UCLA
COMPUTATIONAL AND APPLIED MATHEMATICS

On Wavelet Fundamental Solutions to The heat Equation -- Heatlets

Jianhong Shen
Gilbert Strang

May 1999

CAM Report 99-18

Department of Mathematics
University of California, Los Angeles
Los Angeles, CA. 90095-1555

http://www.math.ucla.edu/applied/cam/index.html

ON WAVELET FUNDAMENTAL SOLUTIONS TO THE HEAT EQUATION — HEATLETS

JIANHONG SHEN AND GILBERT STRANG

ABSTRACT. We present an application of wavelet theory in partial differential equations. We study the wavelet fundamental solutions to the heat equation. The heat evolution of an initial wavelet state is called a *heatlet*. Like wavelets for the L^2 space, heatlets are "atomic" heat evolutions in the sense that any general heat evolution can be "assembled" from a heatlet according to some simple rules. We study the basic properties and algorithms of heatlets and related functions. Serving as a pointer to an interesting topic is our major goal.

Dedicated to Alan Newell.

1. INTRODUCTION

It is a pity that wavelets, in fact, have very little to do with physical waves or the wave equation ($u_{tt} = c^2\Delta u$). But the search for mutual interactions between wavelet theory and differential equations has never stopped. This paper is a part of it. The idea presented here awaits further exploration (and criticism) by both communities of differential equations and wavelet theory.

So far, the application of wavelet theory in (ordinary or partial) differential equations has been mostly focused on numerical computation. Wavelet-related functions are chosen as the basis functions or test functions for numerical solutions. It is now known that a large class of differential operators can be efficiently represented and stored through a wavelet basis (see Beylkin et. al. [1], for example). Despite its successful applications in numerical computation, wavelet theory still seems to keep some distance away from the theory of PDE's.

On the other hand, people have already felt certain theoretical links between PDE and wavelet theory. For example, one contact emerging between the two subjects is the concept of "scale." Both communities see it as they work on scale similarity. But the

Date: May 20, 1999.
To appear in *Journal of Differential Equations.*

gap is still open and no one has seen the bridge yet. Our work shall provide a partial answer.

A possible bridge is the invariant properties of wavelets and differential operators. Once an invariant property is shared, there exists a chance for interesting interactions to occur. It is this simple principle that has guided our following work.

A heatlet is a "fundamental" solution to the heat equation. The fundamentality is with respect to a wavelet decomposition. Consider the 1-D heat equation (homogeneous and with constant conduction coefficient)

$$u_t = \frac{\sigma^2}{2} u_{xx}, \qquad -\infty < x < \infty, \qquad t > 0,$$

with the initial state

$$u(x, 0) = f(x).$$

If $f(x)$ is the Dirac point-source at $x = 0$, i.e. $f(x) = \delta(x)$, then the heat evolution $u(x, t)$ is the classical *fundamental solution.* In this paper, instead of $\delta(x)$, we study the evolution of a wavelet-source $\psi(x)$ (a wavelet). The corresponding heat evolution $u(x, t)$ is called a *wavelet fundamental solution* to the heat equation, or simply a *heatlet.* The name "heatlet" is given on the consideration that a general heat evolution (from an arbitrary initial state $f(x)$) can be assembled from heatlets by simple rules. In this sense, a heatlet is an "atomic" heat evolution, analogous to the role that wavelets have played in the L^2 space.

The interaction between wavelets and the heat equation has been made possible by first applying a wavelet decomposition to the initial state, and then investigating the evolution of each wavelet component. The invariance under translation ($x \to x - b$) and dilation ($x \to ax$), for special a and b, of both wavelets and heat equations make such a dialogue successful.

Another channel that has connected the heat equation and wavelets is their common probabilistic background. The phenomenon of heat diffusion is closely related to Gaussian random process (through the kernel function). Similarly, the fundamental equation in wavelet theory – the refinement equation, also allows probabilistic interpretations (Derfel [3] and Shen [11]). As a result, we are able to view heatlets or related functions

in the framework of probability theory. This idea points to a computational algorithm for heatlets.

We have chosen the simple homogeneous heat equation as the model equation to illustrate our idea. Possible applications of the same idea to other equations are discussed briefly at the end of the article. These equations include the Schrödinger equation with a zero potential.

The presentation has been organized as follows. Section 2 is a brief survey on some fundamental concepts of wavelet theory. Section 3 introduces heatlets and heatlet decompositions. We have restricted ourselves to the 1-D heat equation. The probability interpretation and subdivision algorithm for heatlets constitute the main content of Section 4. In Section 5, we study the applications of multidimensional wavelets in multidimensional heat equations. The key idea is eventually summarized in the last section.

2. Multiresolution, Scaling Functions and Wavelets

To make the paper self-contained, we first give a brief survey on some major concepts and ideas of wavelet theory (See Daubechies [2], or Strang and Nguyen [14]).

The word "multiresolution" first appeared in image processing and was axiomatized by Mallat and Meyer in 1986. A multiresolution of $L^2(\mathbb{R})$ is a chain of *closed* subspaces indexed by all integers:

$$\cdots V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \cdots,$$

subject to the following three axioms

Axiom-1 (Completeness)

$$\overline{\lim_{n\to+\infty} V_n} = L^2(\mathbb{R}), \qquad \lim_{n\to-\infty} V_n = \{0\}.$$

Axiom-2 (Scale Similarity)

$$f(x) \in V_n \iff f(2x) \in V_{n+1}.$$

Axiom-3 (Translation Similarity) V_0 has a Riesz basis consisting of all integral translates of a single function $\phi(x)$: $\{\phi(x-n) : n \in \mathbb{Z}\}$. $\phi(x)$ is conventionally called the *scaling function.*

The axiom on scale similarity determines the whole chain from one subspace V_0, and the existence of the scaling function $\phi(x)$ compresses the information to a single function.

