A Note on Wavelets and Diffusions

J. Shen *

School of Mathematics,
University of Minnesota
Minneapolis, MN 55455
jhshen@math.umn.edu

Abstract

Motivated by image processing and numerical wavelet methods for partial differential equations, we study the theoretical interactions between wavelets and the diffusion equations. Important properties of wavelets, such as the translation and scaling invariance, the p-vanishing-moment condition, and the atomic decomposition, are integrated into the diffusion process and lead to many interesting results.

Key Words: Wavelets, diffusion, heatlets, annihilation.

1 Introduction

Despite the existence of many earlier similar ideas, the modern popularity of wavelets was very much inspired by the thirsty for efficient and characteristic representations of temporally or spatially varying data, such as seismic signals or digital images. Take images for example. Wavelets not only provide an innovative way to encode images and thus an efficient way of Internet transmission (the JPEG-2000 standard), but also lead to a preliminary low-level model for the human vision system and deeper insights into computer vision (Field [13]).

The criterion for the effectiveness of data representation depends on the concrete task, or the corresponding "operator," mathematically speaking. For image analysis, the ultimate operator is the human visual perception. The fact that wavelets representation of visual data is so effective is mostly due to the intrinsic link between the wavelets/multiresolution framework and the configuration of visual neurons in our visual system (Field [13]). For other operators, wavelets representation may be less straightforwardly effective.

One major class of mathematically interesting operators are evolutionary differential equations. The input data is the initial value condition, and the output

*Research supported by NSF Grant DMS-0202565. *J. Comp. Anal. Appl.*, **5**(1): 147-159, 2003.

is its evolutionary behavior at different times. As for images, wavelets can typically lead to sparse representation (important for numerical computations) of certain classes of differential operators (Beylkin, Coifman and Rohklin [2]). On the other hand, the theoretical link between wavelets and differential operators is still weak. This is mostly "due to" Fourier. Fourier modes or harmonic waves are the *eigenstates* of the basic differential operators ∂/∂_x, ∂/∂_y, $\cdots$ In this sense, wavelets cannot compete with harmonic waves. Or, from the quantum mechanical point of view, the harmonic waves are the eigenstates of the momentum operator $p = -i\partial/\partial_x$, in which, the wave-particle duality completely favors the wave behavior. For wavelets, to the best knowledge of the author, such physical interpretation has not been found. Moreover, from the design point of view, except for the *divergence-free* wavelets (Battle et al. [1]), most wavelets constructions are not directly motivated by differential operators.

The curiosity over the connection between wavelets and differential equations never fades away, however. Most work has been focused on numerical computations (for example, the *wavelet homogenization* technique (Engquist [12], Brewster and Beylkin [3])). The theoretical aspect is still understood very little. In a recent paper by Shen and Strang [25], the connection between wavelets and the heat equation is studied from the viewpoint of translation and scaling invariance. The present paper can be seen as its continuation, in which, we shall emphasize the other two major properties of wavelets — *locality* and the *cancelation condition* or the *vanishing-moment condition* (Daubechies [10], Strang and Nguyen [26]), which interact with the *locality* and *smoothing* properties of the diffusion equations.

Our special interest in the connection between wavelets and the diffusion equations are motivated by both numerical computations and image processing. First, by theoretically understanding the evolutionary behavior of the wavelet components (i.e. heatlets), we can gain better *a priori* knowledge on the numerical solution when we apply wavelets based numerical methods (such as the Wavelet-Galerkin method). The importance of such work can be seen from its classical counterpart in other numerical PDE methods such as the finite difference method. For a simple example, by understanding the propagation behavior of the wave equations (diffusion equations in our case) and its effect on the discrete grid (the wavelet basis in our case), we may *a priori* know the importance of choosing the *upwind* difference scheme (see, for example, Golub and Ortega [14]). The second motivation is the PDE methods in image processing (Perona and Malik [19], Rudin, Osher and Fatemi [20], Morel and Solimini [17], Chan and Shen [5]). Various nonlinear and anisotropic diffusion equations have found wide applications in image processing, such as denoising, restoration, and segmentation. In the current literature, however, the PDE method and the wavelet/multiresolution method are quite separated, and a joint understanding is necessary if we intend to combine the advantages of both methods. Some preliminary work has been recently carried out by Coifman and Sowa [8, 9] and Chui and Wang [7]. The present paper is also intended for this important task by first attempting to understand the interactions between wavelets/multiresolution analysis and the linear diffusion equations. This shall serve as a step-stone for the next work on anisotropic diffusions.

This short note consists of two parts. The first one basically reviews the recent work of Shen and Strang [25] on the properties of heatlets (Section 2). The second part studies the diffusion annihilation behavior of wavelets (Section 3), where we discuss the influence of the locality and vanishing-moment conditions of wavelets on linear diffusion equations.

2 Wavelets and Heatlets

In this section, we surf the main ideas in Shen and Strang [25] on integrating wavelets into the spatially homogeneous heat equation

$$u_t(t, \mathbf{x}) = D\Delta u(t, \mathbf{x}), \tag{1}$$

where $\mathbf{x} = (x_1, x_2, \cdots, x_n)$ is the spatial variable and D the diffusivity coefficient. For simplicity, in what follows, we shall only address on the 1-dimensional spatial domain $\mathbb{R}$.

The heat equation can be treated as an operator acting on an input data, namely the initial heat distribution $u(0, x) = f(x)$. The general philosophy expressed in the introduction section naturally leads us to considering a wavelet representation of the input data $f(x)$:

$$f(x) = \sum_{j,k \in \mathbb{Z}} c_{jk}(f) \psi_{j,k}(x), \tag{2}$$

where ψ is a mother wavelet in $L^2(\mathbb{R})$ associated with an orthogonal multiresolution analysis (Daubechies [10]) and

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

with j and k indexing the scale levels and spatial locations.

The wavelet representation differs from two of the most popular and well-practiced representations in the classical theory of PDEs. The first is the Dirac representation:

$$f(x) = \int_{\mathbb{R}} \delta(x - y) f(y) dy, \tag{3}$$

where $\delta(x)$ is the Dirac point source at origin. And the second is the celebrated Fourier representation:

$$f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{ikx} \hat{f}(k) dk, \tag{4}$$

where $\hat{f}(k)$ is the Fourier transform of $f(x)$ and k the *wave-number* variable. The Dirac representation leads to the classical *fundamental solution* (or the Green's function) $G(t, x; y)$, videoing the heat evolution of an initial point source $\delta(x - y)$ located at y. While the Fourier representation reveals the evolution of each harmonic mode e^{ikx} and its associated decaying factor $e^{-Dk^2 t}$.

In this spirit, in [25], we call the heat evolution of an initial wavelet distribution $\psi(x)$ a *heatlet*, and denote it by $\Psi(t, x)$.

The starting point of [25] is the *scaling* and *translation* invariance shared by both a wavelet basis and the heat equation. A wavelet basis

$$\Sigma = \{\psi_{j,k} \mid j, k \in \mathbb{Z}\}$$

is both scaling and translation invariant in the sense that

(i) If $w(x) \in \Sigma$, then $\sqrt{\lambda} w(\lambda x) \in \Sigma$ for any $\lambda = 2^j$, $j \in \mathbb{Z}$;

(ii) Suppose that $w(x) \in \Sigma$ is at the scale level j. Then for any dyadic node $x_0 = k2^{-j}$ at scale j, $w(x - x_0) \in \Sigma$.

On the other hand, the heat equation is apparently both scaling and translation invariant with respect to the spatial variable. That is, for any positive scale parameter λ and real translation parameter x_0, the equation is invariant under the affine transform

$$x \to \lambda x - x_0,$$

though the diffusivity constant D is indeed updated, or it is strictly invariant under the combined temporal and spatial transform with different scaling factors:

$$x \to \lambda x - x_0, \qquad t \to \lambda^2 t. \tag{5}$$

The remarkable result from the interaction of these invariant properties is the existence of a single fundamental heat diffusion – the heatlet $\Psi(t, x)$ introduced above. All heat diffusion processes can be assembled from it in the way that a wavelet serves as the building block for L^2.

First, the heatlet $\Psi(t, x)$ satisfies a *two-scale* similarity relation, explained as follows. Let $\phi(x)$ and $\psi(x)$ be the associated mother scaling function and wavelet for a multiresolution. Then there exist a set of "lowpass filter" coefficients

$$\{h_n \mid n \in \mathbb{Z}\},$$

and "highpass filter" coefficients

$$\{g_n \mid n \in \mathbb{Z}\},$$

(usually both are only finitely many non-zero; see Daubechies [10]) such that,

$$\phi(x) = 2 \sum_{n \in \mathbb{Z}} h_n \phi(2x - n), \tag{6}$$

$$\psi(x) = 2 \sum_{n \in \mathbb{Z}} g_n \phi(2x - n). \tag{7}$$

These are traditionally called *two-scale relations.* Denote the heat diffusion of $\phi(x)$ by $\Phi(t,x)$ and define

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

Suppose we symmetrically extend both $\Phi(t,x)$ and $\Psi(t,x)$ to negative t values. That is, for example,

$$\Psi(-t,x) := \Psi(t,x), \qquad t \geq 0.$$

Then $\Phi(t,x)$ and $\Psi(t,x)$ as functions on $\mathbb{R}^2$ also satisfy the 2-D (space-time) two-scale relations

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

Here $h_{\mathbf{n}} = 0$ unless when $\mathbf{n} = (0,n)'$, in which case $h_{\mathbf{n}} = h_n$. For this reason, $\Phi(t,x)$ is called a "refinable heat" (Shen and Strang [25]).

Secondly, the atomic role of a heatlet is explained as follows. Storing the heatlet $\Psi(x)$ in computer memory *a priori*, for each pair of integers j and n, one generates

$$\Psi_{j,n}(t,x) := 2^{j/2}\Psi(4^j t, 2^j x - n),$$

via spatial translation and spatial-temporal scaling from the mother heatlet. Then we have the following *heatlet decomposition*.

Theorem 1 (Heatlet Decomposition. Shen and Strang [25]) *Let $f(x) \in L^2(\mathbb{R})$ be an initial heat distribution. Then its heat evolution is given by*

$$u(t,x) = \sum_{j,n \in \mathbb{Z}} c_{j,n}(f) \Psi_{j,n}(t,x),$$

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

From the computational point of view, for a given initial heat distribution $f(x)$, to view its heat evolution, one can first apply the fast wavelet transform to $f(x)$ to obtain its wavelet coefficients $c_{j,k}(f)$ (i.e. the *analysis stage*); then the *synthesis stage* retrieves the heatlet $\Psi(t,x)$ in storage and assembles the heat evolution according to this heatlet decomposition theorem.

Shen and Strang [25] also showed that, by taking advantage of the probability interpretation [21, 22] of both the dilation equation and the heat equation, one can apply the continuous subdivision scheme for the construction of the mother heatlet.

3 Wavelets Annihilate Diffusions

In this section, we study the interaction between the smoothing effect of the diffusion process and the annihilation property of wavelets.

3.1 Law of diffusion and wavelets annihilation

The fundamental law associated with the diffusion equation is (see Einstein [11] for example)

Theorem 2 (Square Root Law) *An exceptionally "hot" spot (in terms of heat diffusion) or an exceptionally dense spot (in terms of mass or chemical diffusion) spreads out in a speed proportional to the square root of time.*

The spreading rate is evaluated in the root-mean-square sense. The law results from the Brownian motion microscopically. The importance of the law lies in its linkage of space and time. Because of this law, currently in image processing, there is a booming field called the scale-space theory (see the monograph by Morel and Solimini [17]). One major tool in the scale-space theory is various types of diffusions. In smooth regions of images, time in the diffusions (linear or non-linear, isotropic and anisotropic) automatically provides the scale information due to the Square Root Law. This is perhaps a helpful digression, since recently several authors have been devoted to the joint understanding of wavelets and the PDE method in image processing (see, for instance, Osher and Shen [18], Coifman and Sowa [8, 9], Chan and Zhou [6]).

The Square Root Law also applies to any local perturbation with *positive* concentrations (of heat or mass). It may easily be taken for granted that the law applies to *any* local perturbation. In fact the latter is not true. Wavelet perturbations are the best counterexamples. It is the interaction between the *smoothing* property of the diffusion operator and the *annihilation* action of wavelets that is taking effect.

Wavelets are typically designed to have the *annihilation* property:

$$\int_{\mathbb{R}} x^k \psi(x) dx = 0, \qquad k = 0, 1, \cdots, p-1. \tag{8}$$

Meyer [16] called it the *cancelation* or *oscillation* condition. In the wavelet literature, this is also famously referred to as the p-vanishing-moment condition (Strang and Nguyen [26]). We shall call it here the *annihilation* condition — a wavelet annihilates polynomials, thus smoothes data to some degree. It is directly connected to the smoothness of the wavelets for orthogonal wavelets (Daubechies [10]; Cai and Shen [4]), and the approximation order of the associated multiresolution (Strang and Nguyen [26]).

3.2 Wavelet annihilates diffusion: an example

Our main task in this section is to quantitatively characterize the following "theorem."

"Theorem 3" An annihilating wavelet also annihilates diffusion.

This "theorem" is construed (and proved) first through the following example. General discussion will be given in the next subsection.

Consider a system in an equilibrium state $f_0(x) = 1$ (for a diffusion system, an equilibrium state must be constant so that no flux can exist). Now suppose at time $t = 0$, $f_0(x)$ is perturbed to

$$f(x) = f_0(x) + \psi_{j,k}(x),$$

by a wavelet at the scale level j (say $j = 8$) and located at $x_0 = k2^{-j}$. (Note that in real situations of diffusions, we may require that $f(x)$ be nonnegative to maintain the physical meaning (of mass or heat). But apparently this causes no problem in the analysis that follows.) In addition, we assume that the mother wavelet $\psi(x)$ has exactly p vanishing moments.

Let us see how the perturbation is diffused. By linearity, the diffusion $u(t, x)$ of $f(x)$ consists of two parts accordingly:

$$u(t, x) = u_0(t, x) + u_p(t, x),$$

where $u_0 \equiv 1$ is the diffusion of f_0, and $u_p(t, x)$ that of the perturbation due to $\psi_{j,k}$. Notice that $u_p(t, x)$ is exactly a heatlet. Define $\lambda = 2^{-j}$ to be the scale associated with the perturbation. According to the Square Root Law, it is appropriate to define a critical time

$$t_c = \lambda^2.$$

If the scale of the wavelet perturbation is very small (i.e. $j \gg 1$), so is t_c. Suppose the diffusion time $t \gg t_c$ relatively (though t itself can be also small). Then we will observe the "annihilation" of the diffusion in the following sense.

For simplicity, consider only the case when the diffusivity coefficient is homogeneous and

$$D(x) \equiv 1.$$

Let $G_t(x)$ denote a Gaussian with variance $2t$:

$$G_t(x) := \frac{1}{2\sqrt{\pi t}} \exp\left(-\frac{x^2}{4t}\right).$$

Then $G_t(x - y)$ is the Green's function, and the perturbed component $u_p(t, x)$ (a heatlet) is given by

$$u_p(t, x) = \int_{\mathbb{R}} G_t(x - y)\psi_{j,k}(y)dy. \tag{9}$$

Introduce new variables

$$\epsilon = \sqrt{\frac{t_c}{2t}} \ll 1, \qquad X = \frac{x - x_0}{\sqrt{2t}}, \qquad Y = \frac{y - x_0}{\sqrt{2t}}.$$

Then

$$u_p(t,x) = \frac{1}{\sqrt{2\pi t_c}} \int_{\mathbb{R}} e^{-\frac{(X-Y)^2}{2}} \psi\left(\frac{Y}{\epsilon}\right)\, dY. \tag{10}$$

