig. 1. To estimate the discrepancy between the time-discretized kernel and the continuous time one, one can thus compare the Green's function along such contour. Again, the only arc that requires detailed attention is the arc $BCD$, as the integral over rest of the contour of integration can be bounded from above as in the previous section.

Let $h = h_m$ and let us introduce the two functions

$$\phi_0(t, x, \tau) \equiv 2\mathcal{L}_m(x, x + \tau h) e^{t\mathcal{L}_m(x,x)} 1(t \ge 0), \tag{5.5}$$

$$\phi_{\delta t}(j, x, \tau) \equiv 2\mathcal{L}_m(x, x + \tau h)\big(1 + \delta t \mathcal{L}_m(x, x)\big)^{j-1}. \tag{5.6}$$

and the corresponding Fourier transforms

$$\hat\phi_0(\omega, x, \tau) = \int_0^\infty \phi_0(t, x, \tau) e^{-i\omega t} \frac{d\omega}{2\pi} = \left(\frac{v(x)}{h^2} + \tau \frac{\mu(x)}{h}\right)\left(\frac{v(x)}{h^2} + i\omega\right)^{-1} \tag{5.7}$$

$$\hat\phi_{\delta t}(\omega, x, \tau) = \sum_{j=0}^{\frac{t}{\delta t}} \phi_{\delta t}(j, x, \tau) e^{-i\omega j \delta t} = \left(\frac{v(x)}{h^2} + \tau \frac{\mu(x)}{h}\right)\left(e^{i\omega\delta t} - 1 + \delta t \frac{v(x)}{h^2}\right)^{-1}. \tag{5.8}$$

We have that

$$
\begin{aligned}
\hat{\phi}_{\delta t}(\omega, x, \tau) &= \left(\frac{v(x)}{h^2} + \tau\frac{\mu(x)}{h}\right)\left(i\omega + \frac{v(x)}{h^2} - \frac{\omega^2}{2}\delta t + O(\delta t^2)\right)^{-1} \\
&= \hat{\phi}_0(\omega, x, \tau) + \frac{\omega^2}{2v(x)}h^2\delta t + O(h^2\delta t^2). = \hat{\phi}_0(\omega, x, \tau) + O(h^4), \qquad (5.9)
\end{aligned}
$$

where the last step uses the fact that $\delta t = O(h^2)$.

Let us also introduce the functions

$$
\psi_0(t, x, \tau) \equiv e^{t\mathcal{L}_m(x,x)}1(t \geq 0), \qquad \psi_{\delta t}(j, x, \tau) \equiv \sum_{k=1}^{j}\left(1 + \delta t\mathcal{L}_m(x, x)\right)^{j-1}. \tag{5.10}
$$

and the corresponding Fourier transforms

$$
\hat{\psi}_0(\omega, x, \tau) = \left(\frac{v(x)}{h^2} + i\omega\right)^{-1}, \qquad \hat{\psi}_{\delta t}(\omega, x, \tau) = \left(e^{i\omega\delta t} - 1 + \delta t\frac{v(x)}{h^2}\right)^{-1}. \tag{5.11}
$$

Again we find that

$$
\hat{\psi}_0(\omega, x, \tau) = \hat{\psi}_{\delta t}(\omega, x, \tau) + O(h^4). \tag{5.12}
$$

If $\gamma$ is a symbolic sequence, then let us set

$$
\hat{W}_m(\gamma, q; \omega) = \hat{\psi}_0(\omega, \gamma_q)\prod_{j=0}^{q-1}\hat{\phi}_0(\omega; \gamma_j, \mathrm{sgn}(\gamma_{j+1} - \gamma_j)) \tag{5.13}
$$

$$
\hat{W}_m^{\delta t}(\gamma, q; \omega) = \hat{\psi}_{\delta t}(\omega, \gamma_q)\prod_{j=0}^{q-1}\hat{\phi}_{\delta t}(\omega; \gamma_j, \mathrm{sgn}(\gamma_{j+1} - \gamma_j)). \tag{5.14}
$$

We have that

(5.15)

$$
G_m^{\delta t}(x, y; \omega) - G_m(x, y; \omega) = \frac{1}{h}\sum_{q=1}^{\infty}2^{-q}\sum_{\substack{\gamma \in \Gamma_m : \gamma_0 = x, \gamma_q = y \\ |\gamma_j - \gamma_{j-1}| = 1 \forall j \geq 1}}\left(\hat{W}_m^{\delta t}(\gamma, q; \omega) - \hat{W}_m(\gamma, q; \omega)\right).
$$

The integration over the contour in Fig. 1 can again be split into an integration over the countour $\mathcal{C}_-$ and an integration over $\mathcal{C}_+$. The integral over $\mathcal{C}_+$ can be bounded from above thanks to Lemma 1. Furthermore, we have that

$$
\begin{aligned}
&\left|\int_{\mathcal{C}_-}\left(G_m^{\delta t}(x, y; \omega) - G_m(x, y; \omega)\right)e^{i\omega t}\frac{d\omega}{2\pi}\right| \\
&\leq ch^{-1}\sqrt{q_{\max}}\max_{\substack{q, \gamma \in \Gamma_m : \gamma_0 = x, \gamma_q = y \\ |\gamma_j - \gamma_{j-1}| = 1 \forall j \geq 1}}\left|\int_{\mathcal{C}_-}\left(\hat{W}_m^{\delta t}(\gamma, q; \omega) - \hat{W}_m(\gamma, q; \omega)\right)e^{i\omega t}\frac{d\omega}{2\pi}\right|. \\
&\leq ch^2 \qquad (5.16)
\end{aligned}
$$

To bound the time derivative, we have to consider

$$
\left|\int_{\mathcal{C}_-}\left(\frac{e^{i\omega\delta t} - 1}{\delta t}G_m^{\delta t}(x, y; \omega) - i\omega G_m(x, y; \omega)\right)e^{i\omega t}\frac{d\omega}{2\pi}\right|
$$

(5.17)

But, since $\delta t = O(h^2)$, also this difference is $O(h^2)$.

## 6. Conclusions

We obtained bounds on convergence rates for explicit discretization schemes to the kernel of one-dimensional diffusion equations with continuous coefficients. We consider both semidiscrete triangulations with continuous time and explicit Euler schemes with time step small enough for the method to be stable. The proof is constructive and based on a new technique of path conditioning for Markov chains and a renormalization group argument. Convergence rates depend on the degree of smoothness and Hölder differentiability of the coefficients. The method is of more general applicability and will be extended in future work.

## References


Albanese, C. (2006). Stochastic Integrals and Abelian Processes. *preprint*.

Albanese, C. and A. Mijatovic (2006). Convergence Rates for Diffusions on Continuous-Time Lattices. *preprint, available at www.level3finance.com*.

Baker, G., J. Bramble and V. Thomee (1977). Single Step Galerkin Approximations for Parabolic Problems. *Math. Comp.* **31**, 818–847.

Courant, R., K. Friedrichs and H. Lewy (1928). ber die Partiellen Differenzengleichungen der Mathematischen Physik. *Mathematische Annalen*.

Crank, J. and P. Nicolson (1947). A Practical Method for Numerical Evaluation of Solutions of Partial Differential Equations of the Heat Conduction Type. *Proc. Camb. Phil. Soc.*

Ethier, S.N. and T.G. Kurtz (1986). *Markov processes: Characterization and convergence*. John Wiley and Sons, New York.

Fabes, E. B. and N. M. Riviere (1966). Parabolic Partial Differential Equations with Uniformly Continuous Coefficients. *Bull. Amer. Math. Soc.* **72**, 116–117.

Feller, W. (1936). Zur theorie der stochastischen prozesse [existenz und eindeutigkeitssatze]. *Math. Ann.* **113**, 113–160.

Feller, W. (1952). The parabolic differential equations and tke associated semi-groups of transformutions. *Ann. of Math.* **55**, 468–519.

Friedrichs, K. O. (1958). Symmetric positive linear differential equations. *Cornm. Pure Appl. Math.* **11**, 333–418.

Hille, E. (1948). Functional analysis and semi-groups. *Amer. Math. Soc. Colloquium Publications*.

Ito, K. (1957). Fundamental solutions of parabolic differential equations and boundary value problems. *Jap. J. Math.* **27**, 55–102.

Kato, T. (1966). *perturbation Theory for Linear Operators*. Springer, New York.

Kolmogorov, A. N. (1931). Uber die analytischen methoden in der wahrscheinlichkeitsrechnung. *Math. Ann.* **104**, 415–458.

Kurtz, T. (1969). Extensions of Trotter's operator semigroup approximation theorems. *J. Funct. Anal.* **3**, 354–375.

Lax, P. D. and R. S. Phillips (1960). Local boundary conditions for dissipative symmetric linear differential operators. *Cornm. Pure Appl. Math.* **13**, 427–455.

Luskin, M. and R. Rannacher (1978). On the Smoothing Property of the Galerkin Method for Parabolic Equations. *Proc. Japan Acad. Ser. A Math. Sci.* **54**, 326–331.

Mingyou, H. and V. Thomee (1982). On the Backward Euler Method for Parabolic Equations with Rough Initial Data. *SIAM J. Numer. Anal* **19**, 599–603.

Palencia, C. (1996). Maximum Norm Analysis of Completely Discrete Finite Element Methods for Parabolic Problems. *SIAM Journal on Numerical Analysis* **33**, 1654–1668.

Philips, R. S. (1961). On the integration of the diffusion equation with boundary conditions. *Transactions of the American Mathematical Society* **98**, 62–84.

Sova, M. (1967). Convergence d'operations lineaires non bornees. *Rev. Roumaine Math. Pures Appliq.* **12**, 373–389.

Stroock, D.W. and S.R.S. Varadhan (1969). Diffusion processes with continuous coefficients, i and ii.. *Comm. Pure Appl. Math.* **22**, 345–400, 479–530.

Stroock, D.W. and S.R.S. Varadhan (1979). *Multidimensional diffusion processes*. Springer-Verlag, Berlin.

Suzuki, T. (1978). On the Rate of Convergence of the Difference Finite Element Approximation for Parabolic Equations. *Proc. Japan Acad. Ser. A Math. Sci.* **54**, 326–331.

Yosida, K. (1951). Integration of fokker-planck's equation with a boundary condition. *J. Math. Soc. Japan* **3**, 69–73.