Wavelets generated by using discrete singular convolution kernels



2000 J. Phys. A: Math. Gen. 33 8577

(http://iopscience.iop.org/0305-4470/33/47/317)

J. Phys. A: Math. Gen. **33** (2000) 8577–8596. Printed in the UK PII: S0305-4470(00)14938-8

# Wavelets generated by using discrete singular convolution kernels

G W Wei

Department of Computational Science, National University of Singapore, 117543, Singapore

E-mail: cscweigw@nus.edu.sg

Received 22 June 2000, in final form 12 September 2000

**Abstract.** This paper explores the connection between wavelet methods and an efficient computational algorithm—the discrete singular convolution (DSC). Many new DSC kernels are constructed and they are identified as wavelet scaling functions. Two approaches are proposed to generate wavelets from DSC kernels. Two well known examples, the Canny filter and the Mexican hat wavelet, are found to be special cases of the present DSC kernel-generated wavelets. A family of wavelet generators proposed in this paper are found to form an infinite-dimensional Lie group which has an invariant subgroup of translation and dilation. If DSC kernels form an orthogonal system, they are found to span a wavelet subspace in a multiresolution analysis.

## 1. Introduction

The widespread availability of less expensive high-performance computers has given tremendous impetus to the development of computational methodology in every field of science and engineering [1–19]. Most effort has been centred on developing either global methods or local methods for solving a variety of time-dependent and time-independent computational problems. The most commonly used local methods are finite differences, finite volumes, finite elements and boundary elements. The well known global methods are fast Fourier transform and spectral methods. Global methods are highly localized in spectral (momentum) space, but have very poor spatial localization. In contrast, local methods have high spatial localization, but are delocalized in spectral space. As a consequence, local methods can be easily adapted to complex geometries and boundary conditions, while the major advantage of global methods is their higher accuracy. In ordinary applications, it is relatively safe and efficient to use either a global method or a local one for numerically solving an ordinary differential equation (ODE) or a partial differential equation (PDE). However, when an ODE or PDE has singularities and/or homoclinic orbits, neither global methods nor local methods can be applied without numerical instabilities [20, 21]. The global methods lose their accuracy near the singularities due to Gibbs oscillations. The local methods have to be implemented in an adaptive manner, which greatly limits their accuracy and requires extremely small (spatial and/or time) mesh sizes. Obviously, it is desirable to have methods which combine the accuracy of global methods with the flexibility of local methods for practical applications.

Recently, a discrete singular convolution (DSC) algorithm was proposed as a potential approach for the computer realization of singular convolutions [22–25]. Sequences of approximations to the singular kernels of Hilbert type, Abel type and delta type were constructed. Applications to analytical signal processing, Radon transform and surface

0305-4470/00/478577+20$30.00  

interpolation were discussed. The mathematical foundation of this algorithm is the theory of distribution [26]. Numerical solutions to differential equations are formulated via the singular kernels of the delta type. By appropriately selecting the parameters of a DSC kernel, the DSC approach exhibits the accuracy of global methods for integration and the flexibility of local methods in handling complex geometries and boundary conditions. Recently, the unified features of the discrete singular convolution algorithm have been discussed [25]. It is demonstrated that different forms of implementation for the DSC algorithm, such as global, local, Galerkin, collocation and finite difference, can be deduced from a single starting point.

Many DSC kernels, such as (regularized) Shannon's delta kernel, (regularized) Dirichlet kernel, (regularized) Lagrange kernel and (regularized) de la Vallée Poussin kernel, have been constructed [22]. Practical applications were examined for the numerical solution of the Fokker–Planck equation [22] of statistical theory, and for the Schrödinger equation [24] of quantum mechanics. The DSC algorithm was also utilized in waveguide model analysis, electromagnetic wave propagation [27] and structural analysis [28–30]. Another development in the application of the DSC algorithm is its use in computing numerical solutions for the Navier–Stokes equation. The standard Taylor problem (two-dimensional incompressible Navier–Stokes equation with periodic boundary conditions) was solved to machine precision with a few grid points [28]. A DSC-successive over-relaxation algorithm and a DSC-finite subdomain method were developed for simulating the incompressible viscous flows [31] and for highly irregular geometries [32], respectively. In the context of image processing, DSC kernels were used to facilitate a new anisotropic diffusion operator for image restoration from noise [33]. More recently, the DSC algorithm was used to resolve a few numerically challenging problems. It was utilized to integrate the (nonlinear) sine-Gordon equation with the initial values close to a homoclinic manifold singularity [23], for which conventional local methods encounter great difficulties and result in numerically induced chaos [21]. Another difficult example that was resolved by using the DSC algorithm is the integration of the (nonlinear) Cahn–Hilliard equation in a circular domain, which is challenging because of the fourth-order artificial singularity at the origin and the complex phase space geometry [34].