The existence of such a multiresolution is the *design problem.* To design a multiresolution, it suffices to construct the scaling function $\phi(x)$. The whole subspace chain can then be reconstructed from $\phi(x)$ according to the latter two axioms. From the fact that $V_0 \subset V_1$ and Axiom-2 and 3, it is easy to see that $\phi(x)$ must be a linear combination (maybe infinite) of $\{\phi(2x-n) : n \in \mathbb{Z}\}$. This leads to the so-called *refinement equation*:

$$\phi(x) = 2\sum_{n\in\mathbb{Z}} h_n \phi(2x-n)$$

for a certain set of coefficients $(\cdots, h_{-1}, h_0, h_1, \cdots)$. The design problem often starts with a good choice of this set of coefficients.

From the standpoint of digital signal processing, where computers can only manipulate finitely many numbers, the set should be finite. A typical equation in applications therefore has the form of

$$\phi(x) = 2\sum_{n=0}^{L} h_n \phi(2x-n), \tag{2.1}$$

with $h_0 h_L \neq 0$. Since $\int \phi = 1$ is always assumed in wavelet analysis, the coefficients must satisfy the following so-called *lowpass* condition:

$$h_0 + h_1 + \cdots + h_L = 1.$$

In signal processing, the coefficients set $(h_0, h_1, \cdots, h_L)$ is called a *digital filter* with *finite impulse response* (FIR). The connection between signal (and image) processing and wavelet theory is basically built upon this simple but important concept (see Strang and Nguyen [14]).

Let W_0 denote the orthogonal complement of V_0 in V_1. This is the space storing the "lost" detailed information of a signal (or an image) when one zooms out from a finer resolution V_1 to a coarser resolution V_0. So it is sometimes called the *detail space* at level 0. The good news is that W_0 also has an "atom," conventionally denoted by $\psi(x)$, such that the integral translates of $\psi(x)$ yield an orthonormal basis of W_0. It satisfies

the *wavelet equation* since $\psi(x) \in V_1$:

$$\psi(x) = 2\sum_{n\in\mathbb{Z}} g_n \phi(2x-n) \tag{2.2}$$

for a suitable set of new filter coefficients $(\ldots, g_{-1}, g_0, g_1, \ldots)$. From Axiom-1, 2 and 3, it is clear that

$$\{\psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k) : \ j,k \in \mathbb{Z}\}$$

is an orthonormal basis of $L^2(\mathbb{R})$.

The *wavelet decomposition* refers to the expansion of an arbitrary L^2 function $f(x)$ with respect to this basis of localized waves. Whereas the classical Fourier expansion and transform mainly analyze the *wave* contents of a signal, the wavelet decomposition unfolds a signal at different scales or resolutions. It is this attribute that has made wavelet analysis so powerful in data compression and image processing.

We hope this short survey is useful. For a deeper introduction to wavelet theory, we recommend the classical book by Daubechies [2], and the more recent one by Hernández and Weiss [7].

3. The Heatlet Decomposition

Let $\phi(x)$ and $\psi(x)$ be the scaling function and wavelet associated to a multiresolution. Let $\Phi^h(x,t)$ and $\Psi^h(x,t)$ be the heat evolutions of $\phi(x)$ and $\psi(x)$:

$$\Phi^h_t = \frac{\sigma^2}{2}\Phi^h_{xx}, \qquad \Phi^h(x,0) = \phi(x).$$

We call Ψ^h a *heatlet*, and Φ^h a *refinable heat.*

Proposition 3.1 (Similarity). *Suppose that $\phi(x)$ and $\psi(x)$ satisfy*

$$\phi(x) = 2\sum_{n\in\mathbb{Z}} h_n \phi(2x-n),$$
$$\psi(x) = 2\sum_{n\in\mathbb{Z}} g_n \phi(2x-n).$$

Here $(h_n), (g_n) \in l^2$, the Hilbert space of all square summable real sequences. Then the refinable heat Φ^h and heatlet Ψ^h also satisfy the self-similarity equations:

$$\Phi^h(x,t) = 2\sum_{n\in\mathbb{Z}} h_n \Phi^h(2x-n, 4t),$$
$$\Psi^h(x,t) = 2\sum_{n\in\mathbb{Z}} g_n \Phi^h(2x-n, 4t).$$

Proof. It is easy to check that $\Phi^h(2x-n, 4t)$ is the heat evolution corresponding to the initial state $\phi(2x-n)$. Hence, by linearity,

$$2\sum_{n\in\mathbb{Z}} h_n \Phi^h(2x-n, 4t)$$

is the heat evolution of

$$2\sum_{n\in\mathbb{Z}} h_n \phi(2x-n), \quad \text{or} \quad \phi(x).$$

By uniqueness, it must be Φ^h. This completes the proof of the first equation. The second one is done in the same fashion. □

Remark 1. In most applications, (h_n) and (g_n) are chosen to be FIR filters (see the previous section). Then both $\phi(x)$ and $\psi(x)$ are compactly supported, and the summations on the right-hand side of the preceding four equations are all finite.

One instant application in wavelet theory is in the design of high dimensional refinable functions. Define

$$\mathbf{x} = \begin{pmatrix} x \\ t \end{pmatrix}, \qquad M = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}.$$

For negative t, define $\Phi^h(x,t) = \Phi^h(x,-t)$. Then $\Phi^h(\mathbf{x})$ is a 2-D *refinable function* that satisfies the 2-D refinement equation

$$\Phi(\mathbf{x}) = 2\sum_{\mathbf{n}\in\mathbb{Z}^2} h_{\mathbf{n}} \Phi(M\mathbf{x} - \mathbf{n}).$$

Here $\mathbb{Z}^2$ is the lattice of all integral points (i.e. both coordinates are integers) in the plane, and $h_{\mathbf{n}}$ is zero unless $\mathbf{n} = (n,0)'$, in which case $h_{\mathbf{n}} = h_n$. This justifies the name "refinable heat". Design of high dimensional refinable functions and wavelets is still very challenging in wavelet theory.

Once having a heatlet $\Psi^h(x,t)$, we define

$$\Psi^h_{j,n}(x,t) = 2^{j/2}\Psi^h(2^j x - n, 4^j t).$$

This is obtained from Ψ^h by a dilation and a translation (with different rates for space and time). Then we have the following heatlet decomposition theorem.

Theorem 3.1 (Heatlet Decomposition). *Suppose $f(x) \in L^2(\mathbb{R})$. Then the corresponding heat evolution in $L^2(\mathbb{R})$ from $f(x)$ is given by*

$$u(x,t) = \sum_{j,n\in\mathbb{Z}} C_{j,n}(f)\Psi^h_{j,n}(x,t),$$

where $C_{j,n}(f)$ is the wavelet coefficient of $f(x)$ attached to $\psi_{j,n} = 2^{j/2}\psi(2^j x - n)$. Moreover, the infinite series converges in $L^2(\mathbb{R})$ uniformly with respect to t.

Proof. Let $P^\sigma_t(x)$ be the heat kernel

$$P^\sigma_t(x) = \frac{1}{\sigma\sqrt{2\pi t}} e^{-\frac{x^2}{2t\sigma^2}}.$$

Then the heat evolution of any initial state $f(x)$ is given by

$$u(x,t) = P^\sigma_t * f.$$

Since $f \in L^2(\mathbb{R})$, and the one-parameter family $(P^\sigma_t)_{t\geq 0}$ is an *approximate identity* (see Helson [6]), by Féjer's Theorem, at each time t, we have $u(\cdot,t) \in L^2(\mathbb{R})$ and $u(\cdot,t) \to f(x)$ in $L^2(\mathbb{R})$ as $t \to 0^+$.

Since $(\psi_{j,n})$ is an orthonormal basis of $L^2(\mathbb{R})$, the wavelet expansion of $f(x)$ converges to $f(x)$ in $L^2(\mathbb{R})$. From the heat equation, it is easy to check that $\Psi^h_{j,n}$ is the heat evolution of $\psi_{j,n}$. Finally, since for any time t, the heat kernel P^σ_t is a bounded operator in $L^2(\mathbb{R})$ (through convolution) with the operator norm less than 1, we conclude that the infinite series of the heatlet decomposition converges in $L^2(\mathbb{R})$ uniformly (w.r.t. the temporal parameter t) to the heat evolution of $f(x)$. This completes the proof. □

Likewise, corresponding to the multiresolution decomposition of the initial state

$$f(x) = \sum_{m\in\mathbb{Z}} C'_{0,m}(f)\phi_{0,m}(x,t) + \sum_{j\geq 0,n\in\mathbb{Z}} C_{j,n}(f)\psi_{j,n}(x,t),$$

we have the decomposition for the heat evolution

$$u(x,t) = \sum_{m\in\mathbb{Z}} C'_{0,m}(f)\Phi^h_{0,m}(x,t) + \sum_{j\geq 0,n\in\mathbb{Z}} C_{j,n}(f)\Psi^h_{j,n}(x,t).$$

Here $C'_{j,n}(f)$ is the coefficient of f attached to $\phi_{j,k}(x)$.

For any fixed positive integer J, define f_J to be the projection of $f(x)$ onto the J-th resolution space $V_J(\phi)$. Denote its heat evolution by $u_J(x,t)$. Since

$$\begin{aligned} f_J(x) &= \sum_{m\in\mathbb{Z}} C'_{J,m}(f)\phi_{J,m}(x) \\ &= \sum_{m\in\mathbb{Z}} C'_{0,m}(f)\phi_{0,m}(x) + \sum_{0\le j\le J-1, n\in\mathbb{Z}} C_{j,n}(f)\psi_{j,n}(x), \end{aligned}$$

we have

$$\begin{aligned} u_J(x,t) &= \sum_{m\in\mathbb{Z}} C'_{J,m}(f)\Phi^h_{J,m}(x,t) \\ &= \sum_{m\in\mathbb{Z}} C'_{0,m}(f)\Phi^h_{0,m}(x,t) + \sum_{0\le j\le J-1, n\in\mathbb{Z}} C_{j,n}(f)\Psi^h_{j,n}(x,t) \end{aligned}$$

If, as in most applications, both $\phi(x)$ and the initial state $f(x)$ are compactly supported, then only finitely many coefficients $C'_{J,m}(f), m \in \mathbb{Z}$ are non-zero. Therefore, the projection $u_J(x,t)$ is in fact a finite sum.

Suppose the wavelet $\psi(x)$ has p vanishing moments:

$$\int_{\mathbb{R}} x^k \psi(x)dx = 0, \quad k = 0, 1, \ldots, p-1.$$

With some extra conditions on the decay rate and regularity of $\psi(x)$ we have the following well-known estimation (Strang and Nguyen [14]).

Theorem 3.2. *Suppose the wavelet $\psi(x)$ has p vanishing moments and $f(x) \in C^p(\mathbb{R})$ with a bounded p-th derivative. Then*

$$\|f(x) - f_J(x)\|_{L^\infty(\mathbb{R})} = O(h_J^p),$$

where, $h_J = 2^{-J}$ is the resolution size at level J.

This estimation extends to $u_J(x,t)$ via the heat kernel.

Corollary 3.1 (Uniform Convergence). *Suppose $\psi(x)$ has p vanishing moments and the initial state $f(x) \in C^p(\mathbb{R})$ with $f^{(p)}(x)$ bounded. Then as J tends to $+\infty$,*

$$\|u(x,t) - u_J(x,t)\|_{L^\infty(\mathbb{R}\times\mathbb{R}^+)} = O(h_J^p).$$

Here, $\mathbb{R}\times\mathbb{R}^+$ denotes the upper half plane.

Remark 2. The major advantage of a heatlet and refinable heat is that they are universal, i.e. independent of the initial state. Once they have been worked out (as shown in the next section), we store them somewhere ready to profit from them. Given any initial heat state $f(x)$, we first apply DWT (*discrete wavelet transform*) to get the wavelet coefficients of $f(x)$. This is the only essential computation involved and is much easier in applications due to the many existing wavelet computational techniques such as the FWT (*fast wavelet transform*) and direct sampling. The heat evolution $u(x,t)$ can then be assembled simply based on the wavelet coefficients and the heatlet or refinable heat stored in the memory. An example is in order now.

EXAMPLE.

We compute the simplest heatlet – the Haar heatlet derived from the Haar multiresolution. The Haar wavelet $\psi(x)$ is defined by

$$\psi(x) = \begin{cases} 1, & 0 \le x < 1/2, \\ -1, & 1/2 \le x < 1, \\ 0, & \text{elsewhere.} \end{cases}$$

It was constructed by Haar [5] in 1910 searching for a "good" orthonormal system on $[0,1]$ with respect to which the Fourier expansion of any continuous function converges uniformly. It has been re-explained in terms of wavelets after the wavelet theory emerged. Recall the heat kernel

$$P_t^\sigma(x) = \frac{1}{\sigma\sqrt{2\pi t}} \exp(-\frac{x^2}{2t\sigma^2}). \tag{3.1}$$

Hence the associated Haar heatlet Ψ^h and refinable heat Φ^h are

$$\Psi^h(x,t) = \Phi(\frac{x}{\sigma\sqrt{t}}) + \Phi(\frac{x-1}{\sigma\sqrt{t}}) - 2\Phi(\frac{x-1/2}{\sigma\sqrt{t}}),$$
$$\Phi^h(x,t) = \Phi(\frac{x}{\sigma\sqrt{t}}) - \Phi(\frac{x-1}{\sigma\sqrt{t}}).$$

Here $\Phi(x)$ is the standard error function

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{s^2}{2}} ds.$$

It can be verified directly that

$$\Psi^h(x,t) = \Phi^h(2x,4t) - \Phi^h(2x-1,4t),$$

as predicted in Proposition 3.1.

The Haar heatlet belongs to a general class of "B-spline heatlets," the heat evolutions of the B-spline wavelets. Recall that the n-th order B-spline (scaling function) is defined recursively by

$$B_n = B_{n-1} * B_0, \qquad B_0 = 1_{[0,1)}(x).$$

That is, $B_n(x)$ is the $n+1$-th convolutional power of the indicator of $[0,1)$. As the power increases, so does the smoothness of the splines, which pleases numerical computations demanding smooth basis functions.

Denote the associated refinable heat by $B_n^h(x,t)$. Then the following proposition shows that the convolutional structure is preserved.

Proposition 3.2 (B-spline refinable heat). *For any time t, $B_n^h(x,t)$ is the $n+1$-th convolutional power (with respect to the spatial variable x) of $B_0^h(x,t/(n+1))$, the Haar refinable heat at time $t/(n+1)$.*

Proof. By direct computation,

$$\begin{aligned}
&\frac{\partial}{\partial t}\big[B_0^h(x,\frac{t}{n+1}) * \cdots * B_0^h(x,\frac{t}{n+1})\big] \qquad * \text{ by } n \text{ times}\\
=&(n+1)\big[\frac{\partial}{\partial t}B_0^h(x,\frac{t}{n+1})\big] * B_0^h(x,\frac{t}{n+1}) * \cdots * B_0^h(x,\frac{t}{n+1})\\
=&\big[\frac{\partial}{\partial t}B_0^h\big](x,\frac{t}{n+1}) * B_0^h(x,\frac{t}{n+1}) * \cdots * B_0^h(x,\frac{t}{n+1}).
\end{aligned}$$

Similary,

$$\begin{aligned}
&\frac{\partial^2}{\partial x^2}\big[B_0^h(x,\frac{t}{n+1}) * \cdots * B_0^h(x,\frac{t}{n+1})\big]\\
=&\big[\frac{\partial^2}{\partial x^2}B_0^h\big](x,\frac{t}{n+1}) * B_0^h(x,\frac{t}{n+1}) * \cdots * B_0^h(x,\frac{t}{n+1}).
\end{aligned}$$

Therefore the $n+1$-th convolutional power of $B_0^h(x,t/(n+1))$ does satisfy the heat equation. By letting $t \to 0$, we find that it corresponds to the initial spline state $B_n(x)$. The proof is thus complete. □

Remark 3. Due to the infinite-diffusion-velocity property of the heat equation (a defect caused by the linearity of the model equation, see Kolmogoroff et al. [8]), heatlets and refinable heats are not compactly supported as a 2-D function, even as a 1-D time-parameterized spatial function. However, as explicitly recognizable from the Haar heatlet, the "essential" support at time t is approximately $2\sigma\sqrt{t}$.

4. Probability Method and the Subdivision Algorithm

Except Haar and B-spline wavelets, general scaling functions and wavelets have no closed form. This makes it necessary to study the algorithms for computing quantities that involve the scaling functions or wavelets. In this section, we illustrate how the probability interpretation can lead to the *continuous subdivision algorithm* for computing heatlets. Continuous subdivision generalizes the classical discrete one in wavelet method and computer aided design (Derfel, Dyn and Levin [4]). An early application has appeared in solving refinement differential equations (Shen [11]). Here we are giving another nontrivial applicaton.

By Proposition 3.1, to compute a heatlet, it suffices to work out the associated refinable heat function. Our main task is to avoid using the scaling function $\phi(x)$ directly, and proceed with only the help of the filter coefficients.

Suppose a is a positive scalar. Let $N^a, N_1^a, N_2^a, \cdots$ be a sequence of independent random variables with an identical $N(0,a)$ distribution (i.e. normal distribution with mean 0 and variance a). Define

$$Y = \frac{N_1^a}{2} + \frac{N_2^a}{4} + \frac{N_3^a}{8} + \cdots. \tag{4.1}$$

Then Y is again a normal random variable of type $N(0, a/3)$.

Suppose $\phi(x)$ is the scaling function defined from

$$\phi(x) = 2\sum_{n=0}^{L} h_n \phi(2x-n).$$

In addition, let us assume *temporarily* or *formally* that

"all the filter coefficients h_n are non-negative."

The B-spline case is an example. From the lowpass condition

$$h_0 + h_1 + \cdots + h_L = 1,$$

this set of coefficients define a unique discrete probability density function (p.d.f.). Suppose X^h is such a random variable so that

$$\text{Prob}(X^h = n) = h_n, \qquad n = 0, 1, \cdots, L.$$

Then $\phi(x)$ is the p.d.f. of the random variable

$$X^\phi = \frac{X_1^h}{2} + \frac{X_2^h}{4} + \frac{X_3^h}{8} \cdots,$$

where $X_1^h, X_2^h, \cdots$ are i.i.d. random variables of the same type of X^h (see Shen [11]).

Recall that the refinable heat is the convolution $P_t^\sigma * \phi$, where P_t^σ is the heat kernel. P_t^σ is also the p.d.f. corresponding to $N(0, t\sigma^2)$. Therefore, $\Phi^h(\cdot, t)$ as a spatial function is the p.d.f. of the random variable

$$Z = X^\phi + N^{t\sigma^2}.$$

Theorem 4.1. *For any fixed time $t > 0$, let $\rho^t(x)$ denote the p.d.f. of $X^h + N^{3t\sigma^2}$. Then the refinable heat $\Phi^h(x,t)$ associated with the scaling function $\phi(x)$ solves the following* refinement equation *for an unknown function $F(x)$*

$$F(x) = 2\int_{\mathbb{R}} \rho^t(s) F(2x - s) ds. \tag{4.2}$$

Proof. Suppose $Z_0, Z_1, Z_2, \cdots$ are a sequence of i.i.d. random variables of type $X^h + N^{3t\sigma^2}$. Then

$$Z^* = \frac{Z_1}{2} + \frac{Z_2}{4} + \frac{Z_4}{8} + \cdots$$

is a random variable of type

$$Z = X^\phi + N^{t\sigma^2}.$$

From the definition, it is clear that Z^* has the same distribution as

$$\frac{Z_0 + Z^*}{2}.$$

Note that the p.d.f. of Z_0 and Z^* are $\rho^t(x)$ and $\Phi^h(x,t)$. Therefore, $\Phi^h(x,t)$ solves the given continuous refinement equation for each parameter t. (Recall that the addition of

two independent random variables leads to a convolution of their probability measures.)

□

Remark 4. Continuous refinement equations have been recently studied in Derfel, Dyn and Levin [4], and Shen [11]. The good news about (discrete or continuous) refinement equation is that they can always be solved by the *subdivision scheme.* The algorithm originates in computational geometry and plays an important role in wavelet computations. The continuous version of the algorithm has been applied to refinement differential equations in Shen [11]. Shen [12] also reveals its combinatorial meaning in the context of umbral calculus. Based on the preceding theorem, we can now apply the continuous subdivision algorithm to compute the heatlets and related functions.

Let $P_t^\sigma(x)$ be the heat kernel as defined in (3.1). Then the density function $\rho^t(x)$ in the preceding theorem is

$$\rho^t(x) = \sum_{n=0}^{L} h_n P_{3t}^\sigma(x-n),$$

which is a mixture of Gaussians (or "multi-humps"). If t is small, $\rho^t(x)$ must have totally $L+1$ humps if all the filter coefficients are non-zero (a typical case in practice).

We have assumed for convenience that all the filter coefficients are non-negative, which leads to a probability interpretation of the continuous refinement equation. It is not difficult to see that even when (h_n) changes signs, the continuous refinement equation (4.2) still holds (by Fourier method, for example).

Therefore, the refinable heat $\Phi^h(x,t)$ (and hence the heatlet $\Psi^h(x,t)$) can be obtained by the following continuous subdivision process, which avoids the exact expression of the scaling function and wavelet.

CONTINUOUS SUBDIVISION ALGORITHM FOR $\Phi^h(x,t)$

a) Initialize $F_0(x,t) = \delta(x)$;

b) For $n = 1, 2, \cdots$, define

$$F_n(x,t) = \int_{\mathbb{R}} 2\rho^t(x-2s) F_{n-1}(s,t) ds.$$

Since for any $t > 0$, the continuous filter $\rho^t(x)$ is a C^∞ and rapidly decaying function, the results of Derfel, Dyn and Levin [4] and Shen [11] on continuous subdivision guarantee the following.

Corollary 4.1. *For any $t > 0$, and non-negative integer m,*

$$F_n(2^n \cdot\ ,\ t) \quad \textit{converges uniformly to} \quad \Phi^h(\cdot\ ,\ t) \quad \textit{in } C^m(\mathbb{R}),$$

as $n \to \infty$.

As $t \to 0^+$, $\rho^t(x)$ converges in the distributional sense to the impulse train

$$\sum_{n=0}^{L} h_n \delta(x - n).$$

Then the continuous subdivision scheme degenerates to the ordinary discrete one, which is essentially an interpolation process. The limit produces the scaling function $\phi(x)$, which is consistent with the fact that $\phi(x)$ is the initial state of the refinable heat $\Phi^h(x,t)$.

Remark 5. In deriving the continuous refinement equation for the refinable heats, we have applied the semi-group property $P_t^\sigma * P_s^\sigma = P_{s+t}^\sigma$ of the heat kernel. This makes it possible to "decompose" any normal random variable into the infinite sum (4.1). It also explains our special interest in the heat equation.

5. The High Dimensional Case

Let

$$\Delta = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \ldots + \frac{\partial^2}{\partial x_n^2}$$

denote the n-dimensional Laplacian operator. The associated heat equation becomes

$$u_{tt}(\mathbf{x}, t) = \frac{\sigma^2}{2} \Delta u(\mathbf{x}, 0), \qquad u(\mathbf{x}, t) = f(\mathbf{x}). \tag{5.1}$$

More generally, let $A = (a_{i,j})$ be an n by n *positive definite* matrix. Define the A-weighted Laplacian Δ_A by

$$\Delta_A = \sum_{1 \le i,j \le n} a_{i,j} \frac{\partial^2}{\partial x_i \partial x_j}.$$

Then the associated A-weighted heat equation is

$$u_{tt}(\mathbf{x},t) = \frac{1}{2}\Delta_A u(\mathbf{x},t), \qquad u(\mathbf{x},0) = f(\mathbf{x}). \tag{5.2}$$

Definition 5.1 (Dilation Factor). An n by n matrix D is called a *dilation factor* if

a. all entries of D are integers and the determinant $|D| > 0$; and

b. all eigenvalues of D are outside the unit circle.

If in addition, $|D| = 2$, then D is said to be *tight*.

In wavelet theory, dilation factors are used to design high dimensional multiresolution. The refinement equation associated to a given dilation factor D is given by

$$\phi(x) = |D| \sum_{\mathbf{n}\in\mathbb{Z}^n} h_{\mathbf{n}}\phi(D\mathbf{x} - \mathbf{n}), \tag{5.3}$$

The difficulty is to design the filter coefficients $h_{\mathbf{n}}$. Under some special choice of the dilation matrix and filter coefficients, the scaling function $\phi(\mathbf{x})$ leads to a multiresolution in $L^2(\mathbb{R}^n)$. Then, defining

$$\phi_{j,\mathbf{n}}(\mathbf{x}) = |D|^{j/2}\phi(D^j\mathbf{x} - \mathbf{n}) \quad \text{and} \quad V_j = \overline{\text{span}\{\phi_{j,\mathbf{n}} \mid \mathbf{n} \in \mathbb{Z}^n\}}$$

as for $n = 1$, we have

1. $\lim_{j\to+\infty} V_j$ is dense in $L^2(\mathbb{R}^n)$ and $\lim_{j\to-\infty} V_j = \{0\}$;
2. $(\phi_{0,\mathbf{n}})_{\mathbf{n}\in\mathbb{Z}^n}$ provides a Riesz basis for V_0.

Under these conditions, the existence of multidimensional wavelets is ensured by the theorem of Meyer [10].

Theorem 5.1. *There exist* $|D| - 1$ *wavelets*

$$\psi^1(\mathbf{x}), \psi^2(\mathbf{x}), \ldots, \psi^{|D|-1}(\mathbf{x}),$$

whose lattice translates by $\mathbb{Z}^n$ *provide an orthonormal basis for the orthogonal complement of* V_0 *in* V_1. *Especially, the following functions constitute an orthonormal basis of* $L^2(\mathbb{R}^n)$*:*

$$\{\psi^k_{j,\mathbf{n}} \quad : \quad 1 \le k \le |D| - 1, j \in \mathbb{Z}, \mathbf{n} \in \mathbb{Z}^n\}.$$

Therefore, if D is tight, one wavelet is still sufficient to encode the information of $L^2(\mathbb{R}^n)$.

Definition 5.2 (A-unitary). Suppose A is an n by n positive definite matrix. A non-singular real matrix D is said to be *A-unitary* if

$$D^T A D = |D|^{\frac{2}{n}} A.$$

Notice that D is A-unitary if and only if $D^T A D = dA$ for some non-negative scalar d.

If A is a scalar matrix $\sigma^2 I_n$, then an A-unitary matrix D is also said to be *isotropic*. In fact, D is isotropic if and only if $D/|\det D|^{1/n}$ is an ordinary orthogonal matrix.

Definition 5.3 (Multidimensional Heatlets). Suppose the heat equation (5.2) is given. Let

$$\psi^1(\mathbf{x}), \psi^2(\mathbf{x}), \ldots, \psi^{|D|-1}(\mathbf{x})$$

be the wavelets for a multiresolution of $L^2(\mathbb{R}^n)$ with an A-unitary dilation matrix D. Then the heat evolutions of the wavelets, denoted separately by

$$\Psi^1(\mathbf{x}, t), \Psi^2(\mathbf{x}, t), \ldots, \Psi^{|D|-1}(\mathbf{x}, t),$$

are called the *heatlets* of the heat equation with respect to the given multiresolution.

If, in addition, D is tight, then there is only one heatlet. As in the 1-D case, we denote it by $\Psi^h(\mathbf{x}, t)$. Similary, the refinable heat, i.e. the heat evolution of the scaling function $\phi(\mathbf{x})$, is denoted by $\Phi^h(\mathbf{x}, t)$.

Proposition 5.1 (Self-Similarity). *Suppose that $\phi(\mathbf{x})$ and $\psi^k(\mathbf{x}), k = 1, 2, \cdots, |D| - 1$ satisfy*

$$\phi(\mathbf{x}) = |D| \sum_{\mathbf{n} \in \mathbb{Z}^n} h_{\mathbf{n}} \phi(D\mathbf{x} - \mathbf{n}),$$
$$\psi^k(\mathbf{x}) = |D| \sum_{\mathbf{n} \in \mathbb{Z}^n} g^k_{\mathbf{n}} \phi(D\mathbf{x} - \mathbf{n}),$$

Then the refinable heat $\Phi^h(\mathbf{x}, t)$ and heatlets $\Psi^k(\mathbf{x}, t)$ satisfy the similarity relations

$$\Phi^h(\mathbf{x}, t) = |D| \sum_{\mathbf{n} \in \mathbb{Z}^n} h_{\mathbf{n}} \Phi(D\mathbf{x} - \mathbf{n},\ dt), \tag{5.4}$$

$$\Psi^k(\mathbf{x}, t) = |D| \sum_{\mathbf{n} \in \mathbb{Z}^n} g^k_{\mathbf{n}} \Phi(D\mathbf{x} - \mathbf{n},\ dt), \tag{5.5}$$

with $k = 1, 2, \cdots, |D| - 1$, and $d = |D|^{2/n} > 1$.

Proof. Similar to the 1-D case. Notice that since D is A-unitary, $\Phi(D\mathbf{x}-\mathbf{n},\ dt)$ is the heat evolution of $\phi(D\mathbf{x}-\mathbf{n})$. □

If d is an integer, then $\Phi^h(\mathbf{x},t)$ is an $n+1$-dimensional refinable function corresponding to the dilation matrix

$$\begin{bmatrix} D & 0 \\ 0 & d \end{bmatrix}.$$

For example, if $n=4$ and $|D|=4$, then $d=2$. If D is tight, i.e. $|D|=2$, then only when $n=1$ or 2, can d be an integer.

Remark 6. The A-unitary condition on D is critical for the simple self-similar relation (5.4) to be valid. From another angle, the condition combines the (scale) invariant properties of the heat equation and wavelets. This seems to us a significant link between PDE and wavelet theory.

EXAMPLE.

Consider the simplest 2-D heat equation:

$$u_t=\frac{\sigma^2}{2}(u_{xx}+u_{yy}),\quad u(\mathbf{x},0)=f(\mathbf{x}).$$

Here $\mathbf{x}=(x,y)'$. The weight matrix $A=\sigma^2 I_2$ is a scalar matrix. Therefore, a dilation factor D must be isotropic, or equivalently, $D/\sqrt{|D|}$ is an ordinary orthogonal matrix. Especially, if D is tight, then up to a scalar multiple of -1, D must be

$$R=\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix},$$

or its transpose. Notice that all its eigenvalues lie on the circle $|z|=\sqrt{2}$. A refinement equation associated to R has the following form

$$\phi(\mathbf{x})=2\sum_{\mathbf{n}\in\mathbb{Z}^n} h_{\mathbf{n}}\phi(R\mathbf{x}-\mathbf{n}).$$

Set $\lambda=1+i$, $z=x+iy$ if $\mathbf{x}=(x,y)'$, and $n_z=n+im$ if $\mathbf{n}=(n,m)'$. Then in the notation of complex numbers, the refinement equation simplifies to

$$\phi(z)=2\sum_{n_z} h_{n_z}\phi(\lambda z-n_z).$$

If one chooses the Haar filter: $h_0 = h_1 = 1/2$, and $h_{n_z} = 0$ for the remaining lattice points n_z, then the equation becomes

$$\phi(z) = \phi(\lambda z) + \phi(\lambda z - 1).$$

The solution is the characteristic function of a fractal set famously known as the *twin dragon* (see Louis et al. [9]).

Under the setup of Definition 5.3, we can now state the multi-dimensional heatlet decomposition theorem.

Theorem 5.2 (n-D heatlet decomposition). *Let* $\Phi^k(\mathbf{x}, t)$, $k = 1, \cdots, |D| - 1$ *be the heatlets of the given heat equation* (5.2). *Assume* $f \in L^2(\mathbb{R}^n)$ *has the following wavelet decomposition*

$$f(\mathbf{x}) = \sum_{k,j,\mathbf{n}} c^k_{j,\mathbf{n}}(f)\psi^k_{j,\mathbf{n}}(\mathbf{x}).$$

Then the heat evolution of $f(\mathbf{x})$ *(in* $L^2(\mathbb{R}^n)$*) can be decomposed into*

$$u(\mathbf{x}, t) = \sum_{k,j,\mathbf{n}} c^k_{j,\mathbf{n}}(f)\Psi^k_{j,\mathbf{n}}(\mathbf{x}).$$

Here,

$$\Psi^k_{j,\mathbf{n}}(\mathbf{x}, t) = |D|^{j/2}\Psi^k(D^j\mathbf{x} - \mathbf{n}, d^j t), \quad d = |D|^{2/n}.$$

The probability method and subdivision algorithm apply to the n-D case similary.

6. Conclusions

First we emphasize again that a wavelet basis is invariant under a scaling ($\mathbf{x} \to D\mathbf{x}$) and translation ($\mathbf{x} \to \mathbf{x} - \mathbf{n}$).

A first order (in time) linear evolution equation takes the form

$$au_t(\mathbf{x}, t) + L_{\mathbf{x}}u(\mathbf{x}, t) = 0, \qquad u(\mathbf{x}, 0) = f(\mathbf{x}),$$

where a is a scalar and $L_{\mathbf{x}}$ is a linear spatial differential operator with constant coefficients. Then $L_{\mathbf{x}}$ is automatically translation invariant, i.e.

$$L_{\mathbf{x}}[g(\mathbf{x} + \mathbf{c})] = [L_{\mathbf{x}}g](\mathbf{x} + \mathbf{c}),$$

for any real vector $\mathbf{c}$. Therefore, to take the full advantage of a wavelet basis, it is necessary for $L_{\mathbf{x}}$ to be also scaling invariant in the sense that

$$L_{\mathbf{x}}[g(D\mathbf{x})] = \alpha[L_{\mathbf{x}}g](D\mathbf{x}),$$

where α is a scalar universally for $g(x)$. If so, the factor α can be absorbed into the temporal variable t (by a proper scaling). The A-unitary condition introduced for the multi-dimensional case meets this requirement.

Based on this idea, we make the following closing remarks.

(a) The method works equally well for the following Schrödinger equation with a zero potential:

$$iu_t(\mathbf{x},t) + \Delta u(\mathbf{x},t) = 0, \quad \mathbf{x} \in \mathbb{R}^n, \quad t > 0,$$

with an initial condition $u(\mathbf{x},0) = f(\mathbf{x})$. We require the dilation factor D to be isotropic. For the 1-D case, one can simply take $D = 2$. The kernel function of the equation becomes

$$R_t(\mathbf{x}) = \frac{1}{(2\sqrt{\pi t})^n} e^{-i\frac{\pi n}{4}} e^{i\frac{|\mathbf{x}|^2}{4t}}.$$

Its convolution with a scaling function (or a wavelet) is an integral of Fresnel type and possesses the character of self-similarity (see Proposition 3.1).

(b) The method also works for the homogeneous wave equation:

$$u_{tt}(\mathbf{x},t) - c^2\Delta u(\mathbf{x},t) = 0, \quad \mathbf{x} \in \mathbb{R}^n, \quad t > 0,$$

We can study its *zeroth order* wavelet fundamental solution corresponding to the initial condition

$$u(\mathbf{x},0) = \psi(\mathbf{x}), \qquad u_t(\mathbf{x},0) = 0;$$

and the *first order* wavelet fundamental solution corresponding to the initial state

$$u(\mathbf{x},0) = 0, \qquad u_t(\mathbf{x},0) = \psi(\mathbf{x}).$$

Again, we require the dilation factor D to be isotropic.

(c) We can study the Laplace equation (in the upper-half plane) in the same manner. The interaction between the Laplace equation and wavelets has special meaning and is discussed in Shen [13].

References

[1] G. Beylkin, R. Coifman, and V. Rohklin. Fast wavelet transforms and numerical algorithms, I. *Comm. Pure. Appl. Math.*, 44:141–183, 1991.

[2] I. Daubechies. *Ten Lectures on Wavelets.* SIAM, Philadelphia, 1992.

[3] G. Derfel. Probabilistic method for a class of functional-differential equations. *Ukrain. Math. J.*, 41:1117–1234, 1989.

[4] G. Derfel, N. Dyn, and D. Levin. Generalized refinement equations and subdivision processes. *J. Approx. Theory*, 80:272–297, 1995.

[5] A. Haar. Zur theorie der orthogonalen Funktionen-Systeme. *Math. Ann.*, 69:331–371, 1910.

[6] H. Helson. *Harmonic Analysis.* Addison-Wesley, Reading, Massachusetts, 1983.

[7] E. Hernández and G. Weiss. *A First Course on Wavelets.* CRC Press, 1996.

[8] A. Kolmogoroff, J. Petrovsky, and N. Piscounoff. Étude de l'équation de la diffusion avec croissance de la quantité de matière et son application à une problème biologique. *Bulletin de l'université d'état à Moscou, Série Int. VI*, pages 1–25, 1937.

[9] A. Louis, P. Maaβ, and A. Rieder. *Wavelets: Theory and Applications.* John Wiley & Sons, 1997.

[10] Y. Meyer. *Wavelets and Operators.* Cambridge University Press, 1992.

[11] J. Shen. Refinement differential equations and wavelets. *Methods Appl. Anal.*, 5(3):283–316, 1998.

[12] J. Shen. Combinatorics for wavelets: the umbral refinement equation. *Studies Applied Math.*, 102(1999.

[13] J. Shen. On Cauchy generating functions and analytic wavelets. *Submitted*, 1999.

[14] G. Strang and T. Nguyen. *Wavelets and Filter Banks.* Wellesley-Cambridge Press, Wellesley, MA, 1996.

Department of Mathematics, 7354 Math. Sci. Building, UCLA, Los Angeles, CA 90095.
Department of Mathematics, 2-240, MIT, Cambridge, MA 02139