Without loss of generality, assume that $\psi(x)$ is a compactly supported wavelet (Daubechies' family is the famous example [10]). Suppose that the support is $[a,b]$. Let $H_p(x)$ denote the p-th order Hermit polynomial (orthogonal with respect to the weight $e^{-x^2/2}$; see for example, Wiener [27]) defined by

$$H_p(x) = e^{\frac{x^2}{2}} \frac{d^p}{dx^p} e^{-\frac{x^2}{2}}.$$

Then

$$\begin{aligned}
u_p(t,x) &= \frac{1}{\sqrt{2\pi t_c}} \int_{\epsilon a}^{\epsilon b} e^{-\frac{(X-Y)^2}{2}} \psi\left(\frac{Y}{\epsilon}\right)\, dY. \\
&\simeq \frac{1}{\sqrt{2\pi t_c}} \int_{\epsilon a}^{\epsilon b} \frac{1}{p!} \left[\frac{d^p}{dx^p} e^{-\frac{x^2}{2}}\right]_{x=X} Y^p \psi\left(\frac{Y}{\epsilon}\right)\, dY \\
&= \frac{1}{\sqrt{2\pi t_c}} e^{-\frac{X^2}{2}} H_p(X)\, \frac{\epsilon^{p+1}}{p!} \int_a^b s^p \psi(s)\, ds \\
&= \epsilon^p \cdot \frac{M_p}{p!\sqrt{\pi}} \cdot \frac{1}{\sqrt{2t}} h_p\left(\frac{x-x_0}{\sqrt{2t}}\right),
\end{aligned} \tag{11}$$

where,

(a) M_p is the p-th moment of the wavelet, which is non-zero by assumption.

(b) $h_p(x)$ is the p-th weighted (with weight $e^{-x^2/4}$) Hermitian function $e^{-x^2/2}H_p(x)$, which we shall call the p-th mother *Hermitlet*. As in wavelet theory, we shall call any function

$$\frac{1}{\sqrt{2t}} h_p\left(\frac{x-x_0}{\sqrt{2t}}\right)$$

a *Hermitlet* of scale $\sqrt{2t}$ and located at x_0.

In the second step of Eq. (11), "$\simeq$" means the leading term with respect to the small parameter ϵ.

At a first glance, it may sound a surprising fact, that up to a decaying multiplicative constant (depending on time), a heatlet must be an Hermitlet in the leading term, after the diffusion scale exceeds the scale of the initial wavelet. The fact is independent of the particular form of the wavelet, as long as the mother wavelet has exactly p vanishing moments! Apparently this universal asymptotic result is some kind of "far-field" behavior of the diffusion equation.

The quantitative meaning of Theorem 3 is then given by Eq. (11). The Hermitlet part

$$\frac{1}{\sqrt{2t}} h_p\left(\frac{x-x_0}{\sqrt{2t}}\right)$$

is a regular diffusion process whose spatial L^1 norm is invariant with time. (If h_p were non-negative, then it would be due to the *conversation law* of mass or heat diffusions.) What we have called the "annihilation" action in Theorem 3 is quantized by the *annihilation factor*

$$\epsilon^p = \left(\frac{t_c}{2t}\right)^{\frac{p}{2}} = \left(\frac{\lambda}{\sqrt{2t}}\right)^p ,$$

which diminishes both the L^1 and L^2 norms of the perturbation. For instance, suppose the initial wavelet perturbation is of scale $j = 8$ and with $p = 3$ vanishing moments (very practical numerical values). Consider its diffusion at $t = 0.03125 = 1/32$. Then the annihilation factor is

$$\epsilon^p = \left(\frac{2^{-8}}{\sqrt{2 \times 0.03125}}\right)^3 = 2^{-18} \simeq .000004.$$

which is so small that in ordinary numerical computations, the diffusion of the perturbation can be ignored.

This result on a single wavelet mode is easily generalized via the multiresolution setting, which is carried out in the coming section.

3.3 Wavelet annihilates diffusion: multiresolution setting

Generally, suppose that a system originally in an equilibrium state (0, say) undergoes a perturbation $f(x)$ with scale λ near a spot at $t = 0$ (for example, by a local heating up or the injection of new chemical substances). Without loss of generality, assume $\lambda = 1 = 2^0$. Let ϕ and ψ be the mother scaling function and wavelet of an orthogonal multiresolution analysis:

$$V_0 \subset V_1 \subset \cdots \subset V_j \cdots ,$$

with

$$V_j = \text{Closure}\left\{\text{span}\{\phi_{j,k} \;\middle|\; k \in \mathbb{Z}\}\right\}.$$

Let W_j be the j-th wavelet space, orthogonally complementing V_j in V_{j+1}:

$$V_{j+1} = V_j \oplus W_j.$$

Let P_j and Q_j denote the two orthogonal projections from L^2 onto V_j and W_j. Then the multiresolution analysis is given by

$$P_j = P_0 + Q_0 + Q_1 + \cdots + Q_j,$$

and

$$f(x) = P_0 f + Q_0 f + \cdots + Q_j f + \cdots .$$

We assume that $\psi(x)$ has p vanishing moments. Therefore $f_j = P_j f$ converges to f with error $O(2^{-j(p+1/2)})$ if $f(x)$ is smooth enough. Hence as in image processing and data compression, a finite sum of resolutions is often enough

$$f(x) = P_0 f + Q_0 f + \cdots + Q_J f. \tag{12}$$

For instance, in digital image processing, J is typically near 8. The approximation order is preserved in the diffusion process because of the Young's Inequality on convolutions [15].

The "flat" component at the coarsest scale $f_0 = P_0 f$ stores the net mass of the perturbation, i.e.

$$\int_{\mathbb{R}} f(x)dx = \int_{\mathbb{R}} P_0 f(x)dx.$$

This quantity (usually non-zero) is preserved during the diffusion process due to the *conservation law.* The diffusion of $P_0 f$ is given by the refinable heats (see Section 2) and obeys the Square Root Law.

Since the perturbation is local: $\lambda = 1$, $Q_j f$ at j-th fine scale is a linear combination of approximately $p2^j$ wavelets at scale j:

$$Q_j f = \sum_k c_{jk}(f)\psi_{jk}(x).$$

Consider a given time $0 < t < 1$. Define integer

$$j_t = \lceil -\frac{1}{2}\log_2 t - \frac{1}{2} \rceil.$$

Then according to Eq. (11), for all scale levels $j > j_t$, the leading term of the diffusion of $Q_j f$ is given by

$$[\epsilon_j(t)]^p \cdot \frac{M_p}{p!\sqrt{\pi}} \cdot \sum_k c_{j,k}(f)\frac{1}{\sqrt{2t}} h_p\left(\frac{x - k2^{-j}}{\sqrt{2t}}\right),$$

as long as

$$\epsilon_j(t) = \frac{2^{-j}}{\sqrt{2t}} \leq 2^{-(j-j_t)} \ll 1.$$

This remarkable result says that after the diffusion scale $\sqrt{2t}$ exceeds the scale of a wavelet component, the diffusion magnitude is annihilated by a factor ϵ^p, and its shape is linearly assembled from *Hermitlets* (in the leading term).

3.4 A remark on the universality of the method

As well-known in the literature, in some sense, general wavelet bases "interpolate" the classical two extremal ends, namely, the Haar wavelet basis (rough but with compact supports) and the Shannon sinc wavelet basis (entire functions but with a hyperbolic decaying). (The sense of interpolation is better felt from the asymptotic results by Shen and Strang [23] [24] on Daubechies' minimal phase orthogonal wavelets.) Therefore, they reach an optimality between the locality in the physical domain and in the Fourier domain, and lead to a much more flexible tool of analysis.

The way we have proved the annihilation property benefits from this flexibility. The same results can be obtained by going to the Fourier domain. But the major advantage of the physical-domain method is that it also applies to more general linear diffusion processes,

$$u_t = \nabla \cdot (D(x)\nabla u) , \tag{13}$$

where the diffusion coefficient $D(x)$ is location dependent. Under certain mild conditions on the regularity of $D(x)$, the Green's function is still smooth enough and thus allows a Taylor expansion of order $p+1$. Therefore, formula (11) still holds with a location dependent function h_p. The annihilation factor remains the same, meaning that Theorem 3 also holds for general diffusions given by Eq. (13).

Acknowledgment

The author is pleased to thank Professors Gilbert Strang, Stanley Osher and Tony Chan, for their constant support and encouragement. He also wishes to thank the two referees for their comments and generous help.

References

[1] G. Battle, P. Federbush, and P. Uhlig. Wavelets for quantum gravity and divergence-free wavelets. *Appl. Comp. Harmon. Anal.*, 1(3):295–297, 1994.

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

[3] M. E. Brewster and G. Beylkin. A multiresolution strategy for numerical homogenization. *Appl. Comp. Harmon. Anal.*, 2(4):327–349, 1995.

[4] T. Cai and J. Shen. Boundedness is redundant in a theorem of Daubechies. *Appl. Comp. Harmon. Anal.*, 6:400–404, 1999.

[5] T. Chan and J. Shen. Variational restoration of non-flat image features: models and algorithms. *SIAM J. Appl. Math.*, 61(4):1338-1361, 2000.

[6] T. Chan and H.-M. Zhou. Adaptive ENO-wavelet transforms for discontinuous functions. CAM report 99-21, Math. Dept., UCLA, 1999.

[7] C. Chui and J.-Z. Wang. Spline wavelets in Sobolev spaces and applications to numerical computation of nonlinear diffusion equations in image analysis. Talk at *AMS sectional meeting at Austin* and private communication, Oct. 1999.

[8] R. R. Coifman and A. Sowa. Combining variational calculus and wavelets for image enhancement. *Appl. Comp. Harmon. Anal.,* to appear, 1999.

[9] R. R. Coifman and A. Sowa. New types of Total Variation diminishing flows. Preprint, 1999.

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

[11] A. Einstein. On the movement of small particles suspended in a stationary liquid demanded by the molecular kinetic theory. *Ann. Phys.*, 17, 1905.

[12] B. Engquist. Wavelet based numerical homogenization. *Doc. Math.,* Extra Volume ICM 1998, pages 503–512, 1998.

[13] D. Field. *Wavelets, Fractals, and Fourier Transforms*, chapter Scale-invariance and self-similar "wavelets" transforms: an analysis of natural scenes and mammalian visual systems. Clarendon Press, Oxford, 1993.

[14] G. H. Golub and J. M. Ortega. *Scientific Computing and Differential Equations.* Academic Press, 1992.

[15] J. T. Marti. *Introduction to Sobolev Spaces and Finite Element Solution of Elliptic Boundary Value Problems.* Computational Mathematics and Applications. Academic Press, 1986.

[16] Y. Meyer. *Wavelets and Operators.* Cambridge studies in advanced math. Cambridge University Press, 1992.

[17] J.-M. Morel and S. Solimini. *Variational Methods in Image Segmentation.* Birkhäuser, Boston, 1995.

[18] S. Osher and J. Shen. Digitized PDE method for data restoration. *Analytic-Computational Methods in Applied Mathematics*, Ed. G. A. Anastassiou, pp. 751-771, Chapman & Hall/CRC, FL, 2000.

[19] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Anal. Machine Intell.*, 12:629–639, 1990.

[20] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.

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

[22] J. Shen. Combinatorics for wavelets: the umbral refinement equation. *Studies Applied Math.*, 103(2):121–147, 1999.

[23] J. Shen and G. Strang. Asymptotic analysis of Daubechies polynomials. *Proc. Amer. Math. Soc.*, 124:3819–3833, 1996.

[24] J. Shen and G. Strang. Asymptotics of Daubechies filters, scaling functions, and wavelets. *Appl. Comp. Harmon. Anal.*, 5(3):312–331, 1998.

[25] J. Shen and G. Strang. On wavelet fundamental solutions to the heat equation — heatlets. *J. Diff. Eq.*, 161(2):403-421, 2000.

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

[27] N. Wiener. *The Fourier Integral and Certain of Its Applications.* Dover Publications, New York, 1933.