The objective of this paper is to draw the connection between the DSC algorithm and the method of wavelets [35–45]. The latter has great impact in telecommunications and electronics engineering [46] and its applications are found in a variety of other science and engineering disciplines. Mathematically, wavelets are functions generated from a single function by dilation and translation. They form building blocks for some space, such as $L^2(R)$, either as a frame or as an orthonormal basis. Such building blocks are computationally important when they have certain regularity and localization in both time and frequency domains. Physically, the wavelet transform is a mathematical technique that can be used to split a signal into different frequency bands or components so that each component can be studied with a resolution matching its scale, thus providing excellent frequency and spatial resolution, and also achieving computational efficiency. It was recognized by Holschneider *et al* [40,41] that, if a real-valued, non-negative father wavelet, $\phi \in \mathcal{D}$, provides a smoothened version of a function $f$ over the real line $R$ by means of the convolution product

$$\Phi_{a,b}(f) = \langle \phi_a(b), f \rangle = \int_{-\infty}^{\infty} \mathrm{d}t \, \frac{1}{a} \phi \left( \frac{b-t}{a} \right) f(t) = \phi_a(b) * f \tag{1}$$

then a wavelet transform can be given by

$$W_{a,b}(f) = -a\partial_a \int_{-\infty}^{\infty} \mathrm{d}t \, \frac{1}{a} \phi \left( \frac{b-t}{a} \right) f(t) = \int_{-\infty}^{\infty} \mathrm{d}t \, \psi \left( \frac{b-t}{a} \right) f(t) = \psi_a(b) * f \tag{2}$$

where $\psi_a(x) = \psi(x/a)/a$ and $\psi(x)$ is a wavelet given by

$$\psi(x) = G^1\phi(x) = (x\partial_x + 1)\phi(x) \tag{3}$$

where $G^1 = (x\partial_x + 1)$ is a wavelet generator and it creates the Mexican hat wavelet for $\phi(x) = \frac{1}{\sqrt{\pi}}e^{-x^2}$. The wavelet transform, equation (2), provides a mathematical microscopy of $f$ at length scale $a$. Among their applications, wavelets have been used most widely for data compression in which a low-frequency component is encoded with fewer mesh points by a down-sampling scheme without much loss of information. Signal processing methods such as quadrature mirror filters go hand in hand with wavelet techniques for studying a host of communication problems. Essentially, as discovered by Mallat and Meyer [35, 37], wavelet multiresolution theory has a faithful representation in terms of subband filters. Due to their multiresolution feature, wavelets are a useful tool for analysing fractals and associated dynamical processes. Wavelet packets are found to be useful for local characterization of classical turbulence and for pattern recognition. There have been applications of wavelet theory in statistical mechanics in terms of phase space representation and renormalization group. A general group theory description of wavelet transform has also been studied [39, 41, 47]. However, there are still some areas where the existing wavelet methods encounter enormous difficulties [48]. In particular, computational fluid dynamics, and computational physics and mechanics in general, come under this category. Since wavelets are intimately and significantly related to spline theory, the theory of approximation and basis of minimum support, they should have great potential to lead to entirely new approaches for scientific computation. This paper extends the wavelet generator $G^1$ to a family of new generators $G^n$ ($n = 0, 1, 2, \ldots$). The role of the real-valued, non-negative father wavelet, $\phi \in \mathcal{D}$, is replaced by DSC delta kernels. It is hoped that the connection made in this paper between the DSC algorithm and wavelet methods might be useful for the development of efficient and robust wavelet algorithms for solving differential equations. We also hope that the findings of this paper will promote the use of the DSC as a powerful algorithm for image processing and pattern recognition, for which wavelet theory has found great success.

This paper is organized as following: in section 2, we briefly discuss the DSC algorithm and a number of delta sequence kernels used in the realization of the DSC algorithm. In section 3, DSC delta sequence kernels are identified as wavelet scaling functions. The role of orthogonal DSC kernels in a multiresolution framework is studied. Two general approaches are given for transforming a DSC kernel into a wavelet. The first transformation is new and is shown to form an infinite-dimensional Lie group which includes the group of translation and dilation as an invariant subgroup. A new set of Shannon's wavelets is obtained by this transformation. The other approach is the standard difference method, in which a wavelet is constructed by the difference of two unequally parametrized scaling functions. This paper ends with a discussion.

## 2. Discrete singular convolution

Singular convolutions (SC) are a special class of mathematical transformations, which appear in many science and engineering problems, such as the Hilbert, Abel and Radon transforms. It is most convenient to discuss singular convolutions in the context of the theory of distribution. The latter has important ramifications in mathematical analysis. Not only does it provide a rigorous justification for a number of informal manipulations in physical and engineering science, but it also opens up a new area of mathematics, which in turn gives impetus to many other mathematical disciplines, such as operator calculus, differential equations, functional analysis, harmonic analysis and transformation theory. In fact, the theory of wavelets and

frames, a new mathematical branch developed in recent years, can also find its root in the theory of distributions.

Let $T$ be a distribution and $\eta(t)$ be an element of the space of test functions. A singular convolution is defined as

$$F(t) = (T * \eta)(t) = \int_{-\infty}^{\infty} T(t-x)\eta(x)\,\mathrm{d}x. \tag{4}$$

Here $T(t-x)$ is a singular kernel. Depending on the form of the kernel $T$, the singular convolution is the central issue for a wide range of science and engineering problems. For example, singular kernels of the Hilbert type have a general form of

$$T(x) = \frac{1}{x^n} \qquad (n = 1, 2, \ldots). \tag{5}$$

Here, kernel $T(x) = \frac{1}{x}$ occurs commonly in electrodynamics, theory of linear response, signal processing, theory of analytic functions and the Hilbert transform; $T(x) = \frac{1}{x^2}$ is the kernel used in tomography. Another interesting example is the singular kernels of the Abel type

$$T(x) = \frac{1}{x^\beta} \qquad (0 < \beta < 1). \tag{6}$$

These kernels can be recognized as special cases of the singular integral equations of Volterra type of the first kind. Singular kernels of the Abel type have applications in the area of holography and interferometry with phase objects (of practical importance in aerodynamics, heat and mass transfer, and plasma diagnostics). They are intimately connected with the Radon transform, for example, in determining the refractive index from the knowledge of a holographic interferogram. The other important example is the singular kernel of the delta type

$$T(x) = \delta^{(n)}(x) \qquad (n = 0, 1, 2, \ldots). \tag{7}$$

Here, kernel $T(x) = \delta(x)$ is of particular importance for interpolation of surfaces and curves (including atomic, molecular and biological potential energy surfaces, and for a variety of image processing and pattern recognition problems) and $T(x) = \delta^{(n)}(x)$, $(n = 1, 2, \ldots)$ are essential for numerically solving partial differential equations. However, since these kernels are singular, they cannot be directly digitized in a computer. Hence, the singular convolution, equation (4), is of little numerical merit. To avoid the difficulty of using singular expressions directly in a computer, sequences of approximations $(T_\alpha)$ of the distribution $T$ can be constructed:

$$\lim_{\alpha \to \alpha_0} T_\alpha(x) \longrightarrow T(x) \tag{8}$$

where $\alpha_0$ is a generalized limit. Obviously, in the case of $T(x) = \delta(x)$, the sequence, $T_\alpha(x)$, is a delta sequence kernel. Note that one retains the delta distribution at the limit of a delta sequence kernel, which is a real constant in the frequency space and is the so-called *all pass filter*. Computationally, the delta distribution is a *universal reproducing kernel* which can be used as a starting point for the construction of either band-limited reproducing kernels or approximate reproducing kernels. Furthermore, with a sufficiently smooth approximation, it is meaningful to consider a *discrete singular convolution* (DSC)

$$F_\alpha(t) = \sum_k T_\alpha(t - x_k) f(x_k) \tag{9}$$

where $F_\alpha(t)$ is an approximation to $F(t)$ and $\{x_k\}$ is an appropriate set of discrete points on which the DSC (9) is well defined. Note that the original test function $\eta(x)$ has been replaced by $f(x)$. The mathematical properties or requirements of $f(x)$ are determined by the approximate kernel $T_\alpha$. In general, the convolution is required to be Lebesgue integrable.

The most commonly used DSC kernels are the delta sequence kernels, which are approximations to the delta distribution or so-called Dirac delta function ($\delta$). Delta distribution is a generalized function which is integrable inside a particular interval but in itself does not need to have a value. The theory of distribution has been studied by Schwartz [26], Korevaar [49] and others. The use of delta sequences as probability density estimators was discussed by Walