Illinois Journal of Mathematics
Volume 50, Number 4, Winter 2006, Pages 991–1018
S 0019-2082

ON NUMERICAL SOLUTIONS OF THE STOCHASTIC WAVE EQUATION

JOHN B. WALSH

ABSTRACT. We show that there is a numerical scheme for the stochastic wave equation which converges in L^p at a rate of $O(\sqrt{h})$, and which converges a.s. uniformly on compact sets at a rate $O(\sqrt{h|\log h|^{\varepsilon}})$, for any $\varepsilon > 0$, where h is the step size in both time and space. We show that this is the optimal rate: there is no scheme depending on the same increments of white noise which has a higher order of convergence.

1. Introduction

When one speaks of the numerical solution of stochastic PDEs, one usually, but not always, means their simulation. Although the SPDE may describe a physical system perturbed by noise, the noise itself is seldom observable, so that one can't use it as part of the solution. However, explicit solutions of SPDEs are rare and detailed calculations are difficult, so simulations are important. One can simulate the solution by generating increments of the random driving noise and putting these into a numerical scheme. But one must know if the simulation is good, which means at the very least finding bounds on the error and determining the rate of convergence of the scheme. (There is a deeper question which we shall not address in this article: to what extent does the simulation share the interesting sample-path properties of the true solution?) Needless to say, one does not want to generate more increments than necessary, and the accuracy of the simulation may depend on exactly which increments are generated. This will be an important point when we come to the question of lower bounds on the error.

The solutions of classical PDEs are generally smooth, and there are higher-order numerical schemes which take advantage of that smoothness to give higher-order convergence. In contrast, the solutions of SPDEs are often nowhere-differentiable, and there may be a limit on the rate of convergence. This was shown by Davie and Gaines [4] in the parabolic case. Effectively, there are no higher-order schemes, either explicit or implicit.

Received July 5, 2005; received in final form May 17, 2006.
2000 *Mathematics Subject Classification.* 60H35, 60H15.

Most experience to date comes from parabolic equations, and it suggests that the rate of convergence is governed by the continuity of the paths. The sample paths of the stochastic heat equation, for instance, are roughly Hölder(1/4), and, if k is the size of the time-step, the optimal rate of convergence as $k \to 0$ is $O(k^{1/4})$ [4]. Moreover, there are many schemes which attain this rate [5] [6].

In this article, we consider the stochastic wave equation, which is less studied. (See [8] for some references.) Its solutions are roughly Hölder(1/2), so that if h is the step size, we would expect the rate of convergence to be $O(h^{1/2})$. We show that, indeed, there is a scheme which attains that rate. Moreover, there is a limit on the rate of convergence of numerical schemes, and this rate is optimal: no scheme based on the same increments of white noise converges at a rate faster than $O(\sqrt{h})$.

Our scheme is an adaptation of what is called a "leapfrog" method. It a second order method, but we believe that its order is not important, and that many other schemes share the same rate of convergence.

Other schemes may contain surprises, however. Quer-Sardanyans and Sanz-Sole [8] have investigated the rate of convergence of a semi-discrete scheme (discrete in space, continuous in time, sometimes called the "method of lines"). One might expect this to be better than a fully discrete method, for time-continuity is tantamount to setting the time-step equal to zero. Interestingly enough, it is not better, it is worse: they show that it converges at the unexpectedly slow rate of $O(h^{1/3})$. (More exactly, they prove that it converges at least that fast, and make a strong numerical argument that it is no faster.) It would be interesting to know whether time-discretizations of this method would converge faster.

Finally, let us mention the connection between numerical solutions of white-noise-driven SPDEs and the round-off error of numerical schemes for non-stochastic PDEs.

Each step in the numerical solution of a PDE involves a round-off error. That error can be regarded as random—as much as the output of any random number generator can be considered random—and by symmetry it will have mean zero. There will be an error attached to each step, which means that there is an error attached to each space-time cell, so that the round-off error can be considered as a noise in space-time. It is not hard to see that, suitably normalized, it approximates a white noise when the step size is small. Thus one can treat the round-off error as an autonomous white noise. For instance, if $g \equiv 0$, (1) is a PDE, and the numerical scheme for it can be treated as the exact (i.e., no round-off error) scheme for the SPDE with a particular non-zero g—the exact function depends on how f is calculated. The same remark applies to other types of SPDEs as well, of course.

2. The spde

Consider the non-linear wave equation in one space-dimension, perturbed by white noise. We will treat the case of an unbounded region in detail, and just indicate how to apply our results to bounded domains in §7. Let $u_0(x)$ and $v_0(x)$ be real-valued functions on $\mathbb{R}$ and let $f(x,t,u)$ and $g(x,t,u)$ be real-valued functions on $\mathbb{R}\times\mathbb{R}_+\times\mathbb{R}$, where $\mathbb{R}_+ = [0,\infty)$. Consider the initial-value problem:

$$\begin{cases} \dfrac{\partial^2 u}{\partial t^2} = \dfrac{\partial^2 u}{\partial x^2} + f(x,t,u) + g(x,t,u)\dot{W}\,, \quad x\in\mathbb{R},\ t>0\,, \\ u(x,0) = u_0(x)\,, \quad \dfrac{\partial u}{\partial t}(x,0) = v_0(x); \quad x\in\mathbb{R}\,, \end{cases} \tag{1}$$

where

- $\dot{W}$ is a standard white noise on $\mathbb{R}\times\mathbb{R}_+$, i.e., it is a Gaussian measure W on the Borel sets of $\mathbb{R}\times\mathbb{R}_+$ which has the property that $W(A)$ is $N(0,|A|)$ and if A and B are disjoint Borel sets, $W(A)$ and $W(B)$ are independent and $W(A)+W(B) = W(A\cup B)$.
- u_0 and v_0 are deterministic Hölder-continuous functions of order at least (1/2).
- f and g are Lipschitz continuous: for each N there is a constant L_N such that for all x, y, x' and $y' \in [-N,N]$ and all $z,\ z'\in\mathbb{R}$, and if $\xi = (x,y,z)$ and $\xi' = (x',y',z')$,

$$\begin{aligned} &|f(\xi)| + |g(\xi)| \le L_N(1+|\xi|)\,; \\ &|f(\xi)-f(\xi')| + |g(\xi)-g(\xi')| \le L_N|\xi-\xi'|\,. \end{aligned} \tag{2}$$

Let $C(x,t) = \{(y,s) : 0\le s\le t, |y-x|\le t-s\}$ be the backward light cone with apex (x,t). Then $G(x,t;y,s) \equiv (1/2)I_{C(x,t)}(y,s)$ is the Green's function for the wave equation, and (1) is shorthand for an integral equation: $u(x,t)$ is a solution of (1) if and only if it satisfies the following a.s. for each $(x,t)\in\mathbb{R}\times\mathbb{R}_+$:

$$\begin{aligned} u(x,t) = \frac{1}{2}\big(u_0(x-t)+u_0(x+t)\big) &+ \frac{1}{2}\int_{x-t}^{x+t} v_0(y)\,dy \\ &+ \int_{\mathbb{R}\times[0,t]} G(x,t;y,s) f(y,s,u(y,s))\,dy\,ds \\ &\quad + \int_{\mathbb{R}\times[0,t]} G(x,t;y,s) g(y,s,u(y,s))\,W(dy\,ds)\,. \end{aligned} \tag{3}$$

This is called the *mild form* of (1). It is well-known ([3] and Exercise 3.7 of [6]) that there exists a unique Hölder-continuous solution to this. We can rewrite (3):

$$
(4)\quad 2u(x,t) = u_0(x-t) + u_0(x+t) + \int_{x-t}^{x+t} v_0(y)\,dy + \int_{C(x,t)} f(y,s,u(y,s))\,dy\,ds + \int_{C(x,t)} g(y,s,u(y,s))\,W(dy\,ds)\,.
$$

Note that $u(x,t)$ is entirely determined by what happens in $C(x,t)$. Indeed, only the values of u_0 and v_0 inside $[x-t, x+t]$ and the restriction of $W(dy\,ds)$ to $C(x,t)$ appear in (4). Thus the entire problem can be restricted to the cone $C(x,t)$, and we do not need global conditions on u_0 and v_0, just local conditions.

We wish to find a numerical scheme to solve this problem. Observe that if we rotate coordinates by 45° by letting

$$
(5)\qquad \xi = (t+x)/\sqrt{2}\,, \quad \eta = (t-x)/\sqrt{2}\,,
$$

then

$$
\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} = 2\frac{\partial^2 u}{\partial \xi\,\partial \eta}\,.
$$

This change of variables preserves area, and we can write (1) in the form

$$
(6)\qquad 2\frac{\partial^2 u}{\partial \xi\,\partial \eta} = f + g\dot{W}\,.
$$

If we set $\hat{u}(\xi,\eta) = u(x,t)$, we can approximate the mixed partial derivative by the double difference $[\hat{u}(\xi+k,\eta+k) - \hat{u}(\xi,\eta+k) - \hat{u}(\xi+k,\eta) + \hat{u}(\xi,\eta)]/k^2$. Notice that this equals $k^{-2}\int \frac{\partial^2 u}{\partial\xi\partial\eta}\,d\xi\,d\eta$, where the integral is over the square of side k with lower-left-hand corner at (ξ,η). In terms of the original coordinates, let (x,t) be the center of that square, and let $h = k/\sqrt{2}$, so that the above double difference equals $[u(x,t+h) - u(x+h,t) - u(x-h,t) + u(x,t-h)]/2h^2$. Let Δ be the square—or perhaps we should say diamond—with corners at $(x,t+h)$, $(x+h,t)$, $(x-h,t)$, and $(x,t-h)$. Integrate both sides of (6) over Δ to see that:

$$
(7)\quad 2[u(x,t+h) - u(x+h,t) - u(x-h,t) + u(x,t-h)] = 2\int_{\Delta} \frac{\partial^2 u}{\partial\xi\partial\eta}\,d\xi\,d\eta = \int_{\Delta} f(y,s,u)\,dy\,ds + \int_{\Delta} g(y,s,u)\,W(dy\,ds)\,.
$$

In order to get a recurrence scheme, we must replace the integrals by discrete approximations. Denote the area of Δ by $|\Delta| = 2h^2$ and write

$$\int_\Delta f(y,s,u)\,dy\,ds \sim f\big(x,t,\tfrac{1}{2}(u(x+h,t)+u(x-h,t))\big)\,|\Delta|\,;$$
$$\int_\Delta g(y,s,u)\,W(dy\,ds) \sim g\big(x,t,\tfrac{1}{2}(u(x+h,t)+u(x-h,t))\big)\,W(\Delta)\,.$$

Note that these approximations are exact if f and g are constant. Thus we have

$$\begin{aligned}(8)\quad u(x,t+h) \sim{}& u(x+h,t)+u(x-h,t)-u(x,t-h)\\ &+\frac{1}{2}f\big(x,t,\tfrac{1}{2}(u(x+h,t)+u(x-h,t))\big)|\Delta|\\ &+\frac{1}{2}g\big(x,t,\tfrac{1}{2}(u(x+h,t)+u(x-h,t))\big)\,W(\Delta)\,,\end{aligned}$$

with equality if f and g are constant.

