

Finite Element Methods for Parabolic Stochastic PDE's

John B. Walsh

Department of Mathematics, University of British Columbia
Vancouver, B.C. V6T 1Y4
walsh@math.ubc.ca

Abstract

We study the rate of convergence of some explicit and implicit numerical schemes for the solution of a parabolic stochastic partial differential equation driven by white noise. These include the forward and backward Euler and the Crank-Nicholson schemes. We use the finite element method. We find, as expected, that the rates of convergence are substantially similar to those found for finite difference schemes, at least when the size of the time step k is on the order of the square of the size of the space step h : all the schemes considered converge at a rate on the order of $h^{1/2} + k^{1/4}$, which is known to be optimal. We also consider cases where k is much greater than h^2 , and find that only the backward Euler method always attains the optimal rate; other schemes, even though they are stable, can fail to convergence to the true solution if the time step is too long relative to the space step. The Crank-Nicholson scheme behaves particularly badly in this case, even though it is a higher-order method.

KEYWORDS: Stochastic partial differential equations, numerical SPDE, rate of convergence, finite elements, white noise.

1 Introduction

Let $L > 0$ and set $L_0 = [0, L]$. Let \dot{W} be a standard white noise on $L_0 \times \mathbb{R}_+$, let f and g be real-valued Lipschitz continuous functions on \mathbb{R} , let u_0 be a continuous function on L_0 , and consider the stochastic initial-boundary value problem

$$(1) \quad \begin{cases} \frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial x^2} + f(U)\dot{W} + g(U), & (x, t) \in L_0 \times \mathbb{R}_+ \\ U(x, 0) = u_0(x), & x \in L_0, \\ U(0, t) = U(L, t) = 0, & t > 0 \end{cases}$$

The solutions of (1) are continuous but non-differentiable functions. Since the derivatives don't exist, (1) should just be regarded as shorthand for an integral equation. If ϕ is a C^∞ function in L_0 , set $U(\phi, t) = \int_{L_0} U(y, t) \phi(y) dy$. Then the *weak form* of (1) is

$$(2) \quad U(\phi, t) = u_0(\phi) + \int_0^t U(\phi'', s) ds + \int_0^t \int_{L_0} f(U(y, s)) \phi(y) W(dy ds) \\ + \int_0^t \int_{L_0} g(U(y, s)) \phi(y) dy ds, \quad \forall \phi \in C_0^\infty(L_0),$$

where $C_0^\infty(L_0)$ is the space of infinitely differentiable functions on L_0 which vanish at 0 and L . There is a unique process $\{U(x, t), x \in L_0, t \geq 0\}$ which is continuous in the pair (x, t) and satisfies (2) identically with probability one. We say informally that this is also the solution of (1). Equation (2) is equivalent [??] to the *mild form*

$$(3) \quad U(x, t) = \int_{L_0} G_t(x, y) u_0(y) dy + \int_0^t \int_{L_0} f(U(y, s)) G_{t-s}(x, y) W(dy ds) \\ + \int_0^t \int_{L_0} g(U(y, s)) G_{t-s}(x, y) dy ds.$$

where $G_t(x, y)$ is the Green's function or fundamental solution for the homogeneous equation $\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial x^2}$ with boundary conditions $U(t, 0) = U(t, L) = 0$ for all $t > 0$.

Most numerical work on SPDE's has concentrated on the Euler finite-difference scheme. Gyöngi and Nualart [3] have proved that these schemes converge, and Gyöngi [2] has determined the order of convergence. Davie and Gaines [1], examining a much larger class of schemes, have found a universal lower bound for the rate of convergence: if h is the space step and k is the time step, then any scheme, implicit or explicit, will have an error at least $O(h^{\frac{1}{2}} + k^{\frac{1}{4}})$. This is the bound given by Gyöngi, so the simple Euler scheme achieves the optimal rate of convergence. (For the Euler scheme, k must be smaller than a constant times h^2 , so that one can say that the scheme is roughly of order one-fourth.)

We should mention also that Davie and Gaines [1] have shown that if one wishes to compute the *expectation* of a functional of $U(x, t)$ rather than $U(x, t)$ itself, one can get faster convergence.

We will use the finite element method, and we will consider a family of semi-implicit methods, called *one-step theta schemes*. These depend on a parameter θ , $0 \leq \theta \leq 1$, and include the Euler ($\theta = 0$), the Crank-Nicholson ($\theta = \frac{1}{2}$), and the backward Euler ($\theta = 1$).

There is no a priori reason to expect finite elements to have any great advantages over finite differences—on the contrary, one expects the two to give similar results, particularly in one space dimension, and one expects that our results for the one-step theta schemes have their counterparts in finite difference schemes. Indeed, Gyöngy [2], has found results analogous to ours for one-step theta finite difference schemes.

We have limited ourself to one space dimension. The finite element method is particularly suited for equations in domains in higher dimensions. However the solutions of the corresponding equations—e.g. the heat equation driven by white noise in two or more

space dimensions—are only generalized functions, i.e. Schwartz distributions, so it is not easy to see how numerical solutions of any kind fit in. It would be interesting to see how finite element solutions behave when the white noise is replaced by something smooth enough to give, say, continuous solutions.

The numerical schemes are summarized in Section 4.4. The main results are stated in Section 2 and proved in Section 5. The main tools in the proof are moment estimates on the Green’s functions. These are given in Section 8. These methods may not give the sharpest results in the classical case, but they are well-adapted to stochastic equations, where the biggest errors come from stochastic integrals, which are tightly bound up with L^2 .

All these schemes are most efficient when the time step k and the square of the space step h^2 are roughly of the same magnitude; in that case, they all attain the optimal order of convergence. They behave differently if k is large compared to h^2 . The schemes for $\theta \geq 1/2$ are stable for any value of k , but this is not sufficient to guarantee convergence. Indeed, if $1/2 < \theta < 1$, the error is $O(h^2 + k/h)$, and we show that if $k \geq h$, the schemes need not converge. The Crank-Nicholson scheme is worse: its error is $O(h^{1/2} + k/h^{3/2})$, and one needs $k = o(h^{3/2})$ to be sure the scheme converges to the true solution. The backwards Euler, on the other hand, attains the optimal rate of convergence for all h and k .

These convergence rates may seem slow, but the solutions are only Hölder(1/2) in space and Hölder(1/4) in time [6], so that the convergence rate is perhaps not too surprising. For a similar reason, higher-order methods are not significantly faster than first order methods, for they depend on the smoothness of the solution, and the solutions in this case are non-differentiable. In fact, they can even be slower, as we can see from the Crank-Nicholson scheme. Since it is a second order method, one might think it should be faster, but the opposite is true for large k .

The intuitive reason for this is that the higher modes do not decay quickly in the Crank-Nicholson scheme, and these modes are heavily-weighted in the stochastic solution. The backward Euler, on the other hand, damps out higher modes very quickly, and gives the optimum rate of convergence for any h and k . We have not computed exact constants for the errors, so we can not compare the actual speed of convergence for all the different values of theta, but we would not expect the Crank-Nicholson method to be the best choice.

We can draw some conclusions from this. First, there is little to be gained, and perhaps something to be lost, by going to higher order methods: first order methods already attain the optimal order of convergence. Second, since the rate of convergence is roughly the maximum of $h^{1/2}$ and $k^{1/4}$, there is little to be gained by taking a time-step k much greater than h^2 . Put another way, for a given time step k , there is small profit in taking h very much less than \sqrt{k} ; it costs computation time and adds little to the accuracy. Indeed, for $1/2 \leq \theta < 1$ we have the paradoxical situation that there is a point at which refining the space grid actually increases the error. As a practical matter, then, one should choose h

and k so that h^2 and k are roughly of the same magnitude.

One thing which emerges is that the greatest error comes at the end. That is, if we want to approximate the solution at a particular time T , the greatest contribution to the error comes in the final steps just before T . Indeed, the Laplacian has smoothing effect, and it tends to average out errors from earlier steps. Our results assume uniform step sizes, which is good if we wish to find a uniform estimate for the solution. However, if we just want to find the value at a particular time T , we might get a more accurate estimate with a variable-mesh method: by refining the mesh as we approach T , we might reduce the error due to the most recent steps.

There is one unstated assumption here: if the space-step is h and the time-step is k , then the schemes only use the white noise generated on the grid rectangles of sides h and k . This is implicit in the derivation of the optimum rate by Davie and Gaines [1], who calculate the expected error given the values of white noise on these rectangles. One could compute the stochastic integrals more accurately by using white noise generated on a finer grid, but this would be nearly the same amount of work as computing the whole scheme on the finer grid. Effectively, it would be implicitly using the smaller grid size.

We should emphasize here that we are talking only about the *order* of convergence; one might well be able to improve the *constants* in the exact rate. We have looked at one possible refinement in section 7, namely the inclusion of a quasi-Milstein correction term. This term increases the rate of convergence for ordinary SDEs, but not for SPDEs, or even systems of SDEs. However, it may possibly give a better constant in the convergence rate.

2 The main result

Theorem 2.1 *Let f and g be Lipschitz continuous functions on the line and let U be the solution of (2). Let \tilde{U}^θ be the numerical approximation given by the one-step theta scheme (4.4) with space step h and time-step k . Suppose u_0 is smooth.*

(i) *For all $T > 0$ there is a constant C_T such that if either $\frac{1}{2} \leq \theta \leq 1$ or if $\theta = 0$ and $k < (1 - \varepsilon)h^2/6$ for some $\varepsilon > 0$,*

$$(4) \quad \sup_{x \in L_0, t \leq T} E\{(U_t(x) - \tilde{U}_t^\theta(x))^2\}^{\frac{1}{2}} \leq \begin{cases} C_T h^{\frac{1}{2}} & \text{if } \theta = 0 \\ C_T \left(h^{\frac{1}{2}} + \left(\frac{k^2}{h^3} \right)^{\frac{1}{2}} \right) & \text{if } \theta = \frac{1}{2} \\ C_T \left(h^{\frac{1}{2}} + \left(\frac{k}{h} \right)^{\frac{1}{2}} \right) & \text{if } \frac{1}{2} < \theta < 1 \\ C_T \left(h^{\frac{1}{2}} + k^{\frac{1}{4}} \right) & \text{if } \theta = 1. \end{cases}$$

(ii) *Suppose that $0 \leq \theta \leq 1$, that $h \sim 2^{-m}L$, $m = 1, 2, \dots$ and that k/h^2 is bounded. Then for any $x \in L_0$, $t \leq T$, and $\varepsilon > 0$*

$$(5) \quad \frac{|U(x, t) - \tilde{U}(x, t)|}{h^{\frac{1}{2}-\varepsilon}} \longrightarrow 0 \quad a.e.$$

3 The Discretization in Space

We will construct our numerical scheme in two stages: we first discretize in space using the finite element method, and then discretize in time. The spatial discretization leads to a system of stochastic ODE's.

Divide the interval $L_0 = [0, L]$ into $n + 1$ sub-intervals by the points $0, h, \dots, nh, L$, where $h = \frac{L}{n+1}$. The *finite element functions* ϕ_1, \dots, ϕ_n are the piecewise-linear tent-functions defined by $\phi_j(jh) = 1$, $\phi_j(x) = 0$ if $|x - jh| \geq h$, and ϕ_j is linear on the intervals $[(j-1)h, jh]$ and $[jh, (j+1)h]$. In short, $\phi_j(x) = (1 - |\frac{(n+1)x}{L} - j|)^+$. (So each ϕ_j vanishes except in $(jh - h, jh + h)$, and each ϕ_j satisfies the boundary conditions.)

Let $\mathcal{S}_n \subset L^2(L_0)$ be the space spanned by ϕ_1, \dots, ϕ_n and let

$$\langle f, g \rangle \stackrel{\text{def}}{=} \int_{L_0} f(y)g(y) dy$$

be the inner product in $L^2(L_0)$. Let us approximate the solution U of (1) by a process \hat{U} with values in \mathcal{S}_n , which we write

$$\hat{U}(x, t) = \sum_{j=1}^n a_j(t) \phi_j(x).$$

We adapt the weak form (2) to determine the coefficients a_j :

$$(6) \quad \begin{aligned} \langle \hat{U}_t, \phi \rangle &= \langle U_0, \phi \rangle - \int_0^t \left\langle \frac{\partial \hat{U}(\cdot, s)}{\partial x}, \phi' \right\rangle ds + \int_0^t \int_{L_0} f(\hat{U}(y, s)) \phi(y) W(dy ds) \\ &\quad + \int_0^t \int_{L_0} g(\hat{U}(y, s)) \phi(y) dy ds, \quad t \geq 0, \phi \in \mathcal{S}_n. \end{aligned}$$

We have used the facts that $\phi \in \mathcal{S}_n \implies \phi(0) = \phi(L) = 0$ and that ϕ' is a step function, so that ϕ' and $\frac{\partial \hat{U}}{\partial x}$ are defined a.e. and are bounded as functions of x . It is convenient to write this in vector notation. Let $\hat{A}(t)$ be the vector of coefficients, and $\Phi(x)$ the vector of finite-element functions:

$$\begin{aligned} \hat{A}(t) &= (a_1(t), \dots, a_n(t))^T \\ \Phi(x) &= (\phi_1(x), \dots, \phi_n(x))^T \\ \hat{U}(x, t) &= \hat{A}^T(t) \Phi(x) = \Phi^T(x) \hat{A}(t). \end{aligned}$$

Write (6) for $\phi = \phi_j$, $j = 1, \dots, n$ to see that

$$(7) \quad \langle \hat{U}_t, \Phi \rangle = \langle \hat{U}_0, \Phi \rangle - \int_0^t \left\langle \frac{\partial \hat{U}}{\partial x}, \Phi' \right\rangle ds + \int_0^t \int_{L_0} f(\hat{U}(y, s)) \Phi(y) W(dy ds) \\ + \int_0^t \int_{L_0} g(\hat{U}(y, s)) \Phi(y) dy ds.$$

Let $M = \langle \Phi, \Phi^T \rangle$, $N = \langle \Phi', \Phi'^T \rangle$. These are symmetric, tri-diagonal, positive definite $n \times n$ matrices. Note that $\hat{U}(t, x) = \Phi^T(x) \hat{A}(t)$, so

$$\hat{A}_t(t) = M^{-1} \langle \Phi(\cdot), \hat{U}(\cdot, t) \rangle.$$

Multiply (7) by M^{-1} and set $Q = M^{-1}N$, to get an integral equation for \hat{A} :

$$(8) \quad \hat{A}(t) = \hat{A}(0) - Q \int_0^t \hat{A}(s) ds + \int_0^t \int_{L_0} f(\hat{U}(y, s)) M^{-1} \Phi(y) W(dy ds) \\ + \int_0^t \int_{L_0} g(\hat{U}(y, s)) M^{-1} \Phi(y) dy ds.$$

Let us also derive the mild form of the equation for \hat{U} , which will introduce the finite element Green's function. Notice that the associated homogeneous equation is

$$(9) \quad \hat{A}(t) = \hat{A}(0) - Q \int_0^t \hat{A}(s) ds,$$

which has the solution $\hat{A}(t) = e^{-tQ} \hat{A}(0)$. As with the original equation, one can represent the solution of the inhomogeneous equation as

$$(10) \quad \hat{A}(t) = e^{-tQ} \hat{A}(0) + M^{-1} \int_0^t \int_{L_0} f((\hat{U}(y, s)) e^{-(t-s)Q} \Phi(y) W(dy ds) \\ + M^{-1} \int_0^t \int_{L_0} g((\hat{U}(y, s)) e^{-(t-s)Q} \Phi(y) dy ds.$$

Let $\hat{G}_t(x, y)$ be the *finite element Green's function*, defined by

$$(11) \quad \hat{G}_t(x, y) \stackrel{\text{def}}{=} \Phi^T(x) M^{-1} e^{-tQ} \Phi(y).$$

Multiply (10) by $\Phi^T(x)$, use the fact that $\hat{U}(x, t) = \Phi^T(x)\hat{A}(t)$, and note that Q and M^{-1} commute (they have the same eigenvectors F_j):

$$(12) \quad \hat{U}(x, t) = \int_{L_0} \hat{G}_t(x, y)u_0(y) dy + \int_0^t \int_{L_0} f(\hat{U}(y, s))\hat{G}_{t-s}(x, y) W(dy ds) \\ + \int_0^t \int_{L_0} g(\hat{U}(y, s))\hat{G}_{t-s}(x, y) dy ds,$$

which is the mild form of (7).

4 The Discretization in Time

4.1 The Time-Stepping Scheme

Let us use a tilde to distinguish the discrete-time quantities such as \tilde{U} and \tilde{A} from the continuous-time discrete-space quantities \hat{U} and \hat{A} . It is convenient to work with the coefficient process $\tilde{A}(t)$, since the solution $\tilde{U}(t, x)$ is easily derived from it. There are two stages. Let k be the time step, h the space step, where $h = L/(n + 1)$ for an integer n . We do not assume any relation between h and k for the moment. Let $t_0 = 0$, $t_1 = k$, $t_2 = 2k, \dots$. We will determine $\tilde{A}(t_0), \tilde{A}(t_1), \dots$.

We use (8) to update our scheme. There are three integrals to approximate; let D_j be a numerical approximation (to be determined below) of the second two integrals, i.e.

$$(13) \quad D_j \sim \int_{t_j}^{t_{j+1}} \int_{L_0} f(\hat{U}(y, s))M^{-1}\Phi(y)W(dy ds) + \int_{t_j}^{t_{j+1}} \int_{L_0} g(\hat{U}(y, s))M^{-1}\Phi(y) dy ds.$$

In order to approximate the third integral, $\int_{t_j}^{t_{j+1}} \hat{A}(s) ds$, we replace $\hat{A}(s)$ by $\hat{A}(t_j) + D_j$, leading to $\tilde{A}(t_{j+1}) = \tilde{A}(t_j) - kQ(\tilde{A}(t_j) + D_j) + D_j$ or, if we set $\Gamma_0 = I - kQ$,

$$(14) \quad \tilde{A}(t_{j+1}) = \Gamma_0 \left(\tilde{A}(t_j) + D_j \right).$$

Remark 4.1 The integral $\int_{t_j}^{t_{j+1}} \hat{A}(s) ds$ could be approximated by $k\hat{A}(t_j)$, but it could just as well be approximated by $k\hat{A}(t_{j+1})$. The latter would involve implicit methods if we used it as is, but if we make the preliminary approximation $\hat{A}(t_{j+1}) \sim \hat{A}(t_j) + D_j$ instead, we get the above explicit scheme. It is called *operator splitting*. Its practical effect is to minimize some unpleasantness which happens, even to stable schemes, if k is too large relative to h .

What we have described above is the forward Euler scheme. It is well known to be unstable unless the time step is less than a constant times the square of the space step. However, there are stable schemes which are closely related. Consider the *one-step theta schemes* for the homogeneous equation (i.e. with $D_j \equiv 0$.) Let $0 \leq \theta \leq 1$ and replace $\tilde{A}(t_j)$ on the right-hand side by the convex combination $(1 - \theta)\tilde{A}(t_j) + \theta\tilde{A}(t_{j+1})$. Then set

$$\tilde{A}(t_{j+1}) = \tilde{A}(t_j) - kQ \left((1 - \theta)\tilde{A}(t_j) + \theta\tilde{A}(t_{j+1}) \right) .$$

Put

$$(15) \quad \Gamma_\theta \stackrel{\text{def}}{=} I - kQ(I + \theta kQ)^{-1}$$

and note that for the homogeneous equation

$$(16) \quad \tilde{A}(t_{j+1}) = \Gamma_\theta \tilde{A}(t_j) .$$

For the inhomogeneous equation, put

$$(17) \quad \tilde{A}(t_{j+1}) \stackrel{\text{def}}{=} \Gamma_\theta \left(\tilde{A}(t_j) + D_j \right) .$$

This will be our updating scheme for \tilde{A} . Note it reduces to (14) if $\theta = 0$. If \tilde{A}_0 is given, then $\tilde{A}(t_1) = \Gamma_\theta(\tilde{A}_0 + D_0)$, so that by induction,

$$(18) \quad \tilde{A}(t_m) = \Gamma_\theta^m \tilde{A}(0) + \sum_{j=0}^{m-1} \Gamma_\theta^{m-j} D_j .$$

This, in its turn, leads to the formula for $\tilde{U} = \Phi^T \tilde{A}$:

$$(19) \quad \tilde{U}(x, t_m) = \langle u_0, \Gamma_\theta^m M^{-1} \Phi \rangle + \sum_{j=0}^{m-1} \Phi^T(x) \Gamma_\theta^{m-j} D_j \quad m = 0, 1, 2, \dots$$

The cases $\theta = 0$, $\theta = 1/2$ and $\theta = 1$ correspond respectively to the forward Euler, Crank-Nicholson, and backward Euler schemes. If $\theta \geq 1/2$, these are stable for any h and k ; if $\theta < 1/2$, they are unstable unless $k < h^2/3$.

Note that only the linear part is treated implicitly. The non-linear part, which is contained in the increments D_j , is treated explicitly. This is important, for a fully implicit scheme would force us to use the Stratonovich or some other type of non-adapted stochastic integral.

4.2 The Increment D_m .

Notice that D_m is the sum of a stochastic integral and a Riemann integral. Consider the stochastic integral first. Let $\Delta_{im} = [ih, (i+1)h) \times [t_m, t_{m+1})$ and write the (vector) stochastic integral as

$$\sum_{i=1}^n \int_{\Delta_{im}} f(\hat{U}(y, s)) M^{-1} \Phi(y) W(dy ds).$$

We need only consider the integrals over the Δ_{im} . At time t_m we will already have computed $\hat{U}(\cdot, t_m)$, and we will know the white-noise increments $W(\Delta_{im})$ as well. Now $t_{m+1} - t_m = k$, which will be small, so that we can replace $f(\hat{U}(y, s))$ by $f(\hat{U}(y, t_m))$, or, rather, by some average over $(ih, (i+1)h)$. For any function H , let

$$\overline{[H]}_i \stackrel{\text{def}}{=} \frac{1}{h} \int_{ih}^{(i+1)h} H(y) dy$$

be the average of H over $(ih, (i+1)h)$. Then we will approximate the stochastic integral over Δ_{im} by

$$\overline{[f(\hat{U}(\cdot, t_m))]}_i M^{-1} \overline{[\Phi]}_i W(\Delta_{im}).$$

Add in the same type of approximation for the Riemann integral:

$$(20) \quad D_m = \sum_i \left(\overline{[f(\tilde{U}(\cdot, t_m))]}_i W(\Delta_{im}) + \overline{[g(\tilde{U}(\cdot, t_m))]}_i hk \right) M^{-1} \overline{[\Phi]}_i.$$

Remark 4.2 This choice of D_m is convenient, and it is not hard to compute. Indeed, $x \mapsto \Phi(x)$ is affine on $(ih, (i+1)h)$, so that $\overline{[\Phi]}_i = \Phi((i+\frac{1}{2})h)$. In addition, $x \mapsto \tilde{U}(x, t_m)$ is also affine on the same interval, so if F is the anti-derivative of f , then

$$\overline{[f(\tilde{U}(\cdot, t_m))]}_i = \frac{F(\tilde{U}((i+1)h, t_m)) - F(\tilde{U}(ih, t_m))}{\tilde{U}((i+1)h, t_m) - \tilde{U}(ih, t_m)}.$$

We should point out that this is not the only choice. Instead of averages, one could evaluate all the functions at, say, the midpoint of the interval. A slightly better choice would be $M^{-1} \overline{[f(\tilde{U}(\cdot, t_m))\Phi]}_i W(\Delta_{im})$, for this is the conditional expectation of the integral $\int_{\Delta_{im}} f(\tilde{U}(y, t_m)) M^{-1} \Phi(y) W(dy ds)$ given $\mathcal{F}_{t_m} \times \sigma\{W(\Delta_{im})\}$. Again, the fact that both $U(y, t_m)$ and $\Phi(y)$ are affine on $(ih, (i+1)h)$ makes it possible to express this in terms of the integrals of $f(x)$ and $xf(x)$. However, both these choices—and indeed, almost all reasonable choices—lead to the same order of convergence.

4.3 The Initial Value

We will let the initial value of \tilde{U} be equal to the initial value of \hat{U} , which is the projection of u_0 on \mathcal{S}_n . (It might be more convenient to take it to be the polygonal interpolation of u_0 , which is easy to compute, since it equals $\sum_{j=1}^n u_0(ih)\phi_j(x)$. This would give a slightly larger expected L^2 error, but both lead to errors of the same order of magnitude.)

4.4 The Numerical Scheme Summarized

$$(21) \quad \tilde{U}(x, t_m) = \tilde{A}^T(t_m)\Phi(x);$$

$$(22) \quad \tilde{A}(0) = M^{-1}\langle \Phi u_0 \rangle;$$

$$(23) \quad D_m = \sum_{i=0}^n \left(\overline{[f(\hat{U}(\cdot, t_m))]_i} W(\Delta_{im}) + hk \overline{[g(\hat{U}(\cdot, t_m))]_i} \right) M^{-1} \overline{[\Phi]_i}$$

$$(24) \quad \tilde{A}(t_{m+1}) = \Gamma_\theta \left(\tilde{A}(t_m) + D_m \right) \quad m = 0, 1, 2, \dots,$$

where $M = \langle \Phi, \Phi^T \rangle$.

4.5 The Discrete Green's Function

We will translate (19) into its mild form. Let us define

$$\bar{\Phi}(x) \stackrel{\text{def}}{=} \overline{[\Phi]_i} \text{ if } x \in [ih, (i+1)h), \quad i = 0, \dots, n.$$

Notice that $\bar{\Phi}(x) = (\bar{\phi}_1(x), \dots, \bar{\phi}_n(x))^T$ where $\bar{\phi}(x) = \frac{1}{2}I_{(j-1)h, (j+1)h}(x)$. Similarly, define

$$(25) \quad \begin{cases} \tilde{f}(x, s) \stackrel{\text{def}}{=} \overline{[f(\tilde{U}(\cdot, t_j))]_i} & \text{if } s \in [t_j, t_{j+1}), \quad x \in [ih, (i+1)h), \quad i = 0, \dots, n, \\ \tilde{g}(x, s) \stackrel{\text{def}}{=} \overline{[g(\tilde{U}(\cdot, t_j))]_i} & \text{if } s \in [t_j, t_{j+1}), \quad x \in [ih, (i+1)h), \quad i = 0, \dots, n. \end{cases}$$

These are step functions—basically, they are conditional expectations given the partition of L_0 . We can write D_m in terms of these:

$$D_m = \int_{t_m}^{t_{m+1}} \int_{L_0} \tilde{f}(y, s) M^{-1} \bar{\Phi}(y) W(dy ds) + \int_{t_m}^{t_{m+1}} \int_{L_0} \tilde{g}(y, s) M^{-1} \bar{\Phi}(y) dy ds,$$

so that (19) can be written

$$\begin{aligned} \tilde{U}(x, t_m) &= \langle u_0, \Gamma_\theta^m M^{-1} \Phi \rangle + \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} \int_{L_0} \Phi^T(x) \Gamma_\theta^{m-j} \tilde{f}(y, s) M^{-1} \bar{\Phi}(y) W(dy ds) \\ &\quad + \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} \int_{L_0} \Phi^T(x) \Gamma_\theta^{m-j} \tilde{g}(y, s) M^{-1} \bar{\Phi}(y) dy ds. \end{aligned}$$

Let $[s]$ denote the greatest integer in s and set

$$(26) \quad \tilde{G}_t(x, y) = \Phi^T(x) M^{-1} \Gamma_{\theta}^{\left[\frac{t}{k}\right]} \bar{\Phi}(y), \quad t > 0, \quad x, y \in L_0.$$

We can rewrite this as

$$(27) \quad \begin{aligned} \tilde{U}(x, t_m) = & \int_{L_0} \tilde{G}_{t_m}(x, y) u_0(y) dy + \int_0^{t_m} \int_{L_0} \tilde{G}_{t+k-s}(x, y) \tilde{f}(y, s) W(dy ds) \\ & + \int_0^{t_m} \int_{L_0} \tilde{G}_{t+k-s}(x, y) \tilde{g}(y, s) dy ds. \end{aligned}$$

This is the mild form of the discrete-space discrete-time equation. Note that, unlike $G_t(x, t)$ and $\hat{G}_t(x, y)$, $\tilde{G}_t(x, y)$ is not symmetric in x and y .

5 The Proof of Theorem 2.1

We begin with a variation on Gronwall's inequality.

Lemma 5.1 *Let $\gamma(t) = 1/\sqrt{t}$ and let $c > 0$ be a constant. Let $H(t)$ be a positive function which is integrable on each interval $[0, T]$, and let F be a positive function which satisfies*

$$(28) \quad F(t) \leq H(t) + c\gamma * F(t), \quad t > 0.$$

*Then there exists a function $K_c(t)$ on $(0, \infty)$ such that, if we define $f(t) \equiv H(t) + K_c * H(t)$, $t > 0$ then*

(i) $f(t)$ is the unique solution of

$$(29) \quad f(t) = H(t) + c\gamma * f(t), \quad t > 0;$$

(ii) $F(t) \leq f(t)$, $t > 0$;

(iii) $K_c - \gamma \in C^\infty[0, \infty)$;

(iv) $K_c(t) = c\gamma(t) + \pi c^2 e^{\pi c^2 t} (1 + \operatorname{erf}(c\sqrt{\pi t})) \leq c\gamma + 2\pi c^2 e^{\pi c^2 t}$.

PROOF. Define the convolution powers of $c\gamma$ by $(c\gamma)^{*1} = c\gamma$, $(c\gamma)^{*2} = c^2\gamma * \gamma$ and, by induction, $(c\gamma)^{*n} = c\gamma * (c\gamma)^{*(n-1)}$. Define $K_c(t) = \sum_{j=1}^{\infty} (c\gamma)^{*j}$. This converges; in fact, we can determine it exactly: note first that $\gamma * \gamma \equiv \pi$. The convolution powers of a constant are easy: $(c\gamma)^{*2n} = \pi^n c^{2n} t^{n-1} / (n-1)!$. Thus the even powers can be summed:

$$\sum_{j=1}^{\infty} (c\gamma)^{*2j} = \pi c^2 e^{\pi c^2 t}.$$

The sum of the odd terms is

$$c\gamma + c\gamma * (\pi c^2 e^{\pi c^2 t}) = c\gamma + \pi c^2 e^{\pi c^2 t} \operatorname{erf}(c\sqrt{\pi t})$$

proving (iv), and (iii), follows since the error function is bounded by one. Since H is locally integrable, $K_c * H$ is defined and finite, and $f \equiv H + K_c * H$ is easily seen to solve (29). To see that f is the unique solution, suppose that f_1 and f_2 are both locally integrable functions which satisfy (29). Then

$$\begin{aligned} f_2 - f_1 &= \gamma * (f_2 - f_1) \\ &= \gamma^{*2} * (f_2 - f_1) \\ &= \dots \\ &= \gamma^{*n} * (f_2 - f_1) \longrightarrow 0. \end{aligned}$$

Thus $f_2 = f_1$ and the solution is unique. Now from (28),

$$\begin{aligned} F(t) &\leq H(t) + c\gamma * F(t) \\ &\leq H(t) + c\gamma * H(t) + (c\gamma)^{*2} * F(t) \\ &\leq \dots \\ &\leq H(t) + \sum_{j=1}^{n-1} (c\gamma)^{*j} * H(t) + (c\gamma)^{*n} * F(t). \end{aligned}$$

But $\gamma^{*j} \longrightarrow 0$ uniformly on compacts, so $\gamma^{*j} * H \longrightarrow 0$. Thus we can sum the series inside the convolution to see that this is

$$\leq H(t) + K_c * H(t) \equiv f(t).$$

♣

Proposition 5.2 (i) $E\{U(x, t)^2\}$ and $E\{\tilde{U}(x, t)^2\}$ are bounded in $[0, T] \times L_0$, independent of h and k .

(ii) If u_0 is smooth, there is a constant C_T such that if $t \leq T$, $i = 0, 1, \dots, n-1$,

$$(30) \quad E \left\{ \left(\tilde{U}(x_{i+1}, t) - \tilde{U}(x_i, t) \right)^2 \right\} \leq \begin{cases} C_T \left(h + \frac{k}{h} \right) & \text{if } \theta \neq \frac{1}{2} \\ C_T \left(h + \frac{k^2}{h^3} \right) & \text{if } \theta = \frac{1}{2}. \end{cases}$$

PROOF. Let $\bar{U}(x, t)$ be $U(x, t)$ (resp. $\tilde{U}(x, t)$) and let $\bar{G}_t(x, y)$ be $G_t(x, y)$ (resp. $\tilde{G}_t(x, y)$). Then from (3) and (27),

$$\begin{aligned} \bar{U}(x, t) = \int_{L_0} \bar{U}(y, 0) \bar{G}_t(x, y) dy + \int_0^t \int_{L_0} \bar{G}_t(x, y) \bar{f}(y, s) W(dy ds) \\ + \int_0^t \int_{L_0} \bar{G}_t(x, y) g(\bar{U}(y, s)) dy ds \end{aligned}$$

where $\bar{f}(y, s) = f(\bar{U}(y, s))$ (resp. $\bar{f}(y, x)$ is given by (25)). Note that in either case $\sup_x E\{\bar{f}(x, t)^2\} \leq \sup_x E\{f(\bar{U}(x, t))^2\}$.

The first integral is the solution to the deterministic heat equation with initial condition $\bar{U}(x, 0)$. Now G_t is a contraction in L^∞ , while by (64) $\|\tilde{G}_t\|_\infty$ tends to zero as $t \rightarrow \infty$, so either way the first term is uniformly bounded by $M\|u_0\|_\infty$ for some M and all t , independent of h and k .

Let $F(t) = \sup_x E\{\bar{U}(x, t)^2\}$. Then if $0 \leq t \leq T$,

$$\begin{aligned} F(t) \leq 3M^2\|u_0\|_\infty^2 + 3 \sup_x E \left\{ \left(\int_0^t \int_{L_0} \bar{G}_{t-s}(x, y) \bar{f}(y, s) W(dy ds) \right)^2 \right\} \\ + 3 \sup_x E \left\{ \left(\int_0^t \int_{L_0} \bar{G}_{t-s}(x, y) g(\bar{U}(y, s)) dy ds \right)^2 \right\}. \end{aligned}$$

Since f and g are Lipschitz, there exists κ such that $|f(x)|^2 \leq \kappa^2(1 + x^2)$, hence

$$\sup_x E\{\bar{f}(x, t)^2\} \leq \sup_x E\{f(\bar{U}(x, t))^2\} \leq \kappa^2(1 + \sup_x E\{\bar{U}(x, t)^2\}) = \kappa^2(1 + F(t)),$$

and the expectation of the first integral is bounded by

$$\kappa^2 \int_0^t \int_{L_0} \bar{G}_{t-s}^2(x, y) (1 + F(s)) dy ds.$$

Thus, if we apply the Schwartz inequality to the last integral, we see this is

$$\begin{aligned} (31) \quad &\leq 3M^2\|u_0\|_\infty^2 + 3\kappa^2 \int_0^t \int_{L_0} \bar{G}_{t-s}(x, y)^2 (1 + F(s)) dy ds \\ &\quad + 3\kappa^2 L_0 t \int_0^t \int_L \bar{G}_{t-s}(x, y)^2 (1 + F(s)) dy ds. \end{aligned}$$

But $\int_{L_0} \bar{G}_{t-s}(x, y) dy = \|\bar{G}_{t-s}(x, \cdot)\|_2^2 \leq \frac{C}{\sqrt{t-s}}$ so that

$$F(t) \leq 3M^2\|u_0\|_\infty^2 + 3\kappa^2(1 + L_0t)\sqrt{t} + 3\kappa^2(1 + Lt)F * \gamma(t),$$

and the boundedness follows from Lemma 5.1 with $H(t) = 3M^2\|u_0\|_\infty^2 + 3\kappa^2(1 + Lt)\sqrt{t}$.

To prove (ii), write

$$\begin{aligned} \tilde{U}(x_{i+1}, t) - \tilde{U}(x_i, t) &= \int_{L_0} u_0(y) \left(\tilde{G}_t(x_{i+1}, y) - \tilde{G}_t(x_i, y) \right) dy \\ &+ \int_0^t \int_{L_0} \left(\tilde{G}_{t-s+k}(x_{i+1}, y) - \tilde{G}_{t-s+k}(x_i, y) \right) \left(\bar{f}(y, s) W(dy ds) + g(\tilde{U}(y, s)) \right) dy ds. \end{aligned}$$

The first integral is an approximation to the classical solution of the homogeneous heat equation, which we know is first order (except if $\theta = 1/2$, in which case it is second order.) Write $\tilde{G}_t \hat{u}_0(x)$ and $G_t u_0(x)$ for the approximation and the real solution respectively, and note that

$$\begin{aligned} |\tilde{G}_t \hat{u}_0(x_{i+1}) - \tilde{G}_t \hat{u}_0(x_i)| &\leq |\tilde{G}_t \hat{u}_0(x_{i+1}) - G_t u_0(x_{i+1})| \\ &+ |G_t u_0(x_{i+1}) - G_t u_0(x_i)| + |G_t u_0(x_i) - \tilde{G}_t \hat{u}_0(x_i)|. \end{aligned}$$

The first and last terms on the right hand side are $O(k)$ since the scheme is first order, while, as the solution itself is $C^{(\infty)}$ for $t > 0$ and has a locally bounded first derivative if u_0 is smooth, the second term is $O(h)$. Thus the difference is $O(h + k)$, and for each T there is a constant C_T such that its square is bounded by $C_T(h^2 + k^2)$ in $0 \leq t \leq T$, $x \in L_0$. Thus

$$\begin{aligned} E \left\{ \left(\tilde{U}(x_{i+1}, t) - \tilde{U}(x_i, t) \right)^2 \right\} \\ \leq 3C_T(h^2 + k^2) + 3M_T \int_0^t \|\tilde{G}_{t-s+k}(x_{i+1}, \cdot) - \tilde{G}_{t-s+k}(x_i, \cdot)\|_2^2 ds \\ + 3M_T Lt \int_0^t \|\tilde{G}_{t-s+k}(x_{i+1}, \cdot) - \tilde{G}_{t-s+k}(x_i, \cdot)\|_2^2 ds \end{aligned}$$

By (68), if $\theta \neq 1/2$, the integral over s is bounded by

$$\int_0^t \frac{C}{\sqrt{s}} \left(\frac{\sqrt{s}}{h} \wedge \frac{h}{s} + K_{\theta c}(s) \right) ds \leq C \left(h + \frac{k}{h} \right).$$

and if $\theta = 1/2$, we replace k/h by k^2/h^3 in the bound. Assuming that h and k are less than one, $h^2 + k^2 < h + k/h$, so this proves (30). ■

5.1 L^2 Convergence

We are now ready to prove the main theorem. We begin with the L^2 convergence.

$$\begin{aligned}
(32) \quad U(x, t) - \tilde{U}(x, t) &= \int_{L_0} \left(G_t(x, y)u_0(y) - \tilde{G}_t(x, y)\hat{u}_0(y) \right) dy \\
&+ \int_0^t \int_{L_0} \left(G_{t-s}(x, y)f(U(y, s)) - \tilde{G}_{t-s+k}(x, y)\tilde{f}(x, y) \right) W(dy ds) \\
&+ \int_0^t \int_{L_0} \left(G_{t-s}(x, y)g(U(y, s)) - \tilde{G}_{t-s+k}(x, y)\tilde{g}(\tilde{U}(y, s)) \right) dy ds \\
&= \int_{L_0} \left(G_t(x, y)u_0(y) - \tilde{G}_t(x, y)\hat{u}_0(y) \right) dy \\
&+ \int_0^t \int_{L_0} G_{t-s}(x, y) \left(f(U(y, s)) - f(\tilde{U}(y, s)) \right) W(dy ds) \\
&+ \int_0^t \int_{L_0} f(\tilde{U}(y, s)) \left(G_{t-s}(x, y) - \tilde{G}_{t-s+k}(x, y) \right) W(dy, ds) \\
&+ \int_0^t \int_{L_0} \tilde{G}_{t-s+k}(x, y) \left(f(\tilde{U}(y, s)) - \tilde{f}(x, y) \right) W(dy ds) \\
&+ \int_0^t \int_{L_0} G_{t-s}(x, y) \left(g(U(y, s)) - g(\tilde{U}(y, s)) \right) dy ds \\
&+ \int_0^t \int_{L_0} g(\tilde{U}(y, s)) \left(G_{t-s}(x, y) - \tilde{G}_{t-s+k}(x, y) \right) dy, ds \\
&\stackrel{\text{def}}{=} I_1 + I_2 + I_3 + I_4 + I_5 + I_6.
\end{aligned}$$

Let $F(t) = \sup_x E \left\{ (U(x, t) - \tilde{U}(x, t))^2 \right\}$ and consider the I_j in order.

I_1 is well-known. It involves a solution to the homogeneous equation, and by [5] $|I_1| \leq C(1 + \frac{h}{t})\|u_0 - \hat{u}_0\|_\infty$. It is not hard to see that if u_0 is Lipschitz, $\|u_0 - \hat{u}_0\|_\infty = O(h)$. (Indeed, \hat{u}_0 is the projection of u_0 on \mathcal{S}_n . First compare u_0 with the piecewise linear function u_p which equals u_0 at $0, h, 2h, \dots, L$. The difference is at most $\kappa h/2$, where κ is a Lipschitz constant for u_0 . Since u_p is in \mathcal{S}_n , it follows that $\|\hat{u}_0 - u_p\|_\infty \leq \kappa h$.)

Let M_T be greater than all of $E\{f(U(y, t))^2\}$, $E\{f(\tilde{U}(y, t))^2\}$, $E\{g(U(y, t))^2\}$, and $E\{g(\tilde{U}(y, t))^2\}$, $0 \leq t \leq T$, $y \in L_0$, which is possible by Proposition 5.2. Let κ be a Lipschitz constant for both f and g as well as u_0 . Note that $E\{(f(U(x, s)) - f(\tilde{U}(x, s)))^2\} \leq \kappa^2 E\{(U(x, s) - \tilde{U}(x, s))^2\} \leq \kappa^2 F(s)$. We handle the remaining I_j as we did in the previous proposition. For $t \leq T$,

$$(33) \quad E\{I_2^2\} \leq \kappa^2 \int_0^t \|G_{t-s}(x, \cdot)\|_2^2 F(s) ds \leq \kappa^2 \int_0^t \frac{C}{\sqrt{t-s}} F(s) ds,$$

by (53). Next

$$(34) \quad E\{I_3^2\} \leq M_T \int_0^t \|G_{t-s}(x, \cdot) - \tilde{G}_{t-s+k}(x, \cdot)\|_2^2 ds.$$

Let $x_i = ih$. To bound I_4 , we notice that if $y \in (x_i, x_{i+1})$, that $f(y, s)$ is the average value of $f(\tilde{U}(\cdot, s))$ on that interval. The average value falls between the maximum and minimum, so that $|f(\tilde{U}(\cdot, s)) - \tilde{f}(y, x)| \leq \sup_{y, y' \in (x_i, x_{i+1})} |f(\tilde{U}(y, s)) - f(\tilde{U}(y', s))| \leq \kappa \sup_{y, y' \in (x_i, x_{i+1})} |U(y, s) - U(y', s)| = \kappa |U(x_{i+1}, s) - U(x_i, s)|$, because \tilde{U} is linear on (x_i, x_{i+1}) . Thus, by (64) and (30)

$$(35) \quad E\{I_4^2\} \leq \int_0^t \|\tilde{G}_{t-s}(x, \cdot)\|_2^2 \sup_y E \left\{ \left(f(\tilde{U}(y, s)) - \tilde{f}(s, y) \right)^2 \right\} ds \\ \leq \int_0^t \frac{C}{\sqrt{h^2 \vee (t-s)}} \kappa^2 \left(h + \frac{k}{h} \right) ds$$

The remaining two integrals are similar, once we apply Schwartz' inequality:

$$(36) \quad E\{I_5^2\} \leq Lt \int_0^t \|G_{t-s}(x, \cdot)\|_2^2 \sup_y E \{ (g(U(y, s)) - g(\tilde{U}(y, s)))^2 \} dy ds \\ \leq \kappa^2 Lt \int_0^t \frac{C}{\sqrt{t-s}} F(s) ds$$

and

$$(37) \quad E\{I_6^2\} \leq LtM_T \int_0^t \|G_{t-s}(x, \cdot) - \tilde{G}_{t-s+k}(x, \cdot)\|_2^2 ds.$$

Now $\|G_t - \tilde{G}_t\|_2^2 \leq 3\|G_t - \hat{G}_t\|_2^2 + 3\|\hat{G}_t - \check{G}_{t+k}\|_2^2 + 3\|\check{G}_{t+k} - \tilde{G}_{t+k}\|_2^2$. Use (54), (66), (67) and integrate to see that

$$\int_0^t \|G_{t-s} - \tilde{G}_{t-s+k}\|_2^2 ds \leq \int_0^\infty \|G_s - \tilde{G}_{s+k}\|_2^2 ds \leq C \left(h + \sqrt{k} + \frac{k}{h} I_{\{\theta \neq \frac{1}{2}\}} + \frac{k^2}{h^3} I_{\{\theta = \frac{1}{2}\}} \right).$$

$F(s) \leq 6 \sum_{i=1}^6 E\{I_i^2\}$, so, adding the above and combining constants, we see

$$F(t) \leq C \left(h + \sqrt{k} + \frac{k}{h} I_{\{\theta \neq \frac{1}{2}\}} + \frac{k^2}{h^3} I_{\{\theta = \frac{1}{2}\}} \right) + C \left(\frac{1}{\sqrt{t}} * F \right) (t).$$

Now apply Lemma 5.1 with $H(t) = C \left(h + \sqrt{k} + \frac{k}{h} I_{\{\theta \neq \frac{1}{2}\}} + \frac{k^2}{h^3} I_{\{\theta = \frac{1}{2}\}} \right)$. ■

5.2 Proof of Theorem 2.1 Continued: a.e. Convergence

The L^2 convergence rate is really just an average. To see that the scheme actually converges each time we run it, we need an almost sure result. As a rule of thumb, the almost sure convergence rate will be within epsilon of the L^2 convergence rate, as we will show in this section. Let

$$\Theta_p(x, \theta) = 1 + x^{\frac{p+1}{2}} I_{\{\theta = \frac{1}{2}\}} + x^{\frac{p-1}{2}} I_{\{\frac{1}{2} < \theta < 1\}}.$$

Theorem 5.3 *Let $r > 6$. For each T there exists a constant C_{rT} , independent of h and k , such that*

$$(38) \quad \sup_{x \in L_0, t \leq T} E\{|\tilde{U}(x, t)|^r\} \leq C_{rT};$$

$$(39) \quad \sup_{x \in L_0, t \leq T} E\{|U(x, t) - \tilde{U}(x, t)|^r\} \leq C_{rT} \left(h^{\frac{r-6}{2}} + k^{\frac{r-6}{4}} \Theta_{\frac{2r}{r-2}} \left(\frac{k}{h^2}, \theta \right) \right).$$

Theorem 2.1 (ii) follows almost immediately from this. Indeed, if k/h^2 is bounded, the function $\Theta_{\frac{2r}{r-2}}$ is bounded and can be absorbed into the constant, so that we can write

$$P\left\{|U(x, t) - \tilde{U}(x, t)| > \frac{\delta}{N^{\frac{1}{4}-\varepsilon}}\right\} \leq E\{|U(x, t) - \tilde{U}(x, t)|^r\} \frac{N^{(\frac{1}{4}-\varepsilon)r}}{\delta^r}$$

But $k = O(1/N)$ and $h = O(1/N^2)$, so by Theorem 5.3, $E\{|U(x, t) - \tilde{U}(x, t)|^r\} \leq C/N^{\frac{r-6}{4}}$, and the above probability is

$$\leq C \frac{N^{\frac{3}{2}-\varepsilon r}}{\delta^r}.$$

Choose r large enough so that $\varepsilon r > 3/2$, and let $N = 2^m$:

$$= \frac{C}{\delta^r} 2^{-(\varepsilon r - \frac{3}{2})m},$$

which is summable. By the Borel-Cantelli lemma, $|U(x, t) - \tilde{U}(x, t)| \leq \frac{\delta}{N^{\frac{1}{4}-\varepsilon}}$ for all large enough N . This is true simultaneously for a sequence of δ tending to zero, which proves Theorem 2.1 (ii). ♣

It remains to prove Theorem 5.3. Before starting the proof, let us write down some inequalities for the L^p norms of the Green's functions. These follow readily from the estimates of their L^2 and L^∞ norms in Section 8.

Lemma 5.4 *Suppose $2 < p < 3$. Then there exists a constant C_p such that*

$$(i) \int_0^t \int_{L_0} \tilde{G}_s(x, y)^p dy ds \leq C_p t^{\frac{3-p}{2}} ;$$

$$(ii) \int_0^t \int_{L_0} |G_s(x, y) - \tilde{G}_{s+k}(x, y)|^p dy ds \leq C_p \left(h^{3-p} + k^{\frac{3-p}{2}} \Theta_p \left(\frac{k}{h^2}, \theta \right) \right).$$

PROOF. $\tilde{G}_s(x, y)^p \leq \tilde{G}_s(x, y)^2 \|\tilde{G}_s(x, \cdot)\|_\infty^{p-2}$, so

$$\int_{L_0} |\tilde{G}_s(x, y)|^p dy ds \leq \|\tilde{G}_s(x, \cdot)\|_2^2 \|\tilde{G}_s(x, \cdot)\|_\infty^{p-2}.$$

From (64) we see that $\|\tilde{G}_s(x, \cdot)\|_2^2 \leq C \|\tilde{G}_s(x, \cdot)\|_\infty$, so the above is

$$\leq \|\tilde{G}_s(x, \cdot)\|_\infty^{p-1} \leq C t^{\frac{1-p}{2}},$$

and integration gives (i).

(ii) By similar reasoning,

$$\int_{L_0} |G_s(x, y) - \tilde{G}_{s+k}(x, y)|^p dy = \|G_s(x, \cdot) - \tilde{G}_{s+k}(x, \cdot)\|_p^p \leq \|G_s(x, \cdot) - \tilde{G}_{s+k}(x, \cdot)\|_\infty^{p-1},$$

and $\|G_s - \tilde{G}_{s+k}\|_\infty \leq \|G_s - \hat{G}_s\|_\infty + \|\hat{G}_s - \check{G}_{s+k}\|_\infty + \|\check{G}_{s+k} - \tilde{G}_{s+k}\|_\infty$. By (54), (66) and (68), this is

$$\leq \frac{C}{\sqrt{t}} \left(1 \wedge \frac{h^2}{t} + \sqrt{\frac{t}{h^2 \vee k}} \wedge \frac{h^2 \vee k}{t} + \frac{\sqrt{t}}{h} \wedge \frac{h}{\sqrt{t}} + K_{\theta c}(t) \right).$$

Now $\sqrt{t/h^2 \vee k} \wedge (h^2 \vee k)/t \leq 1 \wedge (h^2/t) + 1 \wedge (k/t)$ and $1 \wedge (h^2/t) \leq 1 \wedge h/\sqrt{t}$, so this is

$$\leq \frac{C}{\sqrt{t}} \left(1 \wedge \frac{h}{\sqrt{t}} + 1 \wedge \frac{k}{t} + K_{\theta c}(t) \right),$$

where $K_{\theta c}$ is given by (63). Thus

$$\|G_s - \tilde{G}_s\|_\infty^{p-1} \leq \frac{C}{t^{\frac{p-1}{2}}} \left(1 \wedge \left(\frac{h^2}{t} \right)^{\frac{p-1}{2}} + 1 \wedge \left(\frac{k}{t} \right)^{p-1} + K_{\theta c}(t)^{p-1} \right)$$

We can explicitly integrate this over t to prove (ii). ♣

This brings us to the proof of the theorem.

PROOF. Suppose $r > 6$ and calculate

$$\begin{aligned}
E\{|U(x, t)|^r\} &\leq 3^{r-1} \left(\int_{L_0} u_0(y) \tilde{G}_t(x, y) dy \right)^r \\
&\quad + 3^{r-1} E \left\{ \left| \int_0^t \int_{L_0} f(\tilde{U}(y, s)) \tilde{G}_{t-s}(x, y) W(dy ds) \right|^r \right\} \\
&\quad + 3^{r-1} E \left\{ \left| \int_0^t \int_{L_0} g(\tilde{U}(y, s)) \tilde{G}_{t-s}(x, y) dy ds \right|^r \right\} \\
&\stackrel{\text{def}}{=} 3^{r-1} (I_1 + I_2 + I_3).
\end{aligned}$$

Let $F_r(t) = \sup_y E\{|\tilde{U}(y, t)|^r\}$. First, we know that I_1 is bounded on compact t -intervals. Apply Burkholder's inequality to the stochastic integral:

$$I_2 \leq E \left\{ \left(\int_0^t \int_{L_0} f^2(\tilde{U}(y, s)) \tilde{G}_{t-s}^2(x, y) dy ds \right)^{\frac{r}{2}} \right\}.$$

By Hölder's inequality with $p = r/2$ and $q = r/(r-2)$, this is

$$\leq \int_0^t \int_{L_0} E \left\{ f(\tilde{U}(y, s))^r \right\} dy ds \left(\int_0^t \int_{L_0} \tilde{G}_{t-s}^{\frac{2r}{r-2}}(x, y) dy ds \right)^{\frac{r-2}{2}}.$$

Now f is Lipschitz, so that $f(x) \leq \gamma(1 + |x|)$ for some constant γ , so that

$$E \left\{ f(\tilde{U}(y, s))^r \right\} \leq (2\gamma)^r \left(1 + E \left\{ |\tilde{U}(y, s)|^r \right\} \right).$$

Apply Lemma 5.4 with $p = 2r/(r-2)$ to see that

$$\left(\int_0^t \int_{L_0} \tilde{G}_s^{\frac{2r}{r-2}}(x, y) dy ds \right)^{\frac{r-2}{2}} \leq Ct^{\frac{r-6}{4}}.$$

Thus

$$I_2 \leq Ct^{\frac{r-6}{4}} \int_0^t (1 + F_r(s)) ds.$$

The third integral is similar, except that we apply Hölder's inequality directly with $p = r$, $q = r/(r-1)$:

$$I_3 \leq Ct^{\frac{2r-3}{2}} \int_0^t (1 + F_r(s)) ds.$$

Thus

$$F_r(t) \leq I_1(t) + C(t^{\frac{r-2}{4}} + t^{\frac{2r-1}{2}}) + C(t^{\frac{r-6}{4}} + t^{\frac{2r-3}{2}}) \int_0^t (1 + F_r(s)) ds.$$

(i) now follows from this and Gronwall's inequality.

To prove (ii), let $H_r(t) = \sup_x E\{|U(x, t) - \tilde{U}(x, t)|^r\}$. From (32),

$$|U(x, t) - \tilde{U}(x, t)|^r \leq 6^{r-1}(|I_1|^r + \cdots + |I_6|^r).$$

By [5], $|I_1|^r \leq C(1 + h/t)^r \|u_0 - \hat{u}_0\|_\infty^r$, since \hat{u}_0 , the projection of u_0 on \mathcal{S}_n , is the initial value of \tilde{U} . We have assumed u_0 is smooth, so it is easy to see that $\|u_0 - \hat{u}_0\|^r$ is dominated by Ch^r , hence $|I_1|^r \leq Ch^r$ for some C .

I_2, I_3 and I_4 are stochastic integrals, which we handle by first using Burkholder's inequality, and then Hölder's inequality and the Lemma. Noting that $|f(U(y, s)) - f(\tilde{U}(y, s))| \leq \gamma|U(y, s) - \tilde{U}(y, s)|$, we see

$$\begin{aligned} E\{|I_2|^r\} &\leq CE \left\{ \left(\int_0^t \int_{L_0} G_t(x, y)^2 |U(y, s) - \tilde{U}(y, s)|^2 dy ds \right)^{\frac{r}{2}} \right\} \\ &\leq C \left(\int_0^t \int_{L_0} G_t(x, y)^{\frac{2r}{r-2}} dy ds \right)^{\frac{r-2}{2}} \int_0^t \int_{L_0} E\{|U(y, s) - \tilde{U}(y, s)|^r\} dy ds \\ &\leq Ct^{\frac{6-r}{4}} \int_0^t H_r(s) ds. \end{aligned}$$

Similarly,

$$E\{|I_2|^r\} \leq C \int_0^t \int_{L_0} E\{|f(\tilde{U}(y, s))|^r\} dy \left(\int_0^t \int_{L_0} (G_{t-s}(s, y) - \tilde{G}_{t-s}(x, y))^{\frac{2r}{r-2}} dy ds \right)^{\frac{r-2}{2}}.$$

By Lemma 5.4, for all t ,

$$\left(\int_0^t \int_{L_0} |G_{t-s}(x, y) - \tilde{G}_{t-s}(x, y)|^{\frac{2r}{r-2}} dy ds \right)^{\frac{r-2}{2}} \leq C(h^{\frac{r-6}{2}} + k^{\frac{r-6}{4}} \Theta_{\frac{2r}{r-2}}(\frac{k}{h^2}, \theta)).$$

Now we have seen that $E\{|f(\tilde{U}(y, s))|^r\}$ is bounded by $C(1 + F_r(s))$, so that

$$E\{|I_3|^r\} \leq C \left(h^{\frac{r-6}{2}} + k^{\frac{r-6}{4}} \Theta_{\frac{2r}{r-2}}(\frac{k}{h^2}, \theta) \right) \int_0^t (1 + F_r(s)) ds.$$

Now from (53) and (64), $\tilde{G}_{t-s} \leq CG_{t-s}$, so $E\{|I_4|^r\} \leq CE\{|I_2|^r\}$. The integrals I_5 and I_6 are similar, except that we use Hölder with $p = r$, $q = r/(r-1)$ to get

$$E\{|I_5|^r\} \leq Ct^{2r-3} \int_0^t H_r(s) ds,$$

and

$$E|I_6|^r \leq C \int_0^t (1 + F_r(s)) ds \left(h^{2r-3} + k^{\frac{2r-3}{2}} \Theta_{\frac{2r}{r-2}}\left(\frac{k}{h^2}, \theta\right) \right).$$

Note that as h and k tend to zero, h^{2r-3} and $k^{\frac{2r-3}{2}}$ are negligible compared to $h^{\frac{r-6}{2}}$ and $k^{\frac{3r-2}{4}}$ respectively, so that I_6 is negligible compared with I_3 . Likewise, $E\{|I_1|^r\}$ is $O(h)$, which is negligible with respect to $E\{|I_3|^r\}$. Adding everything together, we see there are continuous increasing functions $A(t)$ and $B(t)$ such that

$$H_r(t) \leq \left(h^{\frac{r-6}{2}} + k^{\frac{r-6}{4}} \Theta_{\frac{2r}{r-2}}\left(\frac{k}{h^2}, \theta\right) \right) A(t) + B(t) \int_0^t H_r(s) ds.$$

Since A and B are increasing, Gronwall's inequality implies that

$$H_r(t) \leq \left(h^{\frac{r-6}{2}} + k^{\frac{r-6}{4}} \Theta_{\frac{2r}{r-2}}(t) \right) A(T) e^{tB(T)}, \quad 0 \leq t \leq T,$$

which proves the theorem. ♣

6 Non-convergence

The one-step theta scheme is stable if $\theta \geq 1/2$, but stability does not guarantee convergence. In fact Theorem 2.1 suggests that even if h and k tend to zero, the error may not go to zero unless k/h also tends to zero. This is in fact correct, as we can see by considering the linear case.

Let $L_0 = [0, 1]$, $f \equiv 1$ and $g \equiv 0$. We will show that if both h and k tend to zero and if either

- (i) $\frac{1}{2} < \theta < 1$ and $k \geq h$, or
- (ii) $\theta = \frac{1}{2}$ and $k \geq h^{\frac{3}{2}}$,

the numerical scheme does not converge to the true solution, either a.e. or in L^2 .

To see this, note that by (19), the numerical approximation \tilde{U} is

$$\tilde{U}(x, t_m) = \sum_{i=0}^{m-1} \Phi^T(x) \Gamma_{\theta}^{m-i} D_i$$

where $t_m \stackrel{\text{def}}{=} mk$ and D_i is the vector increment

$$D_i = \int_{t_i}^{t_{i+1}} M^{-1} \bar{\Phi}(y) W(dy ds).$$

Then

$$\begin{aligned} E\{D_i D_i^T\} &= \int_{t_i}^{t_{i+1}} M^{-1} \bar{\Phi}(y) \bar{\Phi}^T(y) M^{-1} dy \\ &= k \bar{M} M^{-2}. \end{aligned}$$

Since the D_i are independent,

$$\begin{aligned} E\{\tilde{U}(x, t_m)^2\} &= k \sum_{i=0}^{m-1} \bar{\Phi}^T(x) \bar{M} M^{-2} \Gamma_\theta^{2m-2i} \hat{\psi}_j(x)^2 \\ &= k \sum_{j=1}^{\frac{1}{h}-1} \sum_{i=0}^{m-1} \frac{\lambda_j(\bar{M})}{\lambda_j(M)} \lambda_j(\Gamma_\theta)^{2m-2i} \hat{\psi}_j(x)^2 \end{aligned}$$

by (100). Integrate over x to see that

$$(40) \quad E\{\|\tilde{U}(\cdot, t_m)\|_2^2\} = \frac{k}{3} \sum_{j=1}^{\frac{1}{h}-1} \cos^2 \frac{j\pi h}{2} \sum_{i=0}^{m-1} \lambda_j(\Gamma_\theta)^{2m-2i}.$$

If $\frac{1}{2} < \theta < 1$ and $h = k$, $\hat{\lambda}_j = O(j^2)$ so that

$$\lambda_j(\Gamma_\theta) \sim 1 - \frac{hj^2}{1 + \theta hj^2}.$$

Now $\lambda_j(\Gamma_\theta)$ is decreasing in j , and it is eventually negative if h is small enough, hence for small h , it is bounded away from zero, say $|\lambda_j(\Gamma_\theta)| \geq \delta$ for all $j \geq \frac{1}{2h}$. The sum over i in (40) is larger than its last term, so it is

$$\geq \frac{k\delta}{3} \sum_{j=\frac{1}{2h}}^{\frac{1}{h}} \cos^2 \frac{j\pi h}{2}.$$

The sum can be estimated by an integral: it is approximately $\frac{1}{4h}$, and since $k \geq h$, we conclude that there is $\varepsilon > 0$ such that for all small enough h ,

$$(41) \quad E\{\|\tilde{U}(\cdot, t)\|_2^2\} \geq \varepsilon^2, \quad \forall t > 0.$$

We claim (41) also holds if $\theta = \frac{1}{2}$ and $k \geq h^{\frac{3}{2}}$. To see this, first sum over i in (40), to get

$$E\{\|U(\cdot, t_m)\|_2^2\} = \frac{k}{3} \sum_{j=1}^{\frac{1}{h}-1} \cos^2 \frac{j\pi h}{2} \frac{\lambda_j(\Gamma_{\frac{1}{2}})^2 - \lambda_j(\Gamma_{\frac{1}{2}})^{2m}}{1 - \lambda_j(\Gamma_{\frac{1}{2}})^2}$$

But for large j , $k\lambda_j$ is large, so

$$\lambda_j(\Gamma_{\frac{1}{2}}) = -\left(1 - \frac{4}{2 + k\hat{\lambda}_j}\right) \sim -\left(1 - \frac{4}{k\hat{\lambda}_j}\right),$$

hence, for large j , $1 - \lambda_j(\Gamma_{\frac{1}{2}})^2 \sim \frac{8}{k\lambda_j}$ and

$$e^{-\frac{8}{k\lambda_j}} \leq |\lambda_j(\Gamma_{\frac{1}{2}})| \leq e^{-\frac{4}{k\lambda_j}}$$

Thus the above is

$$\geq \frac{k}{24} \sum_{j=1}^{\frac{1}{h}-1} k\hat{\lambda}_j \cos^2 \frac{j\pi h}{2} e^{-8k\hat{\lambda}_j} \left(1 - e^{-\frac{4m}{k\lambda_j}}\right).$$

Now take $m = 1/\sqrt{h}$ (so $t_m = mk = h$) and note that $\hat{\lambda}_j \geq 4j^2$ so that $e^{-\frac{4m}{k\lambda_j}} \leq 1/e$, and this is

$$\begin{aligned} &\geq C_1 k^2 \sum_{j=1}^{\frac{1}{h}-1} j^2 \cos^2 \frac{j\pi h}{2} \\ &\geq C_2 \frac{k^2}{h^3} \end{aligned}$$

for some positive constants C_1 and C_2 . Since $k^2 \geq h^3$, this verifies (41) for each $t > h$. However, since $h \rightarrow 0$, (41) holds in the limit for all $t > 0$.

But the true solution $U(x, t)$ is continuous and satisfies $U(x, 0) = 0$, so $E\{\|U_t\|_2^2\} \rightarrow 0$ as $t \rightarrow 0$. Thus for small enough t , it is less than $\varepsilon^2/4$, so

$$E\{\|U(\cdot, t) - \tilde{U}(\cdot, t)\|_2^2\}^{\frac{1}{2}} \geq E\{\|\tilde{U}_t\|_2^2\}^{\frac{1}{2}} - E\{\|U_t\|_2^2\}^{\frac{1}{2}} \geq \frac{\varepsilon}{2}.$$

Thus $\tilde{U}(x, t)$ can not converge to $U(x, t)$ as $h \rightarrow 0$, either a.s. or in L^2 .

7 A Milstein-type Correction

The correction term we discuss in this section only applies to the stochastic integral, so we may as well assume that $g \equiv 0$. Let $t_m = mk$, $m = 0, 1, 2, \dots$, let $x_i = ih$ and let $y_i = \frac{1}{2}(x_i + x_{i+1})$ be the midpoint of (x_i, x_{i+1}) . Let Δ_{im} be the rectangle $(x_i, x_{i+1}) \times (t_m, t_{m+1})$. The vector increment D_m at step m of the time-discretization scheme (4.4) was a discrete approximation to

$$(42) \quad \int_{t_m}^{t_{m+1}} \int_{L_0} f(\hat{U}(y, s)) M^{-1} \Phi(y) W(dy ds).$$

For the first approximation, we replaced $f(\hat{U}(y, s))$ by its value at the start of the interval, $f(\hat{U}(y, t_m))$. After discretization, this led to (20) (slightly simplified here):

$$(43) \quad D_m = \sum_i f(\hat{U}(y_i, t_m)) M^{-1} \Phi(y_i) W(\Delta_{im}).$$

(When we use this, we replace the continuous time process \hat{U}_t by the discrete process \tilde{U}_t , but we will leave it in terms of \hat{U}_t here.) Now the best choice for the discrete approximation to the integral in (20) would be its conditional expectation given \mathcal{F}_{t_m} and the $W(\Delta_{ij})$. We can compute that—up to higher order terms—and it leads to what we call the *pseudo-Milstein approximation* \tilde{D}_m :

$$(44) \quad \begin{aligned} \tilde{D}_m &= \sum_i f(\hat{U}(y_i, t_m)) M^{-1} \Phi(y_i) W(\Delta_{im}) \\ &\quad + \frac{1}{2} \sum_{i,j} f'(\hat{U}(y_i, t_m)) f(\hat{U}(y_j, t_m)) \Phi^T(y_i) M^{-1} \Phi(y_j) (W(\Delta_{im}) W(\Delta_{jm}) - hk \delta_{ij}) \end{aligned}$$

To derive this, note that $f(\hat{U}(x, t))$ is not constant in t , so we can improve our estimate by plugging a first order approximation in the stochastic integral. Now $\hat{U}(y, s) = \Phi^T(y) \hat{A}(s)$, so that

$$\begin{aligned} f(\hat{U}(y, t+u)) &\sim f(\hat{U}(y, t_m)) + f'(\hat{U}(y, t_m)) (\hat{U}(y, t_m+u) - \hat{U}(y, t_m)) \\ &= f(\hat{U}(y, t_m)) + f'(\hat{U}(y, t_m)) \Phi^T(y) (\hat{A}(t+u) - \hat{A}(t)). \end{aligned}$$

From (8) again,

$$\hat{A}(t+u) - \hat{A}(t_m) \sim -Q \hat{A}(t) u + \int_{t_m}^{t_m+u} f(\hat{U}(z, t_m)) M^{-1} \Phi(z) W(dz dv).$$

According to a standard rule of thumb, non-stochastic second-order terms are negligible, so we can neglect $Q\hat{A}(t_m)u$ and define

$$(45) \quad \hat{D}_m \stackrel{\text{def}}{=} \int_{t_m}^{t_{m+1}} \int_{L_0} f(\hat{U}(y, s)) M^{-1} \Phi(y) W(dy ds) \\ + \int_{t_m}^{t_{m+1}} \int_{L_0} f'(\hat{U}(y, s)) \int_{t_m}^s \int_{L_0} f(\hat{U}(x, t)) M^{-1} W(dz dv) W(dy ds) + o(k).$$

The second term above is a multiple Wiener white noise integral. It turns out that $\tilde{D}_m = E\{\hat{D}_m \mid \mathcal{F}_t^+\} + o(k)$. We will leave it to the reader to verify that the neglected terms are of $o(k)$, and just note that the claim about the conditional expectation follows from the following proposition.

Proposition 7.1 *Let $H(x)$ and $K(x, y)$ be \mathcal{F}_{t_m} -measurable, square-integrable functions. Then*

$$E \left\{ \int_{L_0 \times (t_m, t_{m+1})} H(x) W(dx ds) \mid \mathcal{F}_{t_m}^+ \right\} = \sum_i \left(\frac{1}{h} \int_{x_i}^{x_{i+1}} H(x) dx \right) W(\Delta_{im})$$

and

$$(46) \quad E \left\{ \int_{(L_0 \times (t_m, t_{m+1}))^2, u < v} K(x, y) W(dx du) W(dy dv) \mid \mathcal{F}_{t_m}^+ \right\} \\ = \frac{1}{2} \sum_{i,j} \left(\frac{1}{h^2} \int_{x_i}^{x_{i+1}} \int_{x_j}^{x_{j+1}} K(x, y) dx dy \right) (W(\Delta_{im}) W(\Delta_{jm}) - \delta_{ij}).$$

The proof is straightforward: first reduce it to the case where H and K are deterministic, then indicator functions, then constants. It then comes down to showing that if B_t and B'_t are independent Brownian motions, that

$$E \left\{ \int_0^t B_s dB_s \mid B_t \right\} = \frac{1}{2}(B_t^2 - t);$$

and

$$E \left\{ \int_0^t B_s dB'_s \mid B_t, B'_t \right\} = \frac{1}{2} B_t B'_t.$$

The first is an identity, and the second follows from symmetry and the fact that

$$B_t B'_t = \int_0^t B_s dB'_s + \int_0^t B'_s dB_s.$$

Remark 7.2 1.) This suggests an alternative to (44), using averages rather than evaluating the functions at points. And indeed, we used this in (4.4).

2) We should note that, although the approximation of the stochastic integral is one of the more important sources of error, there are other errors involved of roughly the same magnitude.

3.) The one-dimensional version of this approximation does indeed give the Milstein correction term. It is not hard to see why this works for a single SDE but not for systems or SPDE's. We know from Ito's analysis that the second order properties are the key to understanding stochastic integrals. A single SDE only involves a single Brownian motion, and the only second order integral is $\int B dB$, which can be written explicitly in terms of the increment of B_t . This implies that the Milstein term gives a very good approximation to the integrals involved. However, in systems, the second-order integrals include terms of the form $\int_0^t B_s dB'_s$ for independent Brownian motions B_t and B'_t . The best approximation of this in terms of B_t and B'_t is $\frac{1}{2}B_t B'_t$. This is not very good: the error $\int_0^t B_s dB'_s - \frac{1}{2}B_t B'_t$ has half the variance of the original integral. Thus it is reasonable to expect that the pseudo-Milstein approximation would improve the result, but probably by a factor of less than two, which is by no means sufficient to improve the order of convergence.

4) One can continue this procedure to get higher order terms for \tilde{D}_m , but one wouldn't expect this to make a noticeable improvement—the terms being added are negligible compared to the existing errors.

8 L^2 and L^∞ Bounds on the Green's Functions

Our results are based on moment estimates of the four Green's functions. There are a large number of special cases to treat, for the cases $\theta = 0$, $\theta = 1/2$, $1/2 < \theta < 1$ and $\theta = 1$ will often have to be treated separately. Most involve sums of the kind treated in the following lemma.

Lemma 8.1 *Let $L > 0$, $p \geq 0$ and $a > 0$ be real numbers and let $m \geq 2$ be an integer. Then*

$$(47) \quad \sum_{j=m+1}^{\infty} e^{-a^2 j^2} \leq \frac{\sqrt{\pi}}{2a} \wedge \frac{e^{-a^2 m^2}}{2ma^2}.$$

Moreover, there exists a constant $C_p \leq \Gamma(\frac{p+1}{2})$ such that

$$(48) \quad \sum_{j=1}^{m-1} j^p e^{-a^2 j^2} \leq C_p \left(m \wedge \frac{1}{a} \right)^{p+1}.$$

In particular, for $b > 0$ and $h = L/m$, there exists a constant C_{pbL} , depending only on p and (p, b, L) respectively, such that

$$(49) \quad h^p \sum_{j=1}^{\frac{L}{h}-1} j^p e^{-b^2 t j^2} \leq \frac{C_{pbL}}{\sqrt{t}} \left(\sqrt{\frac{t}{h^2}} \wedge \left(\frac{h^2}{t} \right)^{\frac{p}{2}} \right).$$

PROOF. Equation (47) follows by comparison with the integral:

$$\sum_{j=m+1}^{\infty} e^{-a^2 j^2} \leq \int_m^{\infty} e^{-a^2 x^2} dx.$$

Indeed, it is smaller than the integral from zero to ∞ , and it is also smaller than $\int_m^{\infty} e^{-a^2 x^2} \frac{x}{m} dx$, which we can do by the substitution $y = x^2$.

The sum (48) is obviously smaller than $\sum_{j=1}^{m-1} j^p \leq \int_0^m y^p dy = m^{p+1}/(p+1)$. Now $j^p e^{-aj^2}$ has its unique maximum when $j = \sqrt{p/2a}$. If this is greater than m , the summands are increasing on $[0, m]$, so the sum is dominated by $I \equiv \int_0^{\infty} y^p e^{-ay^2} dy = \Gamma(\frac{p+1}{2})/2a^{(p+1)/2}$. If the maximum occurs in some interval $[i, i+1]$ with $i < m$, we must add in the integral $\int_i^{i+1} y^p e^{-a^2 y} dy \leq I$, so in any case the sum is dominated by $2I$. Thus the sum is

$$\leq \min \left(\frac{\Gamma(\frac{p+1}{2})}{a^{p+1}}, \frac{m^{p+1}}{p+1} \right) \leq \Gamma \left(\frac{p+1}{2} \right) \left(m \wedge \frac{1}{a} \right)^{p+1},$$

since $\Gamma(\frac{p+1}{2}) \geq \frac{1}{p+1}$ if $p \geq 0$. This proves (48) with $C_p = \Gamma(\frac{p+1}{2})$.

Finally, (49) is a direct consequence of (48), with $C_{pbL} = C_p \max\{L^{p+1}, b^{-(p+1)}\}$. ♣

8.1 Comparing G and \hat{G}

The Green's function, or fundamental solution, of (1) (see (3)) can be expanded in terms of its eigenfunctions ψ_j and eigenvalues λ_j defined by

$$(50) \quad \begin{cases} \psi_j(x) &= \sqrt{\frac{2}{L}} \sin \frac{j\pi x}{L}, \quad 0 \leq x \leq L, \quad j = 1, 2, \dots \\ \lambda_j &= \frac{\pi^2 j^2}{L^2}, \quad j = 1, 2, \dots \end{cases}$$

$$(51) \quad G_t(x, y) = \sum_{j=1}^{\infty} e^{-\lambda_j t} \psi_j(x) \psi_j(y), \quad 0 \leq x \leq L, \quad 0 \leq y \leq L, \quad t \geq 0.$$

The finite element Green's function $\hat{G}_t(x, y)$ was defined in (11), and, by (100), has a similar expansion in terms of the basis $(\hat{\psi}_j)_{j=1}^n$ of \mathcal{S}_n (see (99)) and $\hat{\lambda}_j \stackrel{\text{def}}{=} \lambda_j(Q)$:

$$(52) \quad \hat{G}_t(x, y) = \sum_{j=1}^n e^{-\hat{\lambda}_j t} \hat{\psi}_j(x) \hat{\psi}_j(y).$$

Theorem 8.2 *There exists a constant C such that for all $x \in L_0$ and $h > 0$*

$$(53) \quad \begin{aligned} \|G_t(x, \cdot)\|_2^2 &\leq \frac{C}{\sqrt{t}}; & \|G_t(x, \cdot)\|_\infty &\leq \frac{C}{\sqrt{t}}; \\ \|\hat{G}_t(x, \cdot)\|_2^2 &\leq \frac{C}{\sqrt{t} \vee h^2}; & \|\hat{G}_t(x, \cdot)\|_\infty &\leq \frac{C}{\sqrt{t} \vee h^2}; \end{aligned}$$

$$(54) \quad \begin{aligned} \|G_t(x, \cdot) - \hat{G}_t(x, \cdot)\|_2^2 &\leq \frac{C}{\sqrt{t}} \left(1 \wedge \frac{h^2}{t}\right); \\ \|G_t(x, \cdot) - \hat{G}_t(x, \cdot)\|_\infty &\leq \frac{C}{\sqrt{t}} \left(1 \wedge \frac{h^2}{t}\right). \end{aligned}$$

PROOF. We will first prove the L^2 inequalities. The proofs for the L^∞ norms are similar, but simpler. Extend the definition of the $\hat{\lambda}_j$ and $\hat{\psi}_j$ to $j \geq n+1$ by setting $\hat{\lambda}_j = \hat{\psi}_j = 0$ if $j \geq n+1$. Both $(\psi_j)_{j=1}^\infty$ and $(\hat{\psi}_j)_{j=1}^\infty$ are ortho-normal on $L^2(L_0)$, so from (51) and (52),

$$\|G_t(x, \cdot)\|_2^2 = \sum_{j=1}^\infty e^{-2\lambda_j t} \psi_j^2(x) \leq \frac{2}{L} \sum_{j=1}^\infty e^{-\frac{2\pi^2 j^2 t}{L^2}}.$$

and

$$\|\hat{G}_t(x, \cdot)\|_2^2 = \sum_{j=1}^n e^{-2\hat{\lambda}_j t} \hat{\psi}_j^2(x) \leq \frac{6}{L} \sum_{j=1}^n e^{-\frac{8j^2 t}{L^2}},$$

using Prop 9.2 and Prop 9.3 of the appendix. By Lemma 8.1, these are bounded by $\frac{C}{\sqrt{t}}$ and $C(\frac{1}{h} \wedge \frac{1}{\sqrt{t}})$ respectively.

Let us define

$$\Delta_t(x, y, n) \stackrel{\text{def}}{=} G_t(x, y) - \hat{G}_t(x, y) = \sum_{j=1}^\infty e^{-\lambda_j t} \psi_j(x) \psi_j(y) - \sum_{j=1}^n e^{-\hat{\lambda}_j t} \hat{\psi}_j(x) \hat{\psi}_j(y).$$

We want to estimate the L^2 norm of Δ_t , and to see in particular how it varies with h and t . First, note that

$$\|\Delta_t(x, \cdot, n)\|_2 \leq \|G_t(x, \cdot)\|_2 + \|\hat{G}_t(x, \cdot)\|_2,$$

so that $\|\Delta_t(x, \cdot, n)\|_2^2$ is bounded by C/\sqrt{t} . Write

$$\begin{aligned}
\Delta_t(x, t, n) &= \sum_{j=1}^{\infty} e^{-\lambda_j t} \psi_j(x) (\psi_j(y) - \hat{\psi}_j(y)) + \sum_{j=1}^n e^{-\lambda_j t} (\psi_j(x) - \hat{\psi}_j(x)) \hat{\psi}_j(y) \\
&\quad + \sum_{j=1}^n (e^{-\hat{\lambda}_j t} - e^{-\lambda_j t}) \hat{\psi}_j(x) \hat{\psi}_j(y) \\
(55) \quad &\stackrel{\text{def}}{=} S_1(x, y, t) + S_2(x, y, t) + S_3(x, y, t).
\end{aligned}$$

Let us take these terms in order.

$$(56) \quad \int_{L_0} S_1^2(x, y, t) dy = \int_{L_0} \left(\sum_{j=1}^{\infty} e^{-2\lambda_j t} \psi_j^2(x) (\psi_j(y) - \hat{\psi}_j(y)) \right)^2 dy.$$

Expand this:

$$= \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} e^{-\lambda_k t} e^{-\lambda_j t} \psi_j(x) \psi_k(x) \langle \psi_j - \hat{\psi}_j, \psi_k - \hat{\psi}_k \rangle,$$

and note that, as both $(\psi_j)_{j=1}^{\infty}$ and $(\hat{\psi}_j)_{j=1}^n$ are orthonormal families, (and $\hat{\psi}_j = 0$ if $j \geq n+1$), $\langle \psi_j - \hat{\psi}_j, \psi_k - \hat{\psi}_k \rangle = \delta_{jk} - \langle \psi_j, \hat{\psi}_k \rangle - \langle \hat{\psi}_j, \psi_k \rangle + \delta_{jk} I_{\{j \leq n\}}$, so the above is

$$\begin{aligned}
&= \sum_{j=1}^{\infty} e^{-2\lambda_j t} \psi_j^2(x) - 2 \sum_{j=1}^n e^{-\lambda_j t} \psi_j(x) \sum_{k=1}^{\infty} e^{-\lambda_k t} \langle \psi_k, \hat{\psi}_j \rangle \psi_k(x) \\
&\quad + \sum_{j=1}^{\infty} e^{-2\lambda_j t} \psi_j^2(x) - \sum_{j=n+1}^{\infty} e^{-2\lambda_j t} \psi_j^2(x)
\end{aligned}$$

Remark that $G_t \psi_j = e^{-\lambda_j t} \psi_j$, and identify the sum over k as $G_t \hat{\psi}_j(x)$ to see this is

$$\leq 2 \sum_{j=1}^{\infty} e^{-\lambda_j t} \psi_j(x) G_t (\psi_j - \hat{\psi}_j)(x).$$

Since G_t is a contraction in L^∞ , and $\hat{\psi}_j = 0$ for $j \geq n+1$, this is

$$\leq 2 \sum_{j=1}^n e^{-\lambda_j t} \|\psi_j\|_\infty \|\psi_j - \hat{\psi}_j\|_\infty + 2 \sum_{j=n+1}^{\infty} e^{-2\lambda_j t} \|\psi_j(x)\|_\infty^2.$$

Let $\phi(x)$ be the piecewise-linear function which equals $\sin \frac{j\pi x}{L^2}$ at integer multiples of h . Then, for $j \leq n$, $\|\psi_j - \hat{\psi}_j\|_\infty \leq \|\hat{\psi}_j\|_\infty - \|\psi_j\|_\infty + \sqrt{\frac{2}{L}} \|\sin \frac{j\pi x}{L^2} - \phi(x)\|_\infty$. From (99),

$$\|\hat{\psi}_j\|_\infty - \|\psi_j\|_\infty = \sqrt{\frac{2}{L}} \left(\frac{1}{1 - \frac{2}{3} \sin^2 \frac{j\pi h}{L}} - 1 \right) = \sqrt{\frac{2}{L}} \left(\frac{\pi^2 h^2 j^2}{3L^2} \right) + O(h^4 j^4),$$

while, as the second derivative of $\sin \frac{\pi j x}{L}$ is bounded by $\frac{\pi^2 j^2}{L^2}$, $\|\sin \frac{\pi j x}{L} - \phi\|_\infty \leq \frac{\pi^2 j^2 h^2}{8L^2}$. Add these to see that

$$(57) \quad \|\psi_j - \hat{\psi}_j\|_\infty \leq \sqrt{\frac{2}{L}} \frac{\pi^2 j^2 h^2}{L^2}.$$

Thus

$$\|S_1(x, \cdot, t)\|_2^2 \leq \frac{4\pi^2 h^2}{L^3} \sum_{j=1}^n j^2 e^{-\frac{2\pi^2 t j^2}{L^2}} + \frac{2}{L} \sum_{j=n+1}^{\infty} e^{-\frac{2\pi^2 t j^2}{L^2}}.$$

Once again we apply (49) and (47) to see that this is

$$\leq \frac{C}{\sqrt{t}} \left(\sqrt{\frac{t}{h^2}} \wedge \frac{h^2}{t} \right) + \frac{C}{\sqrt{t}} \left(1 \wedge \frac{he^{-\frac{ct}{h^2}}}{\sqrt{t}} \right).$$

But since $\frac{he^{-\frac{ct}{h^2}}}{\sqrt{t}} \leq \frac{h^2}{t}$ for large t , this means

$$(58) \quad \|S_1(x, \cdot, t)\|_2^2 \leq \frac{C}{\sqrt{t}} \left(1 \wedge \frac{h^2}{t} \right).$$

The calculation of the L^2 norm of S_2 is slightly easier than that of S_1 . The $\hat{\psi}_j$ are ortho-normal, so

$$\begin{aligned} \int_{L_0} S_2^2(x, y, t) dy &= \sum_{j=1}^n e^{-2\lambda_j t} (\psi_j(x) - \hat{\psi}_j(x))^2 \\ &\leq \sum_{j=1}^n e^{-2\lambda_j t} \|\psi_j - \hat{\psi}_j\|_\infty^2 \\ &\leq Ch^4 \sum_{j=1}^{\frac{L}{h}-1} j^4 e^{-\frac{\pi^2 t j^2}{L^2}}, \end{aligned}$$

by (57). From Lemma 8.1 again,

$$(59) \quad \|S_2(x, \cdot, t)\|_2^2 \leq \frac{C}{\sqrt{t}} \left(\sqrt{\frac{t}{h^2}} \wedge \frac{h^4}{t^2} \right).$$

Moving on to S_3 , write

$$\int_{L_0} S_3(x, y, t) dy = \sum_{j=1}^n \left(e^{-\hat{\lambda}_j t} - e^{-\lambda_j t} \right)^2 \hat{\psi}_j(x)^2 \leq \frac{2}{L} \sum_{j=1}^n t^2 |\lambda_j - \hat{\lambda}_j|^2 e^{-2\lambda_j t},$$

since $|e^{-\hat{\lambda}_j t} - e^{-\lambda_j t}| \leq t|\lambda_j - \hat{\lambda}_j|e^{-\lambda_j t}$ by the mean value theorem. But now

$$\hat{\lambda}_j - \lambda_j = \frac{12 \sin^2(j\pi h/2L)}{h^2(3 - 2 \sin^2(j\pi h/2L))} - \frac{\pi^2 j^2}{L^2} = \frac{\pi^4 h^2 j^4}{12L^4} + O(j^6 h^4),$$

so the above is

$$\begin{aligned} &\leq C \sum_{j=1}^n (th^2 j^4)^2 e^{-\frac{2j^2 \pi^2 t}{L^2}} \\ &\leq C_1 t^2 h^4 \left(\frac{1}{h} \wedge \frac{1}{\sqrt{t}} \right)^9 \end{aligned}$$

using Lemma 8.1. Thus for some constant C ,

$$(60) \quad \|S_3(x, \cdot, t)\|_2^2 \leq \frac{C}{\sqrt{t}} \left(\left(\frac{t}{h^2} \right)^{\frac{5}{2}} \wedge \frac{h^4}{t^2} \right).$$

We now add (58), (59), and (60) to see there is a constant C such that

$$(61) \quad \|\Delta_t(x, \cdot, n)\|_2^2 \leq \frac{C}{\sqrt{t}} \left(1 \wedge \frac{h^2}{t} \right)$$

The L^∞ inequalities follow from almost the same calculations as the L^2 inequalities. From (51),

$$\|G_t\|_\infty \leq \sum_{j=1}^{\infty} e^{-\lambda_j t} \|\psi_j\|_\infty^2 \quad \text{and} \quad \|\hat{G}_t\|_\infty \leq \sum_{j=1}^n e^{-\hat{\lambda}_j t} \|\hat{\psi}_j\|_\infty^2$$

which can be reduced to the corresponding sums for the L^2 bounds by replacing t by $t/2$.

This gives the L^∞ bounds in (53). To prove the L^∞ bounds in (54), write

$$\|G_t - \hat{G}_t\|_\infty \leq 2 \sum_{j=1}^{\infty} e^{-\lambda_j t} \|\psi_j\|_\infty \|\psi_j - \hat{\psi}_j\|_\infty + \sum_{j=1}^n |e^{-\lambda_j t} - e^{-\hat{\lambda}_j t}| \|\hat{\psi}_j\|_\infty^2.$$

From (57), $\|\psi_j - \hat{\psi}_j\|_\infty \leq Ch^2 j^2$ and we saw that $|e^{-\lambda_j t} - e^{-\hat{\lambda}_j t}| \leq Ch^2 j^4 e^{-ctj^2}$; the first sum was handled in (58) and is bounded by $\frac{C}{\sqrt{t}}(1 \wedge \frac{h^2}{t})$. The second is bounded by $Cth^2 \sum_{j=1}^n j^4 e^{-ctj^2} \leq \frac{C}{\sqrt{t}}((\frac{t}{h^2})^{3/2} \wedge \frac{h^2}{t})$. Adding these together gives (54). \clubsuit

8.2 Comparing \hat{G} , \check{G} , and \tilde{G}

Let $\hat{G}_t(x, y)$ be the finite element Green's function, defined in (11), and let $\tilde{G}_t^\theta(x, y)$ be the discrete-time-and-space Green's function for the one-step theta semi-implicit scheme, defined in (26). We will also use an intermediate Green's function, $\check{G}_t^\theta(x, y)$. The function \tilde{G}_t corresponds to the process in which the integrals are discretized: the stochastic integrals in particular, are sums of the $W(\Delta_{ij})$. The function \check{G}_t corresponds to the process which is discretized in both space and time, but in which the stochastic integrals are not discretized.

We are ultimately interested in the difference between the original process and the discrete approximation, which can be expressed in terms of G_t and \tilde{G}_t . However, the discretization can be broken into several steps, and it is convenient to check the individual steps to see which kind of errors each introduces. In matrix form:

$$\begin{aligned} \hat{G}_t(x, y) &= \Phi^T(x) M^{-1} e^{-tQ} \Phi(y) \\ \check{G}_t(x, y) &= \Phi^T(x) M^{-1} \Gamma_\theta^{\lceil \frac{t}{k} \rceil} \Phi(y) \\ \tilde{G}_t(x, y) &= \Phi(x)^T M^{-1} \Gamma_\theta^{\lceil \frac{t}{k} \rceil} \bar{\Phi}(y). \end{aligned}$$

This leads to expansions in the orthonormal functions $\hat{\psi}_j$:

$$(62) \quad \begin{aligned} \check{G}_t(x, y) &= \sum_{j=1}^n \lambda_j(\Gamma_\theta)^{\lceil \frac{t}{k} \rceil} \hat{\psi}_j(x) \hat{\psi}_j(y) \\ \tilde{G}_t(x, y) &= \sum_{j=1}^n \lambda_j(\Gamma_\theta)^{\lceil \frac{t}{k} \rceil} \hat{\psi}_j(x) \bar{\psi}_j(y). \end{aligned}$$

Let C and c be strictly positive constants and define $K_{\theta c}(t)$ to be zero if $k \leq h^2/36$ or $t < k$, and otherwise

$$(63) \quad K_{\theta c}(t) = \begin{cases} 0 & \text{if } t \geq k, \theta = 0 \\ \frac{\sqrt{t}}{h} e^{-\frac{ct h^2}{k^2}} & \text{if } t \geq k, \theta = \frac{1}{2}, \\ \frac{\sqrt{t}}{h} e^{-\frac{ct}{k}} & \text{if } t \geq k, \frac{1}{2} < \theta < 1 \\ \frac{\sqrt{t}}{h} \wedge \sqrt{\frac{k}{t}} & \text{if } t \geq k, \theta = 1 \end{cases}$$

Theorem 8.3 *Suppose that either $\theta = 0$ and $k < (1 - \varepsilon)h^2/6$ for some $\varepsilon > 0$, or $\frac{1}{2} \leq \theta \leq 1$. Then there exists constants c and C , depending only on θ and ε such that for all $x \in L_0$, all $k > 0$, $h > 0$ and $t > 0$, for $\theta = 0$ or $\frac{1}{2} \leq \theta \leq 1$:*

$$(64) \quad \begin{aligned} \|\check{G}_t(x, \cdot)\|_2^2 &\leq \frac{C}{\sqrt{t \vee h^2}}; & \|\check{G}_t(x, \cdot)\|_\infty &\leq \frac{C}{\sqrt{t \vee h^2}}; \\ \|\tilde{G}_t(x, \cdot)\|_2^2 &\leq \frac{C}{\sqrt{t \vee h^2}}; & \|\tilde{G}_t(x, \cdot)\|_\infty &\leq \frac{C}{\sqrt{t \vee h^2}}; \end{aligned}$$

$$(65) \quad \begin{aligned} \|\hat{G}_t(x, \cdot) - \check{G}_t^\theta(x, \cdot)\|_2^2 &\leq \frac{C}{\sqrt{t}} \left(\sqrt{\frac{t}{k \vee h^2}} \wedge \frac{k^2 \vee h^4}{t^2} + K_{\theta c} \right); \\ \|\hat{G}_t(x, \cdot) - \check{G}_t^\theta(x, \cdot)\|_\infty &\leq \frac{C}{\sqrt{t}} \left(\sqrt{\frac{t}{k \vee h^2}} \wedge \frac{k \vee h^2}{t} + K_{\theta c} \right); \end{aligned}$$

$$(66) \quad \begin{aligned} \|\hat{G}_t(x, \cdot) - \check{G}_{t+k}^\theta(x, \cdot)\|_2^2 &\leq \frac{C}{\sqrt{t}} \left(\sqrt{\frac{t}{k \vee h^2}} \wedge \frac{k^2 \vee h^4}{t^2} + K_{\theta c} \right); \\ \|\hat{G}_t(x, \cdot) - \check{G}_{t+k}^\theta(x, \cdot)\|_\infty &\leq \frac{C}{\sqrt{t}} \left(\sqrt{\frac{t}{k \vee h^2}} \wedge \frac{k \vee h^2}{t} + K_{\theta c} \right); \end{aligned}$$

$$(67) \quad \begin{aligned} \|\check{G}_t^\theta(x, \cdot) - \tilde{G}_t^\theta(x, \cdot)\|_2^2 &\leq \frac{C}{\sqrt{t}} \left(\sqrt{\frac{t}{h^2}} \wedge \frac{h^2}{t} + K_{\theta c}(t) \right); \\ \|\check{G}_t^\theta(x, \cdot) - \tilde{G}_t^\theta(x, \cdot)\|_\infty &\leq \frac{C}{\sqrt{t}} \left(\sqrt{\frac{t}{h^2}} \wedge \frac{h}{\sqrt{t}} + K_{\theta c}(t) \right); \end{aligned}$$

$$(68) \quad \begin{aligned} \|\tilde{G}_t^\theta((i+1)h, \cdot) - \tilde{G}_t^\theta(ih, \cdot)\|_2^2 &\leq \frac{C}{\sqrt{t}} \left(\sqrt{\frac{t}{h^2}} \wedge \frac{h^2}{t} + K_{\theta c}(t) \right); \\ \|\tilde{G}_t^\theta((i+1)h, \cdot) - \tilde{G}_t^\theta(ih, \cdot)\|_\infty &\leq \frac{C}{\sqrt{t}} \left(\sqrt{\frac{t}{h^2}} \wedge \frac{h}{\sqrt{t}} + K_{\theta c}(t) \right); \end{aligned}$$

PROOF. We will first prove all the L^2 bounds. If $t < k$, $\check{G}_t = \tilde{G}_t = \hat{G}_0$, so let $t = mk$ for some positive integer m , and let $\hat{\lambda}_j \stackrel{\text{def}}{=} \lambda_j(Q)$. All eigenvalues of Q are eigenvalues of the identity, and $\Gamma_\theta = I - kQ(I + \theta kQ)^{-1}$ so that

$$\lambda_j(\Gamma_\theta) = 1 - \frac{k\hat{\lambda}_j}{1 + \theta k\hat{\lambda}_j}.$$

Note that if either $\theta = 0$ and $k\hat{\lambda}_j \leq 2$, or if $\frac{1}{2} \leq \theta \leq 1$, that $|\lambda_j(\Gamma_\theta)| \leq 1$. From (62),

$$(69) \quad \|\check{G}_t(x, \cdot)\|_2^2 \leq \frac{6}{L} \sum_{j=1}^{\frac{L}{h}-1} \lambda_j(\Gamma_\theta)^{2\frac{t}{k}} \leq \frac{6}{h}.$$

A quick application of the triangle inequality shows that (64) follows from this and the proof of (65) plus the estimate (53).

Let us prove (65). This comes down to a series of calculations with the eigenvectors and eigenvalues of the matrices $e^{-tQ} - \Gamma_\theta^{t/k}$. If $t < k$, $\check{G}_t = \check{G}_t = \hat{G}_0$, so let $t = mk$, for some integer $m \geq 1$. Then

$$\hat{G}_t(x, y) - \check{G}_t(x, y) = \Phi^T(x)M^{-1} (e^{-tQ} - \Gamma_\theta^m) \Phi(x) = \sum_{j=1}^n \lambda_j(e^{-tQ} - \Gamma_\theta^m) \hat{\psi}_j(x) \hat{\psi}_j(y).$$

Let $\eta_j = \eta_j(h, k, t, \theta) \stackrel{\text{def}}{=} \lambda_j(e^{-tQ} - \Gamma_\theta^{t/k})$ be the j th eigenvalue of $e^{-tQ} - \Gamma_\theta^{t/k}$. From Proposition 9.4 of the appendix,

$$(70) \quad \|\hat{G}_t(x, y) - \check{G}_t(x, y)\|_2^2 \leq \frac{6}{L} \sum_{j=1}^n \eta_j^2.$$

Now

$$\eta_j = e^{-t\hat{\lambda}_j} - \left(1 - \frac{k\hat{\lambda}_j}{1 + \theta k\hat{\lambda}_j}\right)^{\frac{t}{k}}.$$

We break up the sum into the sum over $k\hat{\lambda}_j \leq 1/3$ and $k\hat{\lambda}_j > 1/3$:

$$\sum_j \eta_j^2 = \sum_{k\hat{\lambda}_j \leq \frac{1}{3}} \eta_j^2 + \sum_{k\hat{\lambda}_j > \frac{1}{3}} \eta_j^2 \stackrel{\text{def}}{=} S_1 + S_2.$$

First consider S_1 . Note that

$$(71) \quad \eta_j^2 = \left(e^{-\hat{\lambda}_j t} - e^{\frac{t}{k} \log\left(1 - \frac{k\hat{\lambda}_j}{1 + \theta k\hat{\lambda}_j}\right)} \right)^2.$$

Expand the logarithm in the exponent in a Taylor series in $k\hat{\lambda}_j$, which we can do since $0 \leq k\hat{\lambda}_j \leq 1/2$. The exponent is

$$(72) \quad -k\hat{\lambda}_j^2 t \left(\theta - \frac{1}{2} + (\theta^2 - \theta + \frac{1}{3})k\hat{\lambda}_j + \dots \right).$$

For fixed $t > 0$, $0 \leq \theta \leq 1$, the term in parentheses is a power series in $k\hat{\lambda}_j$ with radius of convergence $1/(1+\theta) \geq 1/2$, hence for fixed $t > 0$ it is uniformly bounded in $0 \leq \theta \leq 1$, $|k\hat{\lambda}_j| \leq \frac{1}{3}$. Thus there is a constant c such that $\eta_j^2 \leq e^{-2\hat{\lambda}_j t} \left(1 - e^{-ck\hat{\lambda}_j^2 t}\right)^2$ which is in turn smaller than $c^2 e^{-2\hat{\lambda}_j t} k^2 \hat{\lambda}_j^4 t^2$. Using the bounds on $\hat{\lambda}_j$ from Proposition 9.2 (iv), we see there are C and c such that

$$(73) \quad \eta_j^2 \leq Ck^2 t^2 j^8 e^{-ctj^2}, \quad 0 \leq \theta \leq 1, \quad k\hat{\lambda}_j \leq \frac{1}{3}, \quad t = mk, \quad m = 1, 2, \dots$$

If $\theta = \frac{1}{2}$, we can sharpen this: the constant term vanishes, so (72) is $O(k^2 \hat{\lambda}_j^3 t)$, and

$$(74) \quad \eta_j^2 \leq Ck^4 j^{12} t^2 e^{-ctj^2}, \quad \theta = \frac{1}{2}, \quad k\hat{\lambda}_j \leq \frac{1}{3}, \quad t = mk, \quad m = 1, 2, \dots$$

We sum the η_j^2 over j for which $k\hat{\lambda}_j \leq \frac{1}{3}$, or $j \leq \frac{L}{\sqrt{12k}}$, so

$$(75) \quad \sum_{k\hat{\lambda}_j \leq \frac{1}{3}} \eta_j^2 \leq \begin{cases} Ck^2 t^2 \sum_{j=1}^{\frac{L}{h} \wedge \frac{L}{\sqrt{12k}}} j^8 e^{-ctj^2} & \text{if } 0 \leq \theta \leq 1, \quad t = mk, \quad m = 1, 2, \dots \\ Ck^4 t^2 \sum_{j=1}^{\frac{L}{h} \wedge \frac{L}{\sqrt{12k}}} j^{12} e^{-ctj^2} & \text{if } \theta = \frac{1}{2}, \quad t = mk, \quad m = 1, 2, \dots \end{cases}$$

Thus by Lemma 8.1 we see that there are constants c and C such that the sum is

$$(76) \quad \leq \begin{cases} Ck^2 t^2 \left(\frac{1}{h \vee \sqrt{k}} \wedge \frac{1}{\sqrt{t}} \right)^9 & \text{if } \theta \neq \frac{1}{2}, \\ Ck^4 t^2 \left(\frac{1}{h \vee \sqrt{k}} \wedge \frac{1}{\sqrt{t}} \right)^{13} & \text{if } \theta = \frac{1}{2}. \end{cases}$$

Since $t \geq k$ here, this implies the simpler upper bound:

$$(77) \quad S_1 \leq \begin{cases} \frac{C}{\sqrt{t}} \frac{k^2}{t^2} & \text{if } t \geq k, \theta \neq \frac{1}{2}, \\ \frac{C}{\sqrt{t}} \frac{k^4}{t^4} & \text{if } t \geq k, \theta = \frac{1}{2}. \end{cases}$$

Now consider S_2 , noting that since $\hat{\lambda}_j \leq \frac{12}{h^2}$, this sum is empty unless $k > \frac{h^2}{24}$. Then

$$(78) \quad S_2 = \sum_{k\hat{\lambda}_j > \frac{1}{3}} \left(e^{-\hat{\lambda}_j t} - \left(1 - \frac{k\hat{\lambda}_j}{1 + \theta k\hat{\lambda}_j} \right)^{\frac{t}{k}} \right)^2$$

$$(79) \quad \leq 2 \sum_{k\hat{\lambda}_j > \frac{1}{3}} e^{-2\hat{\lambda}_j t} + 2 \sum_{k\hat{\lambda}_j > \frac{1}{3}} \left| 1 - \frac{k\hat{\lambda}_j}{1 + \theta k\hat{\lambda}_j} \right|^{\frac{2t}{k}}$$

$$(80) \quad \stackrel{\text{def}}{=} U_1 + U_2.$$

Note that $\hat{\lambda}_j \geq \frac{4j^2}{L^2}$ so that by (47)

$$U_1 \leq C \sum_{j=\frac{L}{\sqrt{12k}}}^{\infty} e^{-\frac{4tj^2}{L^2}} \leq \frac{C}{\sqrt{t}} \left(1 \wedge \sqrt{\frac{k}{t}} e^{-\frac{t}{2k}} \right).$$

Since $t \geq k$ we see that $U_1 = 0$ if $k < h^2/24$ and

$$(81) \quad U_1 \leq \frac{C}{\sqrt{t}} \sqrt{\frac{k}{t}} e^{-\frac{ct}{k}}, \quad t \geq k \geq \frac{h^2}{24}.$$

Notice that for $\theta \neq 0$, $U_1 \leq CK_{\theta c}(t)$ if $t \geq k \geq h^2/24$. If $\theta = 0$ on the other hand, we have $k < h^2/6$ so $U_1 \leq C(k^2 \vee h^4)/t^2$ for some $C > 0$.

Now consider U_2 . If $k\hat{\lambda}_j > \frac{1}{3}$, then if $\theta > \frac{1}{2}$,

$$\frac{3}{4} \geq 1 - \frac{k \max_j \hat{\lambda}_j}{1 + \theta \max_j \hat{\lambda}_j} > 1 - \frac{1}{\theta} > -1$$

while if $\theta = 0$ and $\frac{k}{h^2} < \frac{1-\varepsilon}{6}$,

$$\frac{2}{3} \geq 1 - k\hat{\lambda}_j \geq 1 - k \max_j \hat{\lambda}_j \geq -1 + 2\varepsilon$$

since $\max_j \hat{\lambda}_j \leq 12/h^2$. Thus if $\theta = 0$ or if $\frac{1}{2} < \theta < 1$, there is $\varepsilon > 0$ such that

$$(82) \quad \left| 1 - \frac{k\hat{\lambda}_j}{1 + \theta k\hat{\lambda}_j} \right| \leq 1 - 2\varepsilon, \quad \forall j,$$

As there are at most L/h terms in the sum,

$$(83) \quad U_2 \leq \frac{L}{h}(1 - 2\varepsilon)^{\frac{2t}{k}} \leq \frac{L}{h}e^{-4\varepsilon\frac{t}{k}}, \quad \theta = 0 \text{ or } \frac{1}{2} < \theta < 1.$$

If $\theta = 1$, and $t \geq k$,

$$U_2 = \sum_{k\hat{\lambda}_j > \frac{1}{3}} \frac{1}{(1 + k\hat{\lambda}_j)^{2[\frac{t}{k}]}}$$

As $\hat{\lambda}_j \sim j^2$, this is bounded by the sum over j from $1/\sqrt{k}$ to infinity, which can be estimated by the integral

$$\begin{aligned} \int_{\frac{1}{\sqrt{k}}}^{\infty} \frac{dx}{(1 + kx^2)^{2[\frac{t}{k}]}} &= \frac{1}{\sqrt{k}} \int_1^{\infty} \frac{du}{(1 + u^2)^{2[\frac{t}{k}]}} \\ &\leq \frac{1}{\sqrt{k}} \int_1^{\infty} \frac{du}{u^{4[\frac{t}{k}]}} \end{aligned}$$

It follows that

$$(84) \quad U_2 \leq C \frac{\sqrt{k}}{t}, \quad \text{if } \theta = 1, t \geq k.$$

If $\theta = 1/2$ then $1 - \frac{1}{\theta} = -1$ so (82) may not hold. Indeed, for large values of j , the eigenvalue $\lambda_j(\Gamma_\theta)$ is close to -1 , so the higher modes decay slowly. In fact, it is easy to see that (82) does indeed hold if k/h^2 is bounded, but if not, we have to consider the contribution for large values of $k\hat{\lambda}_j$ separately. Consider the case in which $\sup_j k\hat{\lambda}_j$ is unbounded as h and $k \rightarrow 0$, or equivalently, where k/h^2 is unbounded. If $k\hat{\lambda}_j > 2$,

$$\left| 1 - \frac{k\hat{\lambda}_j}{1 + \frac{1}{2}k\hat{\lambda}_j} \right| = 1 - \frac{4}{2 + k\hat{\lambda}_j} \leq 1 - \frac{2}{k\hat{\lambda}_j},$$

so the contribution is

$$\begin{aligned}
&\leq \sum_{k\hat{\lambda}_j > 2} \left| 1 - \frac{2}{k\hat{\lambda}_j} \right|^{\frac{2t}{k}} \\
&\leq \sum_{k\hat{\lambda}_j > 2} e^{-\frac{4t}{k^2\hat{\lambda}_j}}.
\end{aligned}$$

Now $\hat{\lambda}_j \leq 12/h^2$ and there are fewer than L/h terms in the sum, so $U_2 \leq \frac{C}{h} e^{-\frac{th^2}{3k^2}}$. Put this together with (83):

$$(85) \quad U_2 \leq \frac{C}{\sqrt{t}} K_{\theta c}(t), \quad t \geq k,$$

and, adding in (81): $S_2 = 0$ if $k/h^2 < 1/36$ and there exist positive constants C and c (depending on θ) such that

$$(86) \quad S_2 \leq \frac{C}{\sqrt{t}} K_{\theta c}(t), \quad t \geq k, \quad \theta = 0 \text{ or } \frac{1}{2} \leq \theta < 1.$$

Adding in (77), we see that

$$\|\hat{G}_0(x, \cdot) - \check{G}_0(x, \cdot)\|_2^2 = 0,$$

$$(87) \quad \|\hat{G}_t(x, \cdot) - \check{G}_t(x, \cdot)\|_2^2 \leq \begin{cases} \frac{C}{\sqrt{t}} \left(\frac{k^2}{t^2} + K_{\theta c}(t) \right) & \text{if } t = mk, \quad m = 1, 2, \dots, \quad \theta \neq \frac{1}{2}, 1, \\ \frac{C}{\sqrt{t}} \left(\frac{k^4}{t^4} + K_{\theta c}(t) \right) & \text{if } t = mk, \quad m = 1, 2, \dots, \quad \theta = \frac{1}{2} \\ \frac{C\sqrt{k}}{t} & \text{if } t \geq k, \quad \theta = 1. \end{cases}$$

This takes care of the discrete values of t . Since \check{G}_t is constant on each interval $[mk, (m+1)k)$, we must interpolate to get (65).

Notice that $\hat{G}_0 = \check{G}_0$. If $t \in [mk, (m+1)k)$ for some integer $m \geq 0$,

$$(88) \quad \|\hat{G}_t - \check{G}_t\|_2^2 \leq 2\|\hat{G}_t - \hat{G}_{mk}\|_2^2 + 2\|\hat{G}_{mk} - \check{G}_{mk}\|_2^2.$$

We have dealt with the second term in (87). We can write the first term on the right-hand

side of (88) in the form

$$\begin{aligned}
(89) \quad \|\Phi^T(x)M^{-1}(e^{-mkQ} - e^{-tQ})\Phi(\cdot)\|_2^2 &\leq \frac{6}{L} \sum_{j=1}^{\frac{L}{h}-1} \lambda_j (e^{-mkQ} - e^{-tQ})^2. \\
&= \frac{6}{L} \sum_{j=1}^{\frac{L}{h}-1} (e^{-mk\hat{\lambda}_j} - e^{-t\hat{\lambda}_j})^2 \\
&= \frac{6}{L} \sum_{j=1}^{\frac{L}{h}-1} e^{-2t\hat{\lambda}_j} (1 - e^{-(t-mk)\hat{\lambda}_j})^2 \\
&\leq \frac{6}{L} \sum_{j=1}^{\frac{L}{h}-1} e^{-\frac{8tj^2}{L^2}} \left(1 - e^{-\frac{3\pi^2 j^2 (t-mk)}{L^2}}\right)^2,
\end{aligned}$$

where we have used Lemma 9.4 and the bounds (96) on $\hat{\lambda}_j$. Now we note that $1 - e^{-x} \leq x \wedge 1$ for $x \geq 0$, and that $t - mk \leq k$,

$$\leq \frac{6}{L} \sum_{j=1}^{\frac{L}{h}-1} e^{-\frac{8tj^2}{L^2}} \left(\frac{3\pi^2 j^2 k}{L^2} \wedge 1\right)^2.$$

Absorb the constants in C and c . If $h > \sqrt{k}$, this is

$$(90) \quad \leq Ck^2 \sum_{j=1}^{\frac{L}{h}-1} j^4 e^{-ctj^2}.$$

Use (48) and substitute $h^2 > k$ to see this is

$$\leq \frac{C}{\sqrt{t}} \left(\sqrt{\frac{t}{h^2}} \wedge \frac{h^4}{t^2}\right).$$

On the other hand, if $h < \sqrt{k}$, break (90) into two sums:

$$\leq Ck^2 \sum_{j=1}^{\frac{1}{\sqrt{k}}-1} j^4 e^{-ctj^2} + \sum_{j=\frac{1}{\sqrt{k}}}^{\infty} e^{-ctj^2}.$$

Use (48) and (47) respectively on the two sums, and remember that the second sum is always bounded by C/h to see this is

$$\leq \frac{C}{\sqrt{t}} \left(\sqrt{\frac{t}{k}} \wedge \frac{k^2}{t^2}\right) + \frac{C}{\sqrt{t}} \left(\sqrt{\frac{t}{h^2}} \wedge 1 \wedge \sqrt{\frac{k}{t}} e^{-\frac{ct}{k}}\right).$$

By increasing C , this is

$$\leq \frac{C}{\sqrt{t}} \left(\sqrt{\frac{t}{k}} \wedge \frac{k^2}{t^2} \right).$$

Combining the two cases,

$$(91) \quad \|\hat{G}_t(x, \cdot) - \hat{G}_{mk}(x, \cdot)\|_2^2 \leq \frac{C}{\sqrt{t}} \left(\sqrt{\frac{t}{k \vee h^2}} \wedge \frac{k^2 \vee h^4}{t^2} \right), \quad |t - mk| \leq k,$$

and (65) follows. To see (66), simply note that $\check{G}_{t+k} = \check{G}_{(m+1)k}$ so that $\|\hat{G}_t - \check{G}_{t+k}\| \leq \|\hat{G}_t - \hat{G}_{(m+1)k}\| + \|\hat{G}_{(m+1)k} - \check{G}_{(m+1)k}\|$, which gives us exactly the same bound as before.

To prove (67), write $\|\check{G}_t(x, \cdot) - \tilde{G}_t(x, \cdot)\|_2^2$ in matrix form:

$$\begin{aligned} &= \int_{L_0} \Phi^T(x) M^{-1} \Gamma_\theta^{\lfloor \frac{t}{k} \rfloor} (\Phi(y) - \bar{\Phi}(y)) (\Phi(y) - \bar{\Phi}(y))^T \Gamma_\theta^{\lfloor \frac{t}{k} \rfloor} M^{-1} \Phi(x) dy \\ &= \Phi^T(x) M^{-1} \Gamma_\theta^{\lfloor \frac{t}{k} \rfloor} \langle \Phi - \bar{\Phi}, \Phi^T - \bar{\Phi}^T \rangle \Gamma_\theta^{\lfloor \frac{t}{k} \rfloor} M^{-1} \Phi(x). \end{aligned}$$

Now $\langle \Phi, \Phi^T \rangle = M$ and $\langle \bar{\Phi}, \bar{\Phi}^T \rangle = \langle \Phi, \bar{\Phi}^T \rangle = \langle \bar{\Phi}, \Phi^T \rangle = \bar{M}$ so this is

$$= \Phi^T(x) M^{-1} \Gamma_\theta^{\lfloor \frac{t}{k} \rfloor} (M - \bar{M}) \Gamma_\theta^{\lfloor \frac{t}{k} \rfloor} M^{-1} \Phi(x).$$

Note that M , \bar{M} , and Γ_θ have the same eigenvectors, so we can apply Lemma 9.4 with $\Lambda = \Gamma_\theta^{\lfloor \frac{t}{k} \rfloor} (M - \bar{M}) \Gamma_\theta^{\lfloor \frac{t}{k} \rfloor}$ to see this is

$$(92) \quad \leq \frac{6}{L} \sum_{j=1}^{\frac{t}{h}-1} \lambda_j(\Gamma_\theta)^{2\lfloor \frac{t}{k} \rfloor} \frac{\lambda_j(M - \bar{M})}{\lambda_j(M)}.$$

Now $\lambda_j(M - \bar{M}) = \frac{h}{3} \sin^2 \frac{\pi j h}{2L}$ while $\lambda_j(M) \geq \frac{h}{3}$, so $\lambda_j(M - \bar{M})/\lambda_j(M) \leq \sin^2 \frac{\pi j h}{2L} \leq \frac{\pi^2 j^2 h^2}{4L^2}$. Thus this is

$$(93) \quad \leq \frac{3\pi^2 h^2}{2L^3} \sum_{j=1}^{\frac{t}{h}-1} j^2 \left(1 - \frac{k \hat{\lambda}_j}{1 + \theta k \hat{\lambda}_j} \right)^{2\lfloor \frac{t}{k} \rfloor}.$$

If $0 \leq t < k$, $\lfloor \frac{t}{k} \rfloor = 0$, and this is bounded by C/h . If $t = mk$ for some integer $m > 1$, we sum over $k \hat{\lambda}_j \leq 1/3$ and $k \hat{\lambda}_j > 1/3$ separately. If $k \hat{\lambda}_j \leq 1/3$ and $0 \leq \theta \leq 1$, then

$2/3 \leq 1 - k\hat{\lambda}_j(1 + \theta k\hat{\lambda}_j)^{-1} \leq 1 - 3k\hat{\lambda}_j/4$, so

$$\begin{aligned}
(94) \quad h^2 \sum_{k\hat{\lambda}_j \leq \frac{1}{3}} j^2 \left(1 - \frac{k\hat{\lambda}_j}{1 + \theta k\hat{\lambda}_j}\right)^{\frac{2t}{k}} &\leq h^2 \sum_{k\hat{\lambda}_j \leq \frac{1}{3}} j^2 e^{-\frac{3t\hat{\lambda}_j}{2}} \\
&\leq h^2 \sum_{j=1}^{\frac{L}{h}-1} j^2 e^{-\frac{6tj^2}{L^2}} \\
&\leq \frac{C}{\sqrt{t}} \left(\sqrt{\frac{t}{h^2}} \wedge \frac{h^2}{t} \right).
\end{aligned}$$

Next, since $h^2 j^2 \leq L^2$, by (85) the sum over $k\hat{\lambda}_j > \frac{1}{3}$ is

$$(95) \quad \leq \frac{3\pi^3}{2L} \sum_{k\hat{\lambda}_j > \frac{1}{3}} \left(1 - \frac{k\hat{\lambda}_j}{1 + \theta k\hat{\lambda}_j}\right)^{\frac{2t}{k}} \leq \frac{3\pi^2}{4L} U_2 \leq \frac{C}{\sqrt{t}} K_{\theta c}(t).$$

Put (94) and (95) together to get (67).

Finally, for (68), let $x_i = ih$ and write

$$\tilde{G}_t(x_{i+1}, y) - \tilde{G}_t(x_i, y) = (\Phi^T(x_{i+1}) - \Phi^T(x_i)) M^{-1} \Gamma_\theta^{\lceil \frac{t}{k} \rceil} \bar{\Phi}(y),$$

from which we see, as before

$$\begin{aligned}
\|\tilde{G}_t(x_{i+1}, \cdot) - \tilde{G}_t(x_i, \cdot)\|_2^2 &= (\Phi(x_{i+1}) - \Phi(x_i))^T M^{-1} \bar{M} \Gamma_\theta^{2\lceil \frac{t}{k} \rceil} (\Phi(x_{i+1}) - \Phi(x_i)) \\
&= \sum_{j=1}^{\frac{L}{h}-1} \frac{\lambda_j(\bar{M})}{\lambda_j(M)} \lambda_j(\Gamma_\theta)^{2\lceil \frac{t}{k} \rceil} (\hat{\psi}_j(x_{i+1}) - \hat{\psi}_j(x_i))^2.
\end{aligned}$$

From Proposition 9.2, $\lambda_j(M) \geq h/3$ and $\lambda_j(\bar{M}) = h \cos^2 \frac{j\pi h}{2L}$. Then $|\hat{\psi}_t(x_{i+1}) - \hat{\psi}_t(x_i)| \leq \hat{\psi}'(x_i)h + O(h^2) \leq \sqrt{\frac{6}{L}} \frac{\pi j}{L} h$. Making free with constants, the above is

$$\leq Ch^2 \sum_{j=1}^{\frac{L}{h}-1} j^2 \lambda_j(\Gamma_\theta)^{2\lceil \frac{t}{k} \rceil}.$$

We dealt with this sum in (93) and found it was

$$\leq \frac{C}{\sqrt{t}} \left(\left(\sqrt{\frac{t}{h^2}} \wedge \frac{h^2}{t} \right) + K_{\theta c}(t) \right),$$

which implies (68). This completes the proof of the L^2 inequalities.

The L^∞ bounds follow from roughly the same calculations, but are somewhat simpler. From (62), $\|\check{G}_t\|_\infty \leq 6/h$, and (64) follows from this, (65) and (53).

To see (64), take $t = mk$, $m = 1, 2, \dots$. Note that $\|\hat{G}_t - \check{G}_t\|_\infty \leq \frac{6}{L} \sum_{j=1}^n |\eta_j|$, so that we must evaluate the same sum as before, with η_j^2 replaced by $|\eta_j|$. Following the path to (77), we see that

$$\sum_{k\hat{\lambda}_j \leq \frac{1}{3}} |\eta_j| \leq \begin{cases} \frac{C}{\sqrt{t}} \frac{k}{t} & \text{if } \theta \neq \frac{1}{2}, t = mk, m = 1, 2, \dots \\ \frac{C}{\sqrt{t}} \frac{k^2}{t^2} & \text{if } \theta = \frac{1}{2}, t = mk, m = 1, 2, \dots \end{cases}$$

Then

$$\sum_{k\hat{\lambda}_j > \frac{1}{3}} |\eta_j| \leq \sum_{k\hat{\lambda}_j > \frac{1}{3}} e^{-\hat{\lambda}_j t} + \sum_{k\hat{\lambda}_j > \frac{1}{3}} \left| 1 - \frac{k\hat{\lambda}_j}{1 + \theta k\hat{\lambda}_j} \right|^{\frac{t}{k}}.$$

which leads to the same sums as in (79) with t replaced by $t/2$. Thus it is bounded by $\frac{C}{\sqrt{t}} K_{\theta c}(t)$ for some c and C . This gives us (64) for t a multiple of k . To extend it to all t , follow (88):

$$\|\hat{G}_t - \hat{G}_{mk}\|_\infty \leq \frac{6}{L} \sum_{j=1}^n |e^{-mk\hat{\lambda}_j} - e^{-t\hat{\lambda}_j}|.$$

We estimated this above. Taking $|t = mk| \leq k$, we see this is

$$\leq ck \sum_{j=1}^n j^2 e^{-ctj^2}.$$

Just as before, we take the cases $h > \sqrt{k}$ and $h \leq \sqrt{k}$ separately, and find that this is

$$\leq \frac{C}{\sqrt{t}} \left(\sqrt{\frac{t}{k \vee h^2}} \wedge \frac{k \vee h^2}{t} \right).$$

The L^∞ bounds in (65) and (66) now follow from this and the fact that $\|\hat{G}_t - \check{G}_t\|_\infty \leq \|\hat{G}_t - \hat{G}_{mk}\|_\infty + \|\hat{G}_{mk} - \check{G}_{mk}\|_\infty$.

For the L^∞ bounds in (67), use (62):

$$\|\check{G}_t - \tilde{G}_t\| \leq \sqrt{\frac{6}{L}} \sum_{j=1}^n |\lambda_j(\Gamma_\theta)| \|\hat{\psi}_j - \bar{\psi}_j\|_\infty.$$

Now if $x_i = ih$, then $\bar{\psi}_j = \frac{1}{2}|\hat{\psi}_j(x_{i+1}) + \psi_j(x_i)|$ on (x_i, x_{i+1}) , so $|\hat{\psi}_j(x) - \bar{\psi}_j(x)| \leq \frac{1}{2}|\hat{\psi}_j(x_{i+1}) - \bar{\psi}_j(x_i)| \leq \sqrt{6/L} \pi j h/L$, as we saw, so that this is

$$\leq Ch \sum_{j=1}^n j |\lambda_j(\Gamma_\theta)|^{\frac{t}{k}},$$

which is dealt with as in (93). It is

$$\leq \frac{C}{\sqrt{t}} \left(\sqrt{\frac{t}{h^2}} \wedge \frac{h}{\sqrt{t}} \right) + \frac{C}{\sqrt{t}} K_{\theta c}(t).$$

Finally, the sum for (68) is

$$\|\check{G}_t(x_{i+1}, \cdot) - \tilde{G}_t(x_i, \cdot)\|_\infty \leq \sqrt{\frac{6}{L}} \sum_{j=1}^n |\lambda_j(\Gamma_\theta)| |\hat{\psi}_j(x_{i+1}) - \hat{\psi}_j(x_i)|,$$

which leads to the series we just handled, and gives the L^∞ bound in (67). This completes the proof. ♣

9 Appendix

9.1 The Matrix J .

Let J be the $n \times n$ tri-diagonal matrix with zeros along the main diagonal and ones along the diagonals above and below it. Most of the matrices we deal with in this paper are easily expressed in terms of J and the identity matrix I . Here are some basic facts.

Proposition 9.1 *The matrix J has eigenvalues*

$$(96) \quad \lambda_j(J) = 2 \cos \frac{j\pi}{n+1} = 2 - 4 \sin^2 \frac{j\pi}{2(n+1)}$$

and corresponding ortho-normal eigenvectors $F_j = (F_{j1}, \dots, F_{jn})^T$, $j = 1, \dots, n$ given by

$$(97) \quad F_{ji} = \sqrt{\frac{2}{n+1}} \sin \frac{\pi ij}{n+1} = \sqrt{\frac{2h}{L}} \sin \frac{\pi ijh}{L}.$$

PROOF. J is a $n \times n$ symmetric matrix, hence it has n orthogonal eigenvectors. If $f = (f(1), \dots, f(n))^T$ is an eigenvector of J , it will satisfy

$$f(m-1) + f(m+1) = \lambda f(m), \quad 2 \leq m \leq n-1,$$

while at 1 and n we have:

$$f(2) = \lambda f(1), \quad f(n-1) = \lambda f(n).$$

We solve this to find that the eigenfunctions f_j are given by

$$f_j(m) = \sin \frac{j\pi m}{(n+1)}, \quad j = 1, \dots, n,$$

and the corresponding eigenvalues are

$$\lambda_j(J) = 2 \cos \frac{j\pi}{n+1} = 2 - 4 \sin^2 \frac{j\pi}{2(n+1)}.$$

The eigenvalues are distinct, so the eigenfunctions are orthogonal. To normalize them, note that $\sum_{i=1}^n \sin^2 \frac{\pi ij}{n+1} = \frac{n+1}{2}$.



9.2 Matrices Associated to the Finite Element Functions

Using the notation of the first four sections, fix n , put $h = L_0/(n+1)$, and let $M = \langle \Phi, \Phi^T \rangle$, $\bar{M} = \langle \bar{\Phi}, \bar{\Phi}^T \rangle$, $N = \langle \Phi', \Phi'^T \rangle$, and $Q = M^{-1}N$. These are the covariance matrices of various functions, hence they are positive definite. From the definition of the finite element functions ϕ_i and $\bar{\phi}$ in Sections 3 and 4.5 respectively, we see that the supports of ϕ_i and ϕ_j intersect only if $j = i$, $i+1$, or $i-1$. Thus M , \bar{M} , and N are tri-diagonal $n \times n$ matrices. More specifically, $\langle \phi_i, \phi_j \rangle = h$ if $i = j$, $h/6$ if $j = i \pm 1$, and 0 otherwise. so that $M = (h/6)(4I + J)$. Similar calculations show

$$(98) \quad \begin{aligned} M &= \frac{h}{6}(4I + J) \\ \bar{M} &= \frac{h}{4}(2I + J) \\ N &= \frac{1}{h}(2I - J). \end{aligned}$$

It is clear from this that M , \bar{M} , N , and Q commute and have the same eigenvectors as J . The eigenvalues are easily calculated from (98). The following proposition summarizes their properties. For any real symmetric matrix S , let $\lambda_j(S)$, $\lambda_{\min}(S)$ and $\lambda_{\max}(S)$ be respectively the j th eigenvalue, the smallest and the largest eigenvalues of S .

Proposition 9.2 (i) M, \bar{M}, N and Q are positive definite symmetric $n \times n$ matrices; M, \bar{M} and N are tri-diagonal.

(ii) The matrices $J, M, \bar{M}, N,$ and Q commute and have the eigenvectors F_j given by (97).

(iii) The eigenvalues of $J, M, \bar{M}, N,$ and Q corresponding to the eigenvector F_j are

$$\begin{aligned}\lambda_j(J) &= 2 - 4 \sin^2 \frac{j\pi}{2(n+1)}; \\ \lambda_j(M) &= h \left(1 - \frac{2}{3} \sin^2 \frac{j\pi}{2(n+1)} \right); \\ \lambda_j(\bar{M}) &= h \cos^2 \frac{j\pi}{2(n+1)}; \\ \lambda_j(N) &= \frac{4}{h} \sin^2 \frac{j\pi}{2(n+1)}; \\ \lambda_j(Q) &= \frac{\lambda_j(N)}{\lambda_j(M)} = \frac{12}{h^2} \frac{\sin^2 \frac{j\pi h}{2L}}{3 - 2 \sin^2 \frac{j\pi h}{2L}} \quad j = 1, \dots, n.\end{aligned}$$

(iv) The eigenvalues of $M, \bar{M}, N,$ and Q satisfy

$$\begin{aligned}\cos^2 \frac{j\pi}{2(n+1)} &\leq \frac{\lambda_j(\bar{M})}{\lambda_j(M)} \leq 3 \cos^2 \frac{j\pi}{2(n+1)}; \\ \frac{4j^2}{L^2} &\leq \lambda_j(Q) \leq \frac{3\pi^2 j^2}{L^2}; \\ \frac{\pi^2}{L^2} &< \lambda_{\min}(Q) < \lambda_{\max}(Q) < \frac{12}{h^2}.\end{aligned}$$

PROOF. The matrices M and \bar{M} and N are symmetric, positive definite, and of the form $aI + bJ$. Since all vectors are eigenvectors of the identity, if $b \neq 0$, a vector is an eigenvector of $aI + bJ$ iff it is an eigenvector of J , and clearly $\lambda_j(aI + bJ) = a + b\lambda_j(J)$. We gave the eigenvalues and eigenvectors of J in Proposition 9.1, so the expressions for the eigenvalues of L, M and Q are immediate. The rest of the proposition follows easily.

9.3 An Orthonormal Basis of \mathcal{S}_n

Define functions $\hat{\psi}_j(x)$ by

$$(99) \quad \hat{\psi}_j(x) = \lambda_j(M)^{-\frac{1}{2}} F_j^T \Phi(x), \quad j = 1, \dots, n.$$

Note that

$$\hat{\psi}_j(x) = F_j^T M^{-1/2} \Phi(x).$$

It follows that if we set $\Psi(x) = (\hat{\psi}_1(x), \dots, \hat{\psi}_n(x))^T$ and let \mathbf{F} be the matrix whose columns are F_1, \dots, F_n , then

$$\Psi(x) = \mathbf{F}^T M^{-1/2} \Phi(x) \quad \text{or} \quad \Phi(x) = M^{1/2} \mathbf{F} \Psi(x).$$

Proposition 9.3 (a) *The functions $\hat{\psi}_j$ are ortho-normal in $L^2(L_0)$, and form a basis of \mathcal{S}_n . Moreover,*

$$\|\hat{\psi}_j\|_\infty \leq \sqrt{\frac{6}{L}}.$$

(b) *If $x_i = ih$, then*

$$\hat{\psi}_j(x_i) = \sqrt{\frac{6}{L}} \frac{\sin \frac{\pi i j h}{L}}{\sqrt{1 + 2 \cos^2 \frac{j \pi h}{2L}}}, \quad j = 1, \dots, n, \quad i = 0, \dots, n+1.$$

PROOF. Note that

$$\langle \hat{\psi}_i, \hat{\psi}_j \rangle = \frac{1}{\sqrt{\lambda_i(M) \lambda_j(M)}} \int_{L_0} (F_i^T \Phi(y)) (\Phi^T(y) F_j) dy.$$

But $\int \Phi(y) \Phi^T(y) dy = M$, so this is

$$= \frac{F_i^T M F_j}{\sqrt{\lambda_i(M) \lambda_j(M)}} = \frac{\lambda_j(M) F_i^T F_j}{\sqrt{\lambda_i(M) \lambda_j(M)}} = \delta_{ij},$$

since the F_i are ortho-normal eigenvectors of M . Now $\hat{\psi}_j(x) = (\sum_i F_{ji} \phi_i(x)) / \sqrt{\lambda_j(M)}$, and for each x , $\phi_i(x) \neq 0$ for at most two values of i , while $\phi_i(x) \geq 0$ and $\phi_i(x) + \phi_{i+1}(x) \leq 1$. Thus the sum is dominated by

$$\frac{\max_i |F_{ji}|}{\sqrt{\lambda_j(M)}} \leq \sqrt{\frac{2h}{L}} \sqrt{\frac{3}{h}} = \sqrt{\frac{6}{L}}.$$

The rest follows from the explicit form for F_j . ♣

The next lemma deals with some expressions which are connected with various Green's functions. Let $(\hat{\psi}_j)$ be the orthonormal basis of \mathcal{S}_n defined in (99).

Lemma 9.4 *Let Λ be a symmetric $n \times n$ matrix with the same eigenvectors as J . Then*

$$(100) \quad \Phi^T(x)M^{-1}\Lambda\Phi(y) = \sum_j \lambda_j(\Lambda)\hat{\psi}_j(x)\hat{\psi}_j(y);$$

$$(101) \quad \Phi^T(x)M^{-1}\Lambda M^{-1}\Phi(x) = \sum_j \frac{\lambda_j(\Lambda)}{\lambda_j(M)}\hat{\psi}(x)^2 \leq \frac{6}{L} \sum_j \frac{\lambda_j(\Lambda)}{\lambda_j(M)}.$$

$$(102) \quad \int_{L_0} (\Phi^T(x)M^{-1}\Lambda\Phi(y))^2 dy = \sum_{j=1}^n \lambda_j^2(\Lambda)\hat{\psi}_j(x)^2 \leq \frac{6}{L} \sum_{j=1}^n \lambda_j^2(\Lambda).$$

PROOF. ΛM^{-1} has eigenvectors F_j and eigenvalues $\lambda_j(\Lambda)/\lambda_j(M)$, so we can expand it:

$$M^{-1}\Lambda = \sum_j \frac{\lambda_j(\Lambda)}{\lambda_j(M)} F_j F_j^T$$

so that

$$\Phi^T(x)M^{-1}\Lambda\Phi(y) = \sum_j \lambda_j(\Lambda) \frac{\Phi^T(x)F_j}{\lambda_j(M)^{1/2}} \frac{F_j^T\Phi(y)}{\lambda_j(M)^{1/2}} = \sum_j \lambda_j(\Lambda)\hat{\psi}_j(x)\hat{\psi}_j(y)$$

as claimed. Now (102) follows from the orthonormality of the $\hat{\psi}_j$ and Proposition 9.3 (a), and a similar calculation gives (101). ♣

References

- [1] Davie, A.M. and Gaines, J. G. Convergence of numerical schemes for the solution of parabolic partial differential equations. *Math. Comp.* 70, (2000) 121–134.
- [2] Gyöngy, I., Lattice approximations for stochastic quasi-linear parabolic partial differential equations driven by space-time white noise II, *Potential Analysis* 11 (1999), 1–37.
- [3] Gyöngy, I. and Nualart, D., Implicit schemes for quasi-linear parabolic partial differential equations perturbed by space-time white noise. *Stoch. Proc. and Appl.* 58, (1995). 57–72.
- [4] Gyöngy, I. and Nualart, D., Implicit schemes for quasi-linear parabolic partial differential equations driven by space-time white noise. *Potential Analysis* 7, (1997), 725–757.

- [5] Luskin, M. and Rannacher, R., On the smoothing property of the Galerkin method for parabolic equations, *SIAM J. Numer. Anal.* 19 (1982), 93–113
- [6] Walsh, J. B., *Introduction to stochastic partial differential equations*, Springer Verlag, LMN 1180, (1986) 256–437.