3. The difference scheme

Let $h>0$, put $x_i = ih$, $t_j = jh$, and define subsets $\mathcal{L}_h$ and $\mathcal{M}_h$ of $h\mathbb{Z}^2$ by

$$\mathcal{L}_h = \{(x_i,t_j) :\ i,j\in\mathbb{Z},\ ij \text{ is even}\},\quad \mathcal{M}_h = \{(x_i,t_j) :\ i,j\in\mathbb{Z},\ ij \text{ is odd}\}.$$

Thus, if $(x_i,t_j)\in\mathcal{L}_h$, i and j are either both even or both odd, so that the points of $\mathcal{L}_h$ on the x-axis are (x_0,t_0), $(x_{\pm2},t_0)$, $(x_{\pm4},t_0),\dots$, and the points on the line $t=h$ are $(x_{\pm1},t_1)$, $(x_{\pm3},t_1)\dots$ Thus $\mathcal{L}_h$ contains "every other point" of the lattice $h\mathbb{Z}^2$, $\mathcal{M}_h$ is the complementary lattice, and $\mathcal{L}_h\cup\mathcal{M}_h = h\mathbb{Z}^2$. See Figure 1 .

Let $u_{i,j}\sim u(x_i,t_j)$, and let $\Delta_{i,j}$ be the square with center (x_i,t_j) and corners at $(x_i,t_{j\pm1})$, $(x_{i\pm1},t_j)$. Then (8) suggests:

$$\begin{aligned}(9)\quad u_{i,j+1} ={}& u_{i+1,j}+u_{i-1,j}-u_{i,j-1}\\ &+h^2 f\big(x_i,t_j,\tfrac{1}{2}(u_{i+1,j}+u_{i-1,j})\big)\\ &+\frac{1}{2}g\big(x_i,t_j,\tfrac{1}{2}(u_{i+1,j}+u_{i-1,j})\big)\,W(\Delta_{ij})\,.\end{aligned}$$

Notice that if $(x_i,t_j)\in\mathcal{M}_h$, then all the coordinates appearing in (8) are in $\mathcal{L}_h$, so that this is in fact a recurrence relation on $\mathcal{L}_h$. If we know the values of $u_{i,j}$ if $(i,j)\in\mathcal{L}_h$, $j\le k$, for instance, we can get the values of $u_{i,k+1}$ for $(i,k+1)\in\mathcal{L}_h$. Note that we need to know $u_{\cdot,j}$ for at least two values of j to update. We are given the initial values of u and v. This gives us the values of $u_{i,0}$, but this is not enough to start the induction process: to get $u_{i,1}$ from (9) we need not only $u_{i,0}$ but $u_{i,-1}$ as well. To construct the values for $j=-1$ we extend u to the lower half plane as the solution of the homogeneous wave equation with the same initial conditions. By D'Alembert's formula,

$$(10)\quad u(x,t) = \frac{1}{2}\big(u_0(x-t) + u_0(x+t)\big) + \frac{1}{2}\int_{x-t}^{x+t} v_0(y)\,dy, \quad t < 0,\ x \in \mathbb{R}\,,$$

so we define

$$(11)\qquad u_{i,-1} \equiv u(x_i, t_{-1}) = \frac{1}{2}\big(u_0(x_{i-1}) + u_0(x_{i+1})\big) - \frac{1}{2}\int_{x_{i-1}}^{x_{i+1}} v_0(y)\,dy\,.$$

Strictly speaking, this is an illegal numerical method, since it contains an integral which we might have to do numerically. But if we cannot integrate v_0 in closed form, we can simply replace (10) by, say, $u_{i,-1} = \frac{1}{2}(u_0(x_{i-1}) + u_0(x_{i+1})) - hv_0(x_i)$, at the cost of a small error. (How small depends on the smoothness of v_0. If v_0 is $C^{(2)}$, this leads to an $O(h^3)$ error for $u_{i,-1}$, which gives an $O(h^2)$ error later on. If v_0 is only Hölder(1/2), as it would be if the initial velocity were Brownian, $u_{i,-1}$ would be off by $O(h^{3/2})$, leading to an $O(h^{1/2})$ error later on. This is the same order as the rest of the error.) We will carry out the rest of our analysis using (11). Notice that once the $u_{i,0}$ and $u_{i,-1}$ are known, the scheme determines the $u_{i,j}$ for all $(i,j) \in \mathcal{L}_h$ by iteration.

This gives us our numerical scheme: use the initial data and (11) to define $u_{i,j}$ for $j = -1$, and $j = 0$. Then use the iterative scheme (9) to determine it for $j = 1,\ 2, \ldots$. One can see from Figure 1 that this does indeed determine u_{ij} on $\mathcal{L}_h$.

REMARK 3.1. Note that the first step ($j = 0$) in the scheme is a special case. Indeed, we have extended u to be a solution of the homogeneous equation in the negative half plane, which is equivalent to assuming that both f and g vanish there. Thus when we determine $u_{i,1}$, we should replace $W(\Delta_{i,0})$ by $W(\Delta_{i,0} \cap \mathbb{R} \times \mathbb{R}_+)$, and we should replace h^2 by $h^2/2$ in the right-hand side of (9). This will make the first step exact in the case where f and g are constant.

Just to avoid any misunderstanding, let us re-state the scheme.

$$(12)\quad \begin{cases} u_{i,-1} \equiv u(x_i, t_{-1}) = \frac{1}{2}\big(u_0(x_{i-1}) + u_0(x_{i+1})\big) \\ \qquad\qquad - \frac{1}{2}\int_{x_{i-1}}^{x_{i+1}} v_0(y)\,dy, \quad i \text{ odd}\,; \\ u_{i,1} = u_0(x_{i-1}) + u_0(x_{i+1}) - u_{i,-1} \\ \qquad\qquad + \frac{h^2}{2} f\big(x_i, 0, \tfrac{1}{2}(u_0(x_{i+1}) + u_0(x_{i-1}))\big) \\ \qquad + \frac{1}{2} g\big(x_i, 0, \tfrac{1}{2}(u_0(x_{i+1}) + u_0(x_{i-1}))\big)\, W(\Delta_{i0} \cap \mathbb{R} \times \mathbb{R}_+)\,,\ i \text{ odd}\,; \\ u_{i,j+1} = u_{i+1,j} + u_{i-1,j} - u_{i,j-1} \\ \qquad\qquad + h^2 f\big(x_i, t_j, \tfrac{1}{2}(u_{i+1,j} + u_{i-1,j})\big) \\ \qquad + \frac{1}{2} g\big(x_i, t_j, \tfrac{1}{2}(u_{i+1,j} + u_{i-1,j})\big)\, W(\Delta_{ij})\,,\ (x_i, t_{j+1}) \in \mathcal{L}_h,\ j \ge 2\,. \end{cases}$$

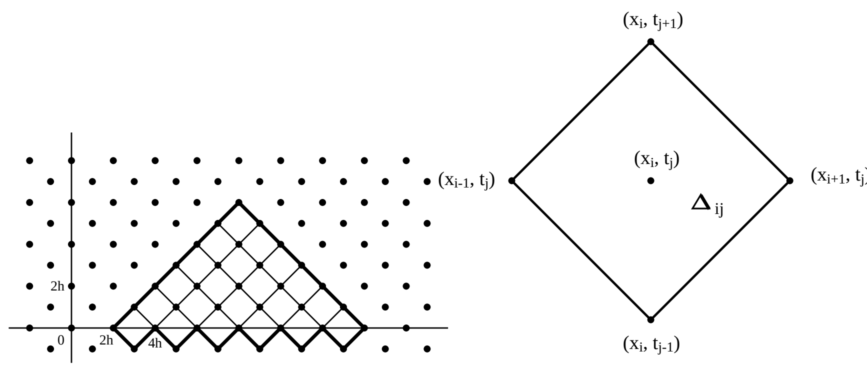

FIGURE 1. L_h and the Δ_{ij}

REMARK 3.2. As a scheme for the solution of PDEs—e.g., with $g \equiv 0$—this is a stable second order method, though its stability is borderline by the Courant-Friedrichs-Lewy criterion. We could evaluate the functions f and g at $u_{i,j-1}$ instead of at $\frac{1}{2}(u_{i+1,j} + u_{i-1,j})$. This would simplify some calculations, and result in a scheme which converges at the same order, at least for SPDEs. The only reason we did not use it is that it is less accurate for PDEs (it is only first order). We apologize to the readers for any extra trouble this causes.

PROPOSITION 3.3. *If the functions f and g are constant, say $f \equiv f_0$ and $g \equiv g_0$, this scheme gives the exact solution at all points of $\mathcal{L}_h$, i.e., $u_{i,j} = u(x_i, t_j)$ for all $(i,j) \in \mathcal{L}_h \cap \{j \geq 0\}$. In particular, the scheme is exact for the homogeneous wave equation.*

Proof. We prove this by induction. It is true for $j = -1$ and $j = 0$ by construction. Suppose it is true for all (i,j) for which $j \leq k$. Modulo some obvious changes, the same argument works whether k is even or odd, so suppose for concreteness that k is even. If i is odd, then $u_{i\pm1,k}$ and $u_{i,k-1}$ equal $u(x_{i\pm1}, t_k)$ and $u(x_i, t_{k-1})$ respectively. Thus we can define $u_{i,k+1}$ by (9). Compare it with $u(x_i, t_{k+1})$:

$$u_{i,k+1} = u_{i-1,k} + u_{i+1,k} - u_{i,k-1} + h^2 f_0 + \frac{1}{2} g_0 \, W(\Delta_{ik}) \, .$$

But $f \equiv f_0$ and $g \equiv g_0$ are constant, so $\frac{1}{2}\int_{\Delta_{ik}} f \, dy \, ds = h^2 f_0$ and $\frac{1}{2}\int_{\Delta_{ik}} g \, W(dy \, ds) = \frac{1}{2} g_0 \, W(\Delta_{ik})$. By the induction hypothesis, this is

$$= u(x_{i-1}, t_k) + u(x_{i+1}, t_k) - u(x_i, t_{k-1}) + \frac{1}{2}\int_{\Delta_{ik}} f \, dy \, ds + \frac{1}{2}\int_{\Delta_{ik}} g \, W(dy \, ds)$$

$$= u(x_i, t_{k+1}) \, .$$

by (7). Thus the result holds for every $(x_i, t_j) \in \mathcal{L}_h$ for which $j \leq k+1$, and therefore, it holds for all $\mathcal{L}_h$. $\square$

4. Preliminaries

Consider a fixed cone $C(0,N)$. Let us write $C_{ij} = C(x_i, t_j)$. The following is known from [3] and [8].

LEMMA 4.1. *Let u be the solution of (3). Then for any N and $p \geq 1$ there is a constant K_{Np} such that:*

(i) $E\{|u(x,y)|^{2p}\} \leq K_{Np}$ *for* $(x,y) \in C(0,N)$ *;*

(ii) $E\{|u(x,t) - u(y,s)|^{2p}\} \leq K_{Np}(|x-y|^p + |t-s|^p)$ *, for* (x,t)*,* $(y,s) \in C(0,N)$*.*

Note that the difference scheme is exact if f and g vanish identically, which means that

$$(13)\quad f \equiv g \equiv 0 \Longrightarrow u_{i,j} = \frac{1}{2}\big(u_0(x_i - t_j) + u_0(x_i + t_j)\big) + \frac{1}{2}\int_{x_i - t_j}^{x_i + t_j} v_0(y)\, dy\,.$$

Now suppose that f and g are no longer zero. Then:

PROPOSITION 4.2. *Let $\tilde{C}_{ij} = \cup\{\Delta_{mn} : \ \Delta_{mn} \cap \mathbb{R} \times \mathbb{R}_+ \subset C_{ij}\}$. For $(x_i, t_j) \in \mathcal{L}_h$,*

$$\begin{aligned}(14)\quad u_{i,j} = {}& \frac{1}{2}\left(u_0(x_i - t_j) + u_0(x_i + t_j)\right) + \frac{1}{2}\int_{x_i - t_j}^{x_i + t_j} v_0(y)\, dy \\ & + \frac{1}{2}\sum_{\{m,n:\Delta_{mn} \subset \tilde{C}_{ij}\}} \Big[f\big(x_m, t_n, \tfrac{1}{2}(u_{m+1,n} + u_{m-1,n})\big)|\Delta_{mn}| \\ & \qquad + g\left(x_m, t_n, \tfrac{1}{2}(u_{m+1,n} + u_{m-1,n})\right) W(\Delta_{mn})\Big].\end{aligned}$$

Proof. Like Proposition 3.3 this can be proved by induction, but we can also see it directly. Note that $\tilde{C}_{ij}$ is the union of C_{ij} with those Δ_{m0} for which $\Delta_{m0} \cap \mathbb{R} \times \mathbb{R}_+$ is contained in C_{ij}. (These are exactly the sets that come up in the first step of the recurrence scheme. The outlined area in Figure 1 is one of the $\tilde{C}_{ij}$.) In particular, $\tilde{C}_{ij}$ is a union of Δ_{mn}. Define $u(\Delta_{mn}) \equiv u_{m,n+1} - u_{m-1,n} - u_{m+1,n} + u_{m,n-1}$. Now notice that if we sum $u(\Delta_{mn})$ over all m, n for which $n \geq 0$ and $\Delta_{mn} \subset \tilde{C}_{ij}$, then, first, the sum is equal to the sum on the right-hand side of (14), and second, it telescopes dramatically: if we write it out in terms of the $u_{m,n}$, the terms from the common corners of any two Δ_{mn} which share an edge cancel; this means that all the terms $u_{m,n}$ cancel, except for $u_{i,j}$ itself and terms of the form $u_{m,n}$ where $n = -1$, (in which case there is no other adjacent Δ_{mn} to cancel it) and where $n = 0$ (in which case there is either no other adjacent Δ_{mn}, or else three Δ_{mn} meet, and only two can cancel.) After the telescoping, we are left with $u_{i,j} - \sum u_{k,0} + \sum u_{\ell,-1}$, where the sums are over k and ℓ for which the points $(k,0)$ and $(\ell,-1)$ are in $\mathcal{L}_h$ and also in $\tilde{C}_{ij}$.

To evaluate these sums, compare this with the solution $\hat{u}$ of the homogeneous wave equation with the same initial conditions. Let $\hat{u}_{m,n}$ be the result of the corresponding numerical scheme. Since the scheme is exact for the homogeneous equation, $\hat{u}_{m,n} = \hat{u}(x_m, t_n)$. The values of $u_{k,0}$ and $u_{\ell,-1}$ are calculated from the initial conditions and do not involve f and g, so $\hat{u}_{m,n} = u_{m,n}$ for $n = 0,\ -1$.

As above, let $\hat{u}(\Delta_{mn}) = \hat{u}_{m,n+1} - \hat{u}_{m-1,n} - \hat{u}_{m+1,n} + \hat{u}_{m,n-1}$ Then by (7), $\hat{u}(\Delta_{mn}) = 0$ for all m,n. Thus $0 = \hat{u}_{i,j} - \sum \hat{u}_{k,0} + \sum \hat{u}_{\ell,-1} = \hat{u}(x_i,t_j) - \sum u_{k,0} + \sum u_{\ell,-1}$; consequently, $\sum u_{k,0} - \sum u_{\ell,-1} = \hat{u}_{x_i,t_j} = 2^{-1}(u(x_i - t_j) + u(x_i + t_j)) + 2^{-1}\int_{x_i-t}^{x_i+t} v_0(y)\,dy$ by (13).

Finally, as remarked above, the sum of the $u(\Delta_{mn})$ is given by the sum on the right-hand side of (14). This proves the proposition. □

Let us interpret this result. Suppose we have carried out the difference scheme, so that the u_{ij} are known. Let

$$\hat{f}(x,t) = \sum_{mn} f(x_m, t_n, \tfrac{1}{2}(u_{m+1,n} + u_{m-1,n})) I_{\Delta_{mn}}(x,t)\,;$$

$$\hat{g}(x,t) = \sum_{mn} g(x_m, t_n, \tfrac{1}{2}(u_{m+1,n} + u_{m-1,n})) I_{\Delta_{mn}}(x,t)\,;$$

Then $f(x_m, t_n, (u_{m-1,n} + u_{m+1,n})/2)|\Delta_{mn}| = \int_{\Delta_{mn}} \hat{f}(y,s)\,dy\,ds$, and the first sum on the right-hand side of (14) is just $\int_{C_{ij}} \hat{f}(y,s)\,dy\,ds$. The second sum is $\int_{C_{ij}} \hat{g}(y,s)\,W(dy\,ds)$. (This is an integral with respect to a martingale measure, but because of the measurability properties of $\hat{g}$, the integral itself is delicate. See §5 and the proof of Lemma 5.3 for details.) Thus we have proved:

Corollary 4.3. *For $x_i, t_j \in \mathcal{L}_h$,*

$$2u_{i,j} = u_0(x_i - t_j) + u_0(x_i + t_j) + \int_{x_i-t_j}^{x_i+t_j} v_0(y)\,dy + \int_{C_{ij}} \hat{f}(y,s)\,dy\,ds + \int_{C_{ij}} \hat{g}(y,s)\,W(dy\,ds)\,. \tag{15}$$

In particular,

$$u_{ij} - u(x_i,t_j) = \frac{1}{2}\int_{C_{ij}} (\hat{f}(y,s) - f(y,s,u(y,s))\,dy\,ds + \frac{1}{2}\int_{C_{ij}} (\hat{g}(y,s) - g(y,s,u(y,s))\,W(dy\,ds)\,. \tag{16}$$

REMARK 4.4. Comparing (15) and (3), we see that the Green's function of the discrete equation is identical to the Green's function of the original equation.

5. Rate of convergence: the upper bound

THEOREM 5.1. *Let u_{ij} be the solution of the scheme (12) for $h > 0$. Let $\varepsilon > 0$. Then:*

(i) *u_{ij} converges as $h \to 0$ in all L^p to the true solution, and, for $p \geq 1$ there exists K_p such that for $(x_i, t_j) \in \mathcal{L}_h \cap C(0, N)$,*

$$\|u_{ij} - u(x_i, t_j)\|_p \leq K_p h^{p/2} . \tag{17}$$

(ii) *For any $\varepsilon > 0$ and $N > 0$, as $h \to 0$ through the sequence (2^{-n}),*

$$\lim_{\substack{h\to 0\\ h=2^{-n}}} \left[\sup_{(x_i,t_j)\in C(0,N)} \frac{|u_{ij} - u(x_i, t_j)|}{\sqrt{h|\log h|^{\varepsilon}}} \right] = 0 \tag{18}$$

REMARK 5.2. We can interpolate the u_{ij}, to, say, a piecewise-linear function $\hat{u}^h(x, t)$ which gives a continuous approximation of $u(x, t)$. It follows from (ii) above and the fact that the paths of $u(x, t)$ are Hölder(1/2) that $|\hat{u}^h(x, t) - u(x, t)|/\sqrt{h|\log h|^{\varepsilon}}$ converges uniformly on compact sets to zero as $h \to 0$ through the sequence (2^{-n}), so that there is uniform convergence on compact sets, not just on the lattice points.

If A is a Borel set, let $\mathcal{F}_A \equiv \sigma\{W(B) : B \subset A\}$ be the sigma field generated by white noise on A. Let $C^*(x, t) \equiv \{(y, s) : s \geq t,\ |y - x| \leq s - t\}$ be the forward light cone with apex (x, t), and denote its complement by $G^*(x, t) \equiv \mathbb{R} \times \mathbb{R}_+ - C^*(x, t)$. Define a martingale measure W^{ij} by $W^{ij}(A) = W(A \cap C^*(x_i, t_{j-1}))$. This is a restriction of white noise, so that $W_t^{ij}(A) \equiv W^{ij}(A \cap [0, t])$ is a martingale measure relative to the filtration $\mathcal{F}_t \equiv \mathcal{F}_{\mathbb{R}\times[0,t]}$. However, as white noise on $C^*(x_i, t_{j-1})$ is independent of white noise on $G^*(x_i, t_{j-1})$, it is also a martingale measure with respect to the enlarged filtration $\mathcal{G}_t^{ij} \equiv \mathcal{F}_t \vee \mathcal{F}_{G^*(x_i,t_{j-1})}$. Thus if $\rho(x, t)$ is a measurable, square-integrable process adapted to the filtration $(\mathcal{G}_t^{ij})$, then $\int_{\Delta_{ij}} \rho\, dW = \int_{\Delta_{ij}} \rho\, dW^{ij}$ is defined and satisfies

$$E\left\{ \int_{\Delta_{ij}} \rho\, dW \mid \mathcal{G}_0^{ij} \right\} = 0 ; \tag{19}$$

$$E\left\{ \left(\int_{\Delta_{ij}} \rho\, dW \right)^2 \mid \mathcal{G}_0^{ij} \right\} = E\left\{ \int_{\Delta_{ij}} \rho^2\, dx\, dt \mid \mathcal{G}_0^{ij} \right\} . \tag{20}$$

The proof of Theorem 5.1 begins with a series of lemmas.

LEMMA 5.3. *Let $N > 0$, $p > 1$. There exists a constant K_{Np} such that for $(x_m, t_n) \in \mathcal{L}_h \cap C(0, N)$*

(i) $E\left\{\left|\int_{C_{mn}} \hat g(x,t)\, W(dx\,dt)\right|^{2p}\right\} \le K_{Np} \int_{C_{mn}} E\{|\hat g(x,t)|^{2p}\}\, dx\, dt\,;$

(ii) $E\left\{\left|\int_{C_{mn}} (\hat g(x,t) - g(x,t,u(x,t))\, W(dx\,dt)\right|^{2p}\right\}$
$$\le K_{Np} \int_{C_{mn}} E\big\{|\hat g(x,t) - g(x,t,u(x,t)|^{2p}\big\}\, dx\, dt\,.$$

Proof. Let $\rho = \hat g$ in (i) and let $\rho = \hat g - g$ in (ii). Let G_0 be the x-axis and let and $G_{j+1} = \bigcup_{i:(x_i,t_j)\in\mathcal{L}_h} C_{ij}$. Let $\mathcal{G}_j = \mathcal{F}_{G_j}$. To shorten the notation, let $\Lambda(j;m,n) = \{i : (x_i,t_j) \in \mathcal{M}_h\,,\ \Delta_{ij} \subset C_{mn}\}$. Define

$$X_j \stackrel{\text{def}}{=} \int_{C_{mn}\cap G_j} \rho\, dW = \sum_{k=0}^{j-1} \int_{C_{mn}\cap(G_{k+1}-G_k)} \rho\, dW^{ij}$$
$$= \sum_{k=0}^{j-1} \sum_{i\in\Lambda(k;m,n)} \int_{\Delta_{ik}} \rho\, dW\,.$$

In particular,

$$X_{j+1} - X_j = \sum_{i\in\Lambda(j;m,n)} \int_{\Delta_{ij}} \rho\, dW\,.$$

Note: if $(x_i,t_j) \in \mathcal{M}_h$, that $u_{i+1,j}$ and $u_{i-1,j}$ are $\mathcal{F}_{G^*(x_i,t_{j-1})}$-measurable, and $g(x,t,u(x,t))$ is $\mathcal{F}_t$-measurable, so in either (i) or (ii), on Δ_{ij}, ρ is adapted to the filtration $(\mathcal{G}^{ij}_t)$, the integrals are defined, and we can apply (19) and (20). Since $\mathcal{G}_j \subset \mathcal{G}^{ij}_t$ for all t,

$$E\{X_{j+1} - X_j \mid \mathcal{G}_j\} = \sum_{i\in\Lambda(j;m,n)} E\left\{E\left\{\int_{\Delta_{ij}} \rho\, dW \mid \mathcal{G}^{ij}_0\right\} \mid \mathcal{G}_j\right\} = 0,$$

$$E\{(X_{j+1}-X_j)^2 \mid \mathcal{G}_j\} = \sum_{i,k\in\Lambda(j;m,n)} E\left\{E\left\{\int_{\Delta_{ik}} \rho\, dW \int_{\Delta_{ij}} \rho\, dW \mid \mathcal{G}^{ij}_0\right\} \mid \mathcal{G}_j\right\},$$

If $k \ne j$, the integral over Δ_{ik} is $\mathcal{G}^{ij}_0$-measurable and can be factored out of the expectation, and the corresponding term vanishes by (19). That leaves the terms for $k = j$, for which (20) applies:

$$= \sum_{i\in\Lambda(j;m,n)} E\left\{\int_{\Delta_{ij}} \rho\, dx\, dt \mid \mathcal{G}_j\right\} = E\left\{\int_{(G_{j+1}-G_j)\cap C_{mn}} \rho^2\, dx\, dt \mid \mathcal{G}_j\right\}.$$

It follows that $\{X_j,\ \mathcal{G}_j\,,\ 0 \le j \le n\}$ is a martingale with increasing process

$$[X]_k = \sum_{j=0}^{k-1} E\left\{\int_{(G_{j+1}-G_j)\cap C_{mn}} \rho^2\, dx\, dt \mid \mathcal{G}_j\right\}.$$

Thus, by Burkholder's inequality, there exists a constant K_p such that

$$E\left\{\left|\int_{C_{mn}} \rho\, dW\right|^{2p}\right\} \le K_p E\{[X]_n^p\}$$
$$\le K_p E\left\{\left(\int_{C_{mn}} \rho^2\, dx\, dt\right)^p\right\}.$$

By Hölder's inequality, this is

$$\le C\int_{C_{mn}} E\left\{\rho^{2p}\right\}\, dx\, dt\,.$$

This proves both (i) and (ii). □

LEMMA 5.4. *There exists a constant K_{Np} such that*

(21) $$E\{|u_{mn}|^{2p}\} \le K_{Np}, \quad (x_m, t_n) \in C(0,N)\,.$$

Proof. Let

$$U(x,t) = \frac{1}{2}(u_0(x-t) + u_0(x+t)) + \frac{1}{2}\int_{x-t}^{x+t} v_0(y)\, dy$$

be the solution of the homogeneous wave equation with the given initial conditions. Let $M = \sup\{|U(x,t)|^{2p},\ (x,t) \in C(0,N)\}$. Then

(22) $$E\{|u_{mn}|^{2p}\} \le 3^{2p}\left(M + E\left\{\left|\int_{C_{mn}} \hat{f}(x,t)\, dx\, dt\right|^{2p}\right\} + E\left\{\left|\int_{c_{mn}} \hat{g}(x,t)\, W(dx\, dt)\right|^{2p}\right\}\right).$$

Apply Hölder's inequality to the first integral and Lemma 5.3 to the second to see that this is

(23) $$\le 3^{2p}\left(M + |C_{mn}|^{2p-1} E\left\{\int |\hat{f}|^{2p}\, dx\, dt\right\} + E\left\{\int |\hat{g}|^{2p}\, dx\, dt\right\}\right).$$

For $(x,t) \in \Delta_{ij}$, the Lipschitz conditions imply that

(24) $$\begin{aligned}|\hat{f}(x,t)|^{2p} &= |f(x_i, t_j, (u_{i-1,j} + u_{i+1,j})/2)|^{2p}\\ &\le L_N^{2p}(1 + |(u_{i-1,j} + u_{i+1,j})/2|)^{2p})\\ &\le K(1 + (|(u_{i-1,j}|^{2p} + |u_{i+1,j}|^{2p})/2)\,.\end{aligned}$$

For $j = 0, 1, 2, \ldots$ let

$$F_j = \sup\{E\{|u_{ik}|^{2p}\} : (i,k) \in \mathcal{L}_h,\ k \le j\,,\ (x_i, y_k) \in C(0,N)\}\,.$$

Then $F_0 = M$ and from (24), for $j \ge 1$,

$$E\left\{\int_{\Delta_{ij}} |\hat{f}|^{2p}\, dx\, dt\right\} \le K(1 + F_j)|\Delta_{ij}|\,.$$

The same is true of $\hat{g}$, so that from (23)

$$E\{|u_{mn}|^{2p}\} \le K\left(M + 2K \sum_{i,j:(x_i,t_j)\in C_{mn}\cap\mathcal{M}_h} (1+F_j)|\Delta_{ij}|\right). \tag{25}$$

Now for a given j there are at most $2n$ values of i for which $(x_i,t_j)\in C_{mn}$ and $(x_i,t_j)\in\mathcal{M}_h$, while $|C_{mn}|\le N$ (for $C_{mn}\subset C(0,N)$), and $|\Delta_{ij}|=2h$, so for a larger value of K this is

$$\le K\left(1 + nh\sum_{j=0}^{n-1} F_j h\right).$$

Now $nh = t_n \le N$ so, absorbing all the constants into K, we have

$$E\{|u_{mn}|^{2p}\} \le K\left(1 + \sum_{j=0}^{n-1} F_j h\right)$$

Set $\tilde{F}(t) = F_{[t/h]}$ and note that since $\tilde{F}$ is increasing, $\tilde{F}(t) \le K(1 + \int_0^t \tilde{F}(s)\,ds)$. By Gronwall, $\tilde{F}(t)\le Ke^{Kt}$, proving the theorem.

The same reasoning can be applied to $\hat{f}-f$ and $\hat{g}-g$. □

This brings us to

Lemma 5.5. *Let $p\ge 1$. There exists $K_p>0$ such that for all $(x_m,t_n)\in C(0,N)$,*

$$E\{|u_{mn} - u(x_m,t_n)|^{2p}\} \le Kh^p\,. \tag{26}$$

Proof. We proceed as in Lemma 5.4.

$$u_{mn} - u(x_m,t_n) = \int_{C_{mn}} (\hat{f}-f)\,dx\,dt + \int_{C_{mn}} (\hat{g}-g)\,W(dx\,dt)\,.$$

On Δ_{ij}

$$\begin{aligned}|\hat{f}-f| &= |f(x_i,t_j,(u_{i-1,j}+u_{i+1,j})/2) - f(x,t,u(x,t))| \\ &\le L_N\big(|x_i-x| + |t_j-t| + |\tfrac{1}{2}(u_{i-1,j}+u_{i+1,j}) - u(x,t)|\big)\,.\end{aligned}$$

Now the last absolute value is

$$\begin{aligned}\le \frac{1}{2}\big(&|u_{i-1,j}-u(x_{i-1},t_j)| + |u_{i+1,j}-u(x_{i+1},t_j)| \\ &+ |u(x_{i-1},t_j)-u(x,t)| + |u(x_{i+1},t_j)-u(x,t)|\big).\end{aligned}$$

Thus by Hölder, we have that on Δ_{ij},

$$\begin{aligned}E\{|\hat f - f|^{2p}\} \le K\big(h^{2p} + E\{|u_{i-1,j} - u(x_{i-1},t_j)|^{2p}\} \\ + E\{|u_{i+1,j} - u(x_{i+1},t_j)|^{2p}\} + E\{|u(x_{i-1},t_j) - u(x,t)|^{2p}\} \\ + E\{|u(x_{i+1},t_j) - u(x,t)|^{2p}\}\big)\,.\end{aligned}$$

Since $|x_i - x| + |t_j - t| \le 4h$ on Δ_{ij}, Lemma 4.1 implies that the last two expectations are $O(h^p)$, while the first two are each bounded by

$$F_j \stackrel{\text{def}}{=} \sup\big\{E\{|u_{ik} - u(x_i,t_k)|^{2p}\} : k \le j,\ (x_i,t_k) \in C(0,N)\big\}\,.$$

Thus,

$$E\bigg\{\int_{\Delta_{ij}} |\hat f - f|^{2p}\,dx\,dt\bigg\} \le K(h^{2p} + h^p + F_j)|\Delta_{ij}|\,.$$

The same is true for $|\hat g - g|$. Thus, as $h^2 < h$ for $h < 1$, it follows as in (25) that

$$E\{|u_{mn} - u(x_m,t_n)|^{2p}\} \le K\Bigg(h^p + \sum_{i,j:(x_i,t_j)\in C_{mn}\cap\mathcal{M}_h} F_j\,|\Delta_{ij}|\Bigg)\,,$$

which leads to the conclusion that $F_0 = Kh^p$ and $F_n \le K(h^p + \sum_{j=0}^{n-1} F_j h)$, $n \ge 1$. As before, if $\tilde F(t) = F_{[t/h]}$,

$$\tilde F(t) \le K\bigg(h^p + \int_0^t \tilde F(s)\,ds\bigg)$$

so that $\tilde F(t) \le Kh^p e^{Kt}$, which proves the lemma. □

Lemma 5.6. *There exists $K_p > 0$ such that*

$$E\Bigg\{\bigg|\sup_{(x_i,t_j)\in C(0,N)} \int_{C_{ij}} (\hat g - g)\,dW\bigg|^p\Bigg\} \le K_p h^p\,.$$

Proof. This is a direct consequence of Cairoli's maximal inequality for two-parameter martingales. To see where the two-parameter martingale enters, define

$$X_{ij} = \int_{\Delta_{ij}} (\hat g - g)\,dW$$

and note that if $C_{ij} \subset C_{pq}$, then

$$E\Big\{X_{pq} \mid \mathcal{F}_{C_{ij}}\Big\} = X_{ij}\,. \tag{27}$$

Indeed, $C_{pq} - C_{ij}$ is a union of $\Delta_{k\ell}$, and if $\Delta_{k\ell}$ is disjoint from C_{ij}, then $C_{ij} \subset G^*(x_k, t_{\ell-1})$, hence $\mathcal{F}_{C_{ij}} \subset \mathcal{G}_0^{k\ell}$, and therefore by (19)

$$E\bigg\{\int_{\Delta_{k\ell}} (\hat g - g)\,dW \mid \mathcal{F}_{C_{ij}}\bigg\} = E\bigg\{E\bigg\{\int_{\Delta_{k\ell}} (\hat g - g)\,dW \mid \mathcal{G}_0^{k\ell}\bigg\} \mid \mathcal{F}_{C_{ij}}\bigg\} = 0\,.$$

To put this in the usual two-parameter setting, let $m = j+i$ and $n = j-i$, (so $j = (m+n)/2$ and $i = (m-n)/2$) and define $\mathcal{T}_{mn} = \mathcal{F}_{C_{ij}}$ and $M_{mn} = X_{ij}$. Note that $m \le m'$ and $n \le n' \iff C_{ij} \subset C_{i'j'}$, so that (27) states that if $m \le m'$ and $n \le n'$ that

$$E\{X_{m'n'} \mid \mathcal{T}_{mn}\} = X_{mn}\,.$$

Thus $\{X_{mn}\,,\ \mathcal{T}_{mn}\,,\ -N \le m, n \le N\}$ is a martingale—we extend it to negative values of m by setting $X_{mn} = 0$ and $\mathcal{T}_{mn} = \{\Omega, \emptyset\}$ if $m < 0$—and it is easily checked that the hypotheses $(F1)$—$(F4)$ of [2] hold, so that Cairoli's maximal inequality implies that

$$E\{\max_{m,n} |X_{mn}|^{2p}\} \le KE\{|X_{NN}|^{2p}\}\,.$$

Lemma 5.6 now follows from Lemma 5.5 and the Lipschitz condition on g. □

Proof of Theorem 5.1. Part (i) has been proved in Lemma 5.5. We must prove a.s. uniform convergence. Consider

$$
\begin{aligned}
(28)\quad & E\Big\{\sup_{(x_i,t_j)\in C(0,N)} |u_{ij} - u(x_i,t_j)|^{2p}\Big\}^{1/(2p)} \\
&\le E\left\{\sup_{(x_i,t_j)\in C(0,N)} \left|\int_{C_{ij}} |\hat f - f|\,dx\,dt\right|^{2p}\right\}^{1/(2p)} \\
&\qquad + E\left\{\left|\sup_{(x_i,t_j)\in C(0,N)} \int_{C_{ij}} (\hat g - g)\,dW\right|^{2p}\right\}^{1/(2p)}.
\end{aligned}
$$

The first integrand is positive, so by Hölder,

$$
\begin{aligned}
\sup\left(\int_{C_{ij}} |\hat f - f|\,dx\,dt\right)^{2p} &\le \left(\int_{C(0,N)} |\hat f - f|\,dx\,dt\right)^{2p} \\
&\le K\int_{C(0,N)} |\hat f - f|^{2p}\,dx\,dt,
\end{aligned}
$$

which is bounded in expectation by Kh^p. But the second term on the right-hand side of (28) is also bounded by Kh^p by Lemma 5.6, so

$$E\left\{\sup |u_{ij} - u(x_i,t_j)|^{2p}\right\} \le Kh^p\,.$$

Now let $h = 2^{-n}$. Then for $\epsilon > 0$, $\delta > 0$

$$
\begin{aligned}
P\left\{\frac{\sup|u_{ij} - u(x_i,t_j)|}{\sqrt{2^{-n}n^{\epsilon}}} > \delta\right\} &\le \frac{E\{\sup|u_{ij} - u(x_i,t_j)|^{2p}\}}{\delta^{2p}\,2^{-pn}\,n^{p\varepsilon}} \\
&\le \frac{K}{\delta^{2p}\,n^{p\varepsilon}}\,.
\end{aligned}
$$

Choose p large enough that $p\varepsilon > 1$. This is then summable, and the a.s. convergence follows from the Borel-Cantelli Lemma. □

6. Rate of convergence: the lower bound

The rate of convergence of the numerical scheme (9) is on the order of the square root of the step size. We will prove in this section that no scheme converges at a faster rate than root h for all choices of f and g.

It is important to emphasize that we are talking of schemes which only use certain increments of the driving white noise. It is possible to do better if we can use more information about the white noise, as pointed out in Kloeden and Platen [7]. However, we often use these equations to simulate the solution, in which case we only want to simulate the increments we actually need, and we do not have access to more information about the white noise.

Let us speak heuristically for a moment. There are three sources of error which concern us here.

(1) Think of a PDE as an infinite system of ODEs. A numerical scheme approximates it by a finite system. In our analysis, this shows up in the Green's functions. The numerical scheme has a Green's function which may only approximate the true Green's function of the PDE.

This is already a principal concern in the non-stochastic case, but there are two further sources of errors in stochastic equations. These tend to be dominant in this type of SPDE.

(2) To solve the SPDE one must approximate terms like $\int G(x,t)\,W(dx\,dt)$ by sums of white noise increments.

(3) Multiple stochastic integrals such as $\int W(dx\,ds)\,W(dy\,dt)$ may enter the picture, and, for reasons associated with the Lévy area, these are ill-approximated by products of white noise increments.

As it turns out, neither (1) nor (2) happens for the leapfrog lattice $\mathcal{L}_h$ of Section 1 , at least at the lattice points—they do in between—but, as in Davie and Gaines [4], (2) does happen for the more usual rectangular lattice. We will treat both cases. They require quite different proofs.

We use an idea of Davie and Gaines [4]: any numerical approximation constructed from a certain set of white-noise increments is measurable with respect to the sigma field $\mathcal{F}_\Delta$ that they generate. Among all $\mathcal{F}_\Delta$-measurable random-variables, the best L^2 approximation to the true solution is its conditional expectation given $\mathcal{F}_\Delta$. No numerical scheme can do better, so the conditional expectation gives a lower bound for the L^2 error. Thus to prove that $\sqrt{h}$ is the best possible rate, we need only exhibit an equation for which the conditional expectation of the solution converges no faster than $O(\sqrt{h})$.

This difference between the two schemes comes from the geometry of the lattices. In the rectangular lattice, the solution is based on white noise increments over rectangles with sides parallel to the axes. (This includes the semi-discretization of [8] as a limiting case.) As with [4], these increments do not approximate the Green's function well, and we can get the lower bound by considering an equation with constant coefficients. In this case the solution is Gaussian and the conditional expectation is readily calculated.

With the leapfrog lattice, the solution is based on the increments $W(\Delta_{ij})$, which do approximate the integral of the Green's function accurately, and the solution is exact for equations of constant coefficients. Thus we must consider an equation with non-constant coefficients. Its solution is non-Gaussian, and the proof hinges on properties of the Lévy area.

There is one more remark to make before we begin: we showed that there was uniform convergence of order (almost) root h. It is easy to show that no method can do better than this *uniformly* in (x,t). Indeed, the process varies by that much between lattice points, so the error halfway between lattice points will be $O(\sqrt{h})$. The convergence at lattice points is much more delicate.

THEOREM 6.1. *The rate of $O(\sqrt{h})$ is best possible on the lattice $\mathcal{L}_h$ in the following sense. There are Lipschitz functions f and g in (1) such that any numerical scheme which uses only the increments $W(\Delta_{ij})$ for a step-size h cannot converge in L^2 at a faster rate than $O(\sqrt{h})$, even at the lattice points.*

THEOREM 6.2. *The rate of $O(\sqrt{h})$ is best possible on a uniform rectangular lattice in the following sense. Let h and k be strictly positive, and let $\square_{ij} = (ih,(i+1)h]\times(jk,(j+1)k]$. Then there are Lipschitz functions f and g such that any numerical scheme for (1) which depends only on the increments $W(\square_{ij})$ cannot converge in L^2 at a rate faster than $O\big(\max(\sqrt{h},\sqrt{k})\big)$, even at the lattice points.*

6.1. Proof of Theorem 6.1. If we rotate coordinates by 45°, the stochastic wave equation falls into the purview of the two-parameter stochastic calculus as given in, for instance, [2]. We will use this fact heavily, and it will be cleaner if we make this change of variables at the outset. Therefore, let $\xi = (x+t)/\sqrt{2}$ and $\eta = (x-t)/\sqrt{t}$ as in (5).

The original problem is posed on the upper half plane in xt space; this becomes the half-plane $\mathfrak{H}_+ \stackrel{\text{def}}{=} \{(\xi,\eta) : \xi+\eta \geq 0\}$. The cone $C(x,t)$ transforms into $R_{\xi\eta} \stackrel{\text{def}}{=} \mathfrak{H}_+ \cap (-\infty,\xi]\times(-\infty,\eta]$. Let $\mathcal{F}_{\xi\eta} = \sigma\{W(R_{uv}) : u \leq \xi, v \leq \eta\}$. Then $\mathcal{F}_{\xi\eta}$ is the sigma field generated by white noise on $R_{\xi\eta}$.

The following lemma will be used repeatedly, usually without mention. It can be found in a more general form in [2].

LEMMA 6.3. *If $\varphi_i = \varphi_i(\xi,\eta,\omega)$, $i=1,2$, are measurable in all variables, adapted to $\mathcal{F}_{\xi\eta}$, and square integrable with respect to $d\xi\,d\eta\,dP$, then the stochastic integrals $\int \varphi_i\,dW$ exist, have mean zero, and satisfy*

$$E\left\{\int \varphi_1\,dW \int \varphi_2\,dW\right\} = \int E\{\varphi_1\varphi_2\}\,d\xi\,d\eta\,.$$

Proof of Theorem 6.1. Consider the problem in the xt-coordinates:

$$\begin{cases} u_{tt} = u_{xx} + 2u\dot{W} \\ u(x,0) = 1 \\ u_t(x,0) = 0\,. \end{cases} \tag{29}$$

whose mild form is $u(x,t) = 1 + \int_{C(x,t)} u(y,s)\, W(dy\, ds)$. In the $\xi\eta$-coordinates, it becomes

$$u(\xi,\eta) = 1 + \int_{R_{\xi\eta}} u(x,y)\, W(dx\, dy)\,. \tag{30}$$

This has a unique continuous solution $u(\xi,\eta)$ which is locally bounded in L^2. For the purposes of this proof, we can restrict ourselves to a bounded domain, which we will take to be R_{11} without loss of generality.

We will use the $\xi\eta$-coordinates for the remainder of this proof, and we will re-cycle the letters x, y, s, and t for dummy variables without reference to the previous coordinate system. In order to simplify the notation, we will omit dummy variables when there is no danger of confusion. The letters C and K below will denote strictly positive constants whose value may change from line to line.

From Lemma 4.1:

$$\sup_{R_{11}} E\{u(x,y)^2\} \le C, \tag{31}$$

$$E\{(u(x_1,y_1) - u(x_2,y_2))^2\} \le C\sqrt{(x_1-x_2)^2 + (y_2-y_2)^2}\,. \tag{32}$$

Fix $h > 0$ and let $x_i = ih$, $y_j = jh$, $i,j = 0, \pm 1, \pm 2, \dots$. Call the (x_i, y_j) *lattice points.* Define rectangles $\Delta_{ij} \stackrel{\text{def}}{=} (x_i, x_{i+1}] \times (y_j, y_{j+1}]$. Let $\mathcal{F}_\Delta = \sigma\{W(\Delta_{ij}),\ i,j = 0, \pm 1, \pm 2, \dots\}$, and set

$$\hat{u}(x,y) = E\{u(x,y) \mid \mathcal{F}_\Delta\}\,.$$

By our remarks above, the theorem will be proved if we can show that there is $C > 0$ such that $E\{(u(x,y) - \hat{u}(x,y))^2\} \ge Ch$ for lattice points $(x,y) \in R_{11}$.

Let us assume the conclusion is false and derive a contradiction. That is, assume that there exists a function $\rho(h)$ such that $\rho(h)/h \longrightarrow 0$ as $h \to 0$ and

$$E\{(u(x_i,y_j) - \hat{u}(x_i,y_j))^2\} \le \rho(h),\ (x_i,y_j) \in R_{11}\,. \tag{33}$$

We know that at least one scheme produces a $\sqrt{h}$ error, which implies that for some $K > 0$,

$$E\{(u(x,y) - \hat{u}(x,y))^2\} \le Ch, \quad \text{for all } (x,y) \in R_{11}\,. \tag{34}$$

For a lattice point (x_p, y_q) we have

$$u(x_p,y_q) = 1 + \int_{R_{x_p y_q}} u(\xi,\eta)\, W(d\xi\, d\eta)\,. \tag{35}$$

The plan of the proof is to decompose this integral and identify the important terms. As it happens, there are many terms and most of the proof is spent dealing with the small ones. The important terms, as we will see below, are I_{31} and I_{32}.

Let $\Gamma^h_{pq} = \{(m,n) : m+n \geq 0, m \leq p,\ n \leq q\}$ and $\Gamma_h = \cup \Gamma^h_{mn} = \{(m,n) : m+n \geq 0\}$. Then $(m,n) \in \Gamma^h_{pq} \iff (x_m, y_n) \in R_{x_p y_q}$, and $(m,n) \in \Gamma^h \iff (x_m, y_n) \in \mathfrak{H}_+$.

For $m+n \geq 0$, $\Delta_{mn} \subset R_{x_p y_q}$ iff $(m,n) \in \Gamma^h_{p-1,q-1}$. However, if $m+n = -1$ and $-q \leq m < q$, Δ_{mn} intersects both $R_{x_p y_q}$ and its complement, and $\Delta_{mn} \cap R_{x_p, y_q} = \Delta_{mn} \cap \mathfrak{H}_+$. Thus (35) is

$$\begin{aligned} &= 1 + \sum_{(m,n)\in\Gamma^h_{p-1,q-1}} \int_{\Delta_{mn}} u\,dW + \sum_{m=-q}^{p-1} \int_{R_+ \cap \Delta_{m,-m-1}} u\,dW \\ &\stackrel{\text{def}}{=} 1 + S_1 + S_2\,. \end{aligned}$$

Note that S_1 is a sum of $O(1/h^2)$ orthogonal terms, while S_2 is a sum of $O(1/h)$ terms of the same kind, so S_2 is negligible compared to S_1. Define the double difference operator $\Box_{ij}$ by

$$\Box_{ij} u = u(x_{i+1}, y_{j+1}) - u(x_i, y_{j+1}) - u(x_{i+1}, y_j) + u(x_i, y_j),$$

and note that by (30),

$$\Box_{mn} u = \int_{\Delta_{mn}} u\,dW\,. \tag{36}$$

Thus from (35)

$$u(x_p, y_q) = 1 + \sum_{(m,n)\in\Gamma^h_{p-1q-1}} \Box_{mn} u + S_2. \tag{37}$$

Our aim is to find $\Box_{mn} u - E\{\Box_{mn} u \mid \mathcal{F}_\Delta\} = \Box_{mn} u - \Box_{mn}\hat{u}$. Use (30) again on the integrand in (36) see that

$$\begin{aligned} \Box_{mn} u &= \int_{\Delta_{mn}} \left(1 + \int_{R_{xy}} u(\xi,\eta)\,W(d\xi\,d\eta)\right) W(dx\,dy) \\ &= \int_{\Delta_{mn}} \left(1 + \int_{R_{x_m y_n}} u\,dW\right) W(dx\,dy) \\ &\quad + \int_{\Delta_{mn}} \left(\int_{R_{xy}\cap\Delta_{mn}} u\,dW\right) W(dx\,dy) \\ &\quad + \int_{\Delta_{mn}} \left(\int_{R_{xy_n} - R_{x_m y_n}} u\,dW\right) W(dx\,dy) \\ &\quad + \int_{\Delta_{mn}} \left(\int_{R_{x_m y} - R_{x_m y_n}} u\,dW\right) W(dx\,dy) \\ &\stackrel{\text{def}}{=} I_1(m,n) + I_2(m,n) + I_3(m,n) + I_4(m,n)\,. \end{aligned}$$

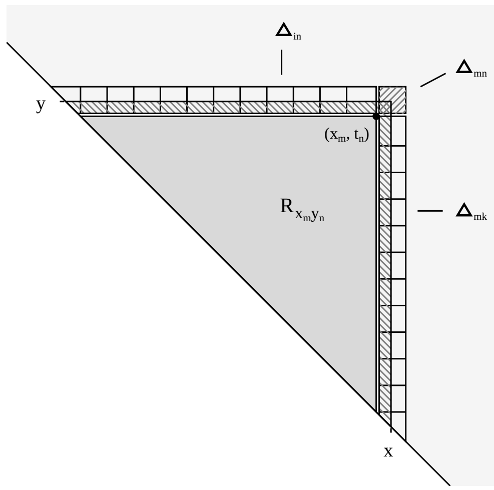

FIGURE 2. R_{x_m,y_n} and the strips $R_{x,y_n} - R_{x_m,y_n}$ and $R_{x_n,y} - R_{x_n y_m}$

See Figure 2 for a picture of the areas of integration. Note that for each k the sets $\{I_k(m,n),\ m,n = 0, \pm 1, \pm 2 \dots\}$ are families of stochastic integrals over the disjoint sets Δ_{mn}, and are therefore orthogonal and have mean zero.

Now $I_1 = u(x_m, y_n)\, W(\Delta_{mn})$, so let us write $I_1 = I_{11} + I_{12}$, where $I_{11}(m,n) = \hat{u}(x, m, y_n)\, W(\Delta_{mn})$ and $I_{12}(m,n) = (u(x_m, y_n) - \hat{u}(x_m, y_n)) W(\Delta_{mn})$. Then I_{11} is $\mathcal{F}_\Delta$-measurable. Both $u(x_m, y_n)$ and $\hat{u}(x_m, y_n)$ are $\mathcal{F}_{x_m y_n}$-measurable, so

$$
\begin{aligned}
(38)\quad E\{I_{12}(m,n)^2\} &= E\big\{E\{I_{12}^2 \mid \mathcal{F}_{x_m y_n}\}\big\} \\
&= E\big\{(u(x_m, y_n) - \hat{u}(x_m, y_n))^2 E\{W(\Delta_{mn})^2 \mid \mathcal{F}_{x_m y_n}\}\big\} \\
&\le \rho(h) |\Delta_{mn}|
\end{aligned}
$$

by hypothesis. Next,

$$
E\{I_2^2\} = \int_{\Delta_{mn}} \left(\int_{R_{xy} \cap \Delta_{mn}} E\{u(\xi, \eta)^2\}\, d\xi\, d\eta \right) dx\, dy\,.
$$

Since $E\{u(\xi,\eta)^2\} \le C$ we see that

$$
(39)\qquad E\{I_2(m,n)^2\} \le C \frac{h^2}{4} |\Delta_{mn}|\,.
$$

This brings us to I_3 and I_4. By symmetry, it suffices to consider I_3.

$R_{xy_n} - R_{x_m y_n}$ is a vertical strip. (See Figure 2.) Write it in terms of its intersections with the Δ_{mk} which lie below Δ_{mn}, and note that if $k < n$, that $(R_{xy_n} - R_{x_m y_n}) \cap \Delta_{mk} = R_{x,y_n} \cap \Delta_{mk}$. Thus

$$
\begin{aligned}
I_3(m,n) &= \sum_{\substack{k<n\\(m,k)\in\Gamma^h}} \int_{\Delta_{mn}} \left(\int_{\Delta_{mk}\cap R_{xy_n}} u(\xi,\eta)\, W(d\xi\, d\eta)\right) W(dx\, dy)\\
&= \sum_{\substack{k<n\\(m,k)\in\Gamma^h}} \int_{\Delta_{mn}} \hat u(x_m,y_n) W(\Delta_{mk}\cap R_{xy_n})\, W(dx\, dy)\\
&+ \sum_{\substack{k<n\\(m,k)\in\Gamma^h}} \int_{\Delta_{mn}} \left(\int_{\Delta_{mn}\cap R_{xy_n}} (u(\xi,\eta)-\hat u(x_m,y_n)) W(d\xi, d\eta)\right) W(dx\, dy)\\
&\overset{\text{def}}{=} I_{31}(m,n) + I_{32}(m,n)\,.
\end{aligned}
$$

Now both $\{I_{31}(m,n),\ (m,n)\in\Gamma^h\}$ and $\{I_{32}(m,n),\ (m,n)\in\Gamma^h\}$ are orthogonal families, and, moreover,

$$
\begin{aligned}
&E\{I_{32}(m,n)^2\}\\
&= \int_{\Delta_{mn}} E\left\{\left(\sum_{\substack{k<n\\(m,k)\in\Gamma^h}} \int_{\Delta_{mk}\cap R_{xy_n}} (u(\xi,\eta)-\hat u(x_m,y_n))\, W(d\xi\, d\eta)\right)^2\right\} dx\, dy\,.
\end{aligned}
$$

If $(\xi,\eta)\in\Delta_{mk}$, it is within a distance h of (x_m,y_k), so by (32) and (33)

$$
\begin{aligned}
\|u(\xi,\eta)-\hat u(x_m,y_k)\|_2 &\le \|u(\xi,\eta)-u(x_m,y_k)\|_2 + \|u(x_m,y_k)-\hat u(x_m,y_k)\|_2\\
&\le C\sqrt{h} + \sqrt{\rho(h)} \le K\sqrt{h}
\end{aligned}
$$

for some constant K. Thus $E\{(u(\xi,\eta)-\hat u(x_m,y_n))^2\}\le K^2 h$ and

$$
E\{I_{32}(m,n)^2\} \le h^2 \sum_{\substack{k<n\\(m,k)\in\Gamma^h}} K^2 h|\Delta_{mn}| \le (|m|+|n|)K^2h^3|\Delta_{mn}|\,. \tag{40}
$$

This brings us to the key term, I_{31}. Let $X_{ij}(t) = W((x_i, x_i+t]\times(y_j,y_{j+1}])$, $0\le t\le h$, and $\tilde X_{ij}(t) = W((x_i,x_{i+1}]\times(y_j,y_j+t])$, $0\le t\le h$. Note that the (X_{ij}) are independent Brownian motions with $d\langle X_{ij}\rangle_t = hdt$, and

$$
\hat u(x_m,y_k)W(\Delta_{mj}\cap R_{x,y_n}) = \hat u(x_m,y_k)X_{mk}(x-x_m).
$$

Thus we can let $t = x-x_m$ and identify the white noise integral above with a Brownian stochastic integral:

$$
I_{31}(m,n) = \sum_{\substack{k<n\\(m,k)\in\Gamma^h}} \hat u(x_m,y_k)\int_0^h X_{mk}(t)\, dX_{mn}(t)\,. \tag{41}
$$

This is the key observation. Doing the same for I_4, we see that $I_4 = I_{41}+I_{42}$ where

$$I_{41}(m,n) = \sum_{\substack{i<m \\ (i,m)\in\Gamma^h}} \hat{u}(x_i,y_n)\int_0^h \tilde{X}_{in}(t)\,d\tilde{X}_{mn}(t), \tag{42}$$

$$E\{I_{42}(m,n)^2\} \le (m+n)K^2h^3|\Delta_{mn}|\,. \tag{43}$$

To summarize, we have shown that

$$\begin{aligned} \Box_{mn}u = \hat{u}(x_m,y_n)W(\Delta_{mn}) &+ \sum_{\substack{k<n \\ (m,k)\in\Gamma^h}} \hat{u}(x_m,y_k)\int_0^h X_{mk}(t)\,dX_{mn}(t) \\ &+ \sum_{\substack{i<m \\ (i,m)\in\Gamma^h}} \hat{u}(x_i,y_n)\int_0^h \tilde{X}_{in}(t)\,d\tilde{X}_{mn}(t) + e(m,n)\,, \end{aligned} \tag{44}$$

where $e(m,n) = I_{12}(m,n) + I_2(m,n) + I_{32}(m,n) + I_{42}(m,n)$.

Let us consider the conditional expectation of $\Box_{mn}u$ given $\mathcal{F}_\Delta$. The first term on the right-hand side of (44) is $\mathcal{F}_\Delta$-measurable. To handle the stochastic integrals, note that for each i, j, k, and ℓ, $X_{ij}(h)X_{k\ell}(h) = \int_0^h X_{ij}\,dX_{k\ell} + \int_0^h X_{k\ell}\,dX_{ij}$. It follows by symmetry and the Markov property that

$$\begin{aligned} E\Big\{\int_0^h X_{ij}\,dX_{k\ell} \mid \mathcal{F}_\Delta\Big\} &= E\Big\{\int_0^h X_{ij}\,dX_{k\ell} \mid X_{ij}(h),\ X_{k\ell}(h)\Big\} \\ &= \frac{1}{2}X_{ij}(h)X_{k\ell}(h)\,. \end{aligned}$$

Moreover,

$$\int_0^h X_{ij}\,dX_{k\ell} - \frac{1}{2}X_{ij}(h)X_{k\ell}(h) = \frac{1}{2}\int_0^h X_{ij}\,dX_{k\ell} - \frac{1}{2}\int_0^h X_{k\ell}\,dX_{ij}\,.$$

The right-hand side is just half the Lévy area of X_{ij} and $X_{k\ell}$, where the *Lévy area* $\mathcal{A}_t(X,Y)$ of two semi-martingales (X_t) and (Y_t) is

$$\mathcal{A}_t(X,Y) \stackrel{\text{def}}{=} \int_0^t X\,dY - \int_0^t Y\,dX\,.$$

In terms of the Lévy area, we have

$$
\begin{aligned}
(45)\quad \Box_{mn} u - \Box_{mn}\hat{u} = & \sum_{\substack{k<n \\ (m,k)\in\Gamma^h}} \hat{u}(x_m, y_k)\mathcal{A}_h(X_{mk}, X_{mn}) \\
& + \sum_{\substack{i<m \\ (i,m)\in\Gamma^h}} \hat{u}(x_i, y_n)\mathcal{A}_h(\tilde{X}_{in}, \tilde{X}_{mn}) \\
& \qquad + e(m,n) - E\{e(m,n) \mid \mathcal{F}_\Delta\}\,.
\end{aligned}
$$

The following lemma will help with the Lévy areas.

Lemma 6.4.

(i) $E\{\mathcal{A}_h(X_{ij}, X_{mn})\} = 0$;
(ii) *if* $(i,j) \neq (m,n)$, $E\{\mathcal{A}_h(X_{ij}, X_{mn})^2\} = h^4$;
(iii) *Unless either* $(i,j) = (m,n)$ *and* $(k,\ell) = (p,q)$ *or* $(i,j) = (p,q)$ *and* $(k,\ell) = (m,n)$,

$$
\begin{aligned}
E\{\mathcal{A}_h(X_{ij}, X_{k\ell})A_h(X_{mn}, X_{pq})\} &= E\{\mathcal{A}_h(\tilde{X}_{ij}, \tilde{X}_{k\ell})A_h(\tilde{X}_{mn}, \tilde{X}_{pq})\} \\
&= E\{\mathcal{A}_h(X_{ij}, X_{k\ell})A_h(\tilde{X}_{mn}, \tilde{X}_{pq})\} = 0\,.
\end{aligned}
$$

Proof. The Lévy area is a stochastic integral, so (i) is immediate and (ii) follows from the independence of the different Brownian motions. To see (iii), write the Lévy area as a limit of Riemann sums: $\mathcal{A}_h(X_{ij}, X_{k\ell})\mathcal{A}_h(X_{mn}, X_{pq})$ is a limit of sums of the form $W(A_1)W(A_2)W(A_3)W(A_4)$, where $A_1 \subset \Delta_{ij}$, $A_2 \subset \Delta_{k\ell}$, $A_3 \subset \Delta_{mn}$ and $A_4 \subset \Delta_{pq}$. Under the hypotheses, one of the Δ's, say Δ_1, is distinct from the other three, hence $W(A_1)$ is independent of the others and

$$
E\Big\{\prod_{i=1}^{4} W(A_i)\Big\} = E\{W(A_1)\}E\Big\{\prod_{i=2}^{4} W(A_i)\Big\} = 0\,.
$$

Thus the Riemann sums have mean zero, and (i) follows in the limit. □

From (37) and (45)

$$\begin{aligned}
u(x_p, y_q) - \hat{u}(x_p, y_q) &= \sum_{(m,n)\in\Gamma^h_{p-1q-1}} (\Box_{mn} u - \Box_{mn}\hat{u}) + S_2 - E\{S_2 \mid \mathcal{F}_\Delta\} \\
&= \sum_{(m,n)\in\Gamma^h_{p-1,q-1}} \Bigg(\sum_{\substack{k<n \\ (m,k)\in\Gamma^h}} \hat{u}(x_m, y_k)\mathcal{A}_h(X_{mk}, X_{mn}) \\
&\qquad + \sum_{\substack{i<m \\ (i,m)\in\Gamma^h}} \hat{u}(x_i, y_n)\mathcal{A}_h(\tilde{X}_{in}, \tilde{X}_{mn}) \Bigg) \\
&\qquad + \sum_{(m,n)\in\Gamma^h_{p-1,q-1}} \big(e(m,n) - E\{e(m,n) \mid \mathcal{F}_\Delta\}\big) \\
&\qquad + S_2 - E\{S_2 \mid \mathcal{F}_\Delta\}.
\end{aligned}$$

By Lemma 6.4 the Lévy areas are orthogonal mean zero random variables, their variances add, and the expected square of the sums inside parentheses is

$$\sum_{(m,n)\in\Gamma^h_{p-1q-1}} \Bigg[\sum_{\substack{k<n \\ (m,k)\in\Gamma^h}} E\{\hat{u}(x_m, y_k)^2\}h + \sum_{\substack{i<m \\ (i,m)\in\Gamma^h}} E\{\hat{u}(x_i, y_n)^2\}h \Bigg] h^3$$

The bracketed terms are Riemann sums for integrals, so this is

$$\sim h \sum_{(m,n)\in\Gamma^h_{p-1q-1}} \left[\int_{-x_m}^{y_n} E\{\hat{u}(x_m, y)^2\}\, dy + \int_{-y_n}^{x_m} E\{\hat{u}(x, y_n)^2\}\, dx \right] h^2 \,,$$

which is itself a Riemann sum:

$$\sim h \int_{R_{x_p y_q}} \left[\int_{-t}^{s} E\{\hat{u}(t, y)^2\}\, dy + \int_{-s}^{t} E\{\hat{u}(x, s)^2\}\, dx \right] ds\, dt \,.$$

One can see from (30) that (u) is a martingale, so (u^2) is a sub-martingale, and $E\{(u-\hat{u})^2\} \le ch$, hence for small h, $\|\hat{u}(x,y)\|_2 \ge \|u(x,y)\|_2 - \sqrt{ch} \ge \|u(0,0)\|_2 - \sqrt{ch} = 1 - \sqrt{ch} \ge 1/2$. Thus the above is

$$\ge h \int_{R_{x_p y_q}} (t+s)\, ds\, dt \ge \frac{h}{6}(x_p + y_q)^3 \,,$$

and of course, $x_p + y_q > 0$ in $\mathfrak{H}_+$.

This gives us a lower bound on the main terms. It remains to check that the remaining terms are all small. Write

$$\sum_{m,n} e(m,n) = \sum_{m,n} I_{12}(m,n) + \sum_{m,n} I_2(m,n) + \sum_{m,n} I_{32}(m,n) + \sum_{m,n} I_{42}(m,n) \,.$$

The summands in each of the last four sums are orthogonal, mean-zero random variables, and therefore their variances add. By (38) and (33)

$$\begin{aligned} E\Big\{\Big(\sum_{mn} I_{12}(m,n)\Big)^2\Big\} &= \sum_{mn} E\{I_{12}(m,n)^2\} \\ &\le \rho(h)\sum_{mn}|\Delta_{mn}| = \rho(h)|R_{x_p y_q}| \\ &\le \frac{1}{2}\rho(h)\,, \end{aligned}$$

this last because $(x_p, y_q) \in R_{11}$ by hypothesis, so $R_{x_p y_q} \subset R_{11}$, which has area 1/2.

By (39), the variance of $I_2(m,n)$ is bounded by $Ch^2|\Delta_{mn}|/4$, so the same calculation tells us that

$$E\Big\{\Big(\sum_{mn} I_2(m,n)\Big)^2\Big\} \le Ch^2/8.$$

Now $|m|$ and $|n|$ are bounded by $1/h$ since $R_{mn} \subset R_{11}$, so by (40) the $I_{32}(m,n)$ have variances bounded by $(m+n)K^2h^3|\Delta_{mn}| \le 2K^2h^2|\Delta_{mn}|$. Therefore

$$E\Big\{\Big(\sum_{mn} I_{32}(m,n)\Big)^2\Big\} \le Kh^2/2.$$

The same is true of the $I_{42}(m,n)$. Thus we conclude that there is a constant $K > 0$ such that for all points $(x_p, y_q) \in R_{11}$,

$$\begin{aligned} E\Big\{\Big(\sum_{mn} e(m,n) - E\Big\{\sum_{mn} e(m,n) \mid \mathcal{F}_\Delta\Big\}\Big)^2\Big\} &\le E\Big\{\Big(\sum_{mn} e(m,n)\Big)^2\Big\} \\ &\le K(\rho(h)+h^2)\,. \end{aligned}$$

Thus for any lattice point (x_p, y_q) in R_{11}

$$\|u(x_p,y_q) - \hat{u}(x_p,y_q)\| \ge \sqrt{(x_p+y_q)h/6} - \sqrt{K(\rho(h)+h^2)}.$$

But $\rho(h) = o(h)$ so for small h the first square root is bounded below by, say, $\sqrt{(x_p+y_q)h/3}$. But this is strictly greater than $\rho(h)$ for small h. This contradicts the assumption (33) that $E\{(u(x_p,y_q) - \hat{u}(x_p,y_q))^2\} \le \rho(h)$ for small h. This proves Theorem 6.1. □

6.2. Proof of Theorem 6.2. Consider the linear problem

$$\begin{cases} u_{tt} = u_{xx} + 2\dot{W} \\ u(x,0) = 0 \\ u_t(x,0) = 0\,. \end{cases} \tag{46}$$

The solution is

$$u(x,t) = W(C(x,t))\,,$$

which is essentially a Brownian sheet.

Fix strictly positive h and k and let $D_{ij} = (ih,(i+1)h] \times (jk,(j+1)k]$. Let $\mathcal{F}_\Delta = \sigma\{W(D_{ij}) : i = 0, \pm 1,\ \pm 2, \dots,\ j = 0, 1,\ 2, \dots\}$. Fix x and t and let $D = C(x,t)$. It is straightforward to calculate $E\{W(D) \mid \mathcal{F}_\Delta\}$ since the variables are jointly Gaussian. If $|A|$ denotes the area of a set A,

$$E\{W(D) \mid \mathcal{F}_\Delta\} = \sum_{ij} \frac{|D \cap D_{ij}|}{|D_{ij}|} W(D_{ij})\,,$$

and the squared L^2 error is given by

$$\begin{aligned} e(x,t;h,k) &\stackrel{\text{def}}{=} E\Big\{\big(W(D) - E\{W(D) \mid \mathcal{F}_\Delta\}\big)^2\Big\} \\ &= \sum_{ij} |D \cap D_{ij}| \left(1 - \frac{|D \cap D_{ij}|}{|D_{ij}|}\right). \end{aligned} \tag{47}$$

The only non-zero contributions to the sum in (47) come from those D_{ij} which intersect both D and its complement, and which consequently intersect one or both of the two slanted segments of the boundary of D.

Let us remark that if we halve k, we increase the sigma field $\mathcal{F}_\Delta$, and therefore decrease $e(x,t;h,k)$. Thus, $e(x,t;h,k) \geq \liminf_{k\to 0} e(x,t;h,k)$. In fact the limit exists. Let us calculate it.

Fix (x_0,t_0) and consider the left-hand half of the boundary of $D = C(x_0,t_0)$ (the right-hand half is similar), which has the equation $x = x_0 - t_0 + t$. Let $[x]$ denote the greatest integer not exceeding x. If $t = (j+\frac{1}{2})k$, then $(x(t),t) \in D_{ij}$, where $i = [x/h]$. Let $\ell(x) \stackrel{\text{def}}{=} x - [x/h]\,h$ be the distance from x to ih. Then $|D_{ij}| = hk$ and $|D \cap D_{ij}| = (h - \ell(x))k$. Then the part of the sum (47) coming from the left-hand boundary is

$$\sum_{j=0}^{[t/k]} \ell(x((j+\frac{1}{2})k))\,[h - \ell(x((j+\frac{1}{2})k))]\,\frac{k}{h}\,.$$

This is a Riemann sum. As $k \to 0$, it tends to the integral

$$\int_0^{t_0} \ell(x(t))(h - \ell(x(t)))\,\frac{dt}{h} = \frac{1}{h}\int_{x_0-t_0}^{x_0} \ell(x)(h-\ell(x))\,dx\,. \tag{48}$$

If $(p-1)h < x_0 - t_0 \leq ph$ and $qh < x_0 \leq (q+1)h$, we can write the integral as $\int_{x_0-t_0}^{ph} + \sum_{j=p}^{q-1}\int_{jh}^{(j+1)h} + \int_{qh}^{x_0}$ and note that

$$\frac{1}{h}\int_{jh}^{(j+1)h} \ell(x)(h-\ell(x))\,dx = \frac{1}{h}\int_0^h x(h-x)\,dx = \frac{h^2}{6}\,.$$

The first and last integrals are no larger than this, and $q-p \le t_0/h \le q-p+2$, so that (48) is no smaller than $t_0h/6 - h^2/3$. Taking the other half of the boundary into account, we see that for any (x,t),

$$E\{(u(x,t)-\hat{u}(x,t))^2\} \ge \frac{th}{3} - \frac{2h^2}{3}\,.$$

We can now interchange the roles of h and k, by fixing k and letting h tend to zero. By symmetry, we get the lower bound $tk/3 - 2k^2/3$, which proves the theorem. □

Remark 6.5. The Courant-Friedrichs-Lewy condition for the numerical stability of this method is that $k \le h$. This assures that the domain of dependence of the numerical solution is contained in the physical domain of dependence, i.e., the numerical approximation at (x_i, t_j) depends on the values inside C_{ij}. Thus one would normally choose $k \le h$ above; however the case where $k \ge h$ is included in the theorem for symmetry.

7. Bounded regions: the vibrating string

So far we have only treated unbounded domains. However, it is possible to treat several boundary value problems by the same method. We will just consider one here, namely the wave equation on $D \equiv [0,1] \times \mathbb{R}_+$ with Dirichlet boundary conditions:

$$\begin{cases} \dfrac{\partial^2 u}{\partial t^2} = \dfrac{\partial^2 u}{\partial x^2} + f(x,t,u) + g(x,t,u)\,\dot{W}\,, \quad 0 \le x \le 1,\ t > 0\,, \\ u(x,0) = u_0(x)\,, \quad \dfrac{\partial u}{\partial t}(x,0) = v_0(x); \quad 0 \le x \le 1\,, \\ u(0,t) = u(1,t) = 0,\ t > 0\,. \end{cases}$$

E. Cabaña [1] introduced this to model a vibrating string perturbed by space-time white noise. It can be handled numerically by a minor modification of the scheme of Section 3, and it has the same rate of convergence. We will just outline this here.

Fix n, choose $h = 1/2n$, put $x_i = ih$, $t_j = jh$ and let u_{ij} be the discrete approximation of $u(x_i, t_j)$ as before. We observe that we need only determine u_{ij} for (x_i, t_j) in the interior of D, since the boundary conditions determine it on the boundary. Thus:

- Extend the initial values u_0 and v_0 to functions $\bar{u}_0$ and $\bar{v}_0$ on $\mathbb{R}$ which are odd functions of period two which equal u_0 and v_0 respectively on $[0,1]$.
- Calculate the $u_{i,-1}$ and determine u_{ij} for those $(x_i, t_j) \in \mathcal{L}_h$ which are in the interior of D by the scheme (9) with the following modifications:
  - u_0 and v_0 are replaced by $\bar{u}_0$ and $\bar{v}_0$ respectively;
  - $u_{ij} = 0$ if $(x_i, t_j) \in \mathcal{L}_h$ and x_i equals zero or one.

The main observation is that (7) holds for $\Delta = \Delta_{ij}$ if $(x_i, t_j) \in \mathcal{M}_h$ and $\Delta_{ij} \subset \bar{D}$. In particular it holds if $x_{i-1} = 0$ or $x_{i+1} = 1$, in which case $u_{i-1,j}$ or $u_{i+1,j}$ vanishes. In particular, (8) is exact if f and g are constant.

One can deduce from this that the scheme is exact if f and g are constant. This implies as before that the Green's function for the discrete scheme is the same as the Green's function for the PDE. (It is a difference of indicator functions of rectangles and triangles [1], and is bounded on compact sets.) This allows one to write the SPDE in its mild form and check that Lemma 4.1 holds for the solution of the SPDE. The analysis of the unbounded case then adapts straightforwardly to show that the scheme converges at the same rate, $O(\sqrt{h})$, as the unbounded case, and that that rate is the best possible.

References

[1] E. M. Cabaña, *On barrier problems for the vibrating string*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **22** (1972), 13–24. MR 0322974 (48 #1332)

[2] R. Cairoli and J. B. Walsh, *Stochastic integrals in the plane*, Acta Math. **134** (1975), 111–183. MR 0420845 (54 #8857)

[3] R. Carmona and D. Nualart, *Random nonlinear wave equations: smoothness of the solutions*, Probab. Theory Related Fields **79** (1988), 469–508. MR 966173 (90f:60112)

[4] A. M. Davie and J. G. Gaines, *Convergence of numerical schemes for the solution of parabolic stochastic partial differential equations*, Math. Comp. **70** (2001), 121–134 (electronic). MR 1803132 (2001h:65012)

[5] I. Gyöngy, *Lattice approximations for stochastic quasi-linear parabolic partial differential equations driven by space-time white noise. I*, Potential Anal. **9** (1998), 1–25. MR 1644183 (99j:60091)

[6] J. B. Walsh, *Introduction to stochastic partial differential equations*, École d'Éte de Probabilités de Saint-Flour, XIV, Lecture Notes in Mathematics, vol. 1180, Springer, Berlin, 1984, 1180, pp. 265–439. MR 0876085 (88a:60114)

[7] P. E. Kloeden and E. Platen, *Numerical solution of stochastic differential equations*, Applications of Mathematics (New York), vol. 23, Springer-Verlag, Berlin, 1992. MR 1214374 (94b:60069)

[8] L. Quer-Sardanyons and M. Sanz-Solé, *Space semi-discretisations for a stochastic wave equation*, Potential Analysis, to appear.

Department of Mathematics, University of British Columbia, Vancouver, BC V6T 1Z2, Canada

E-mail address: `walsh@math.ubc.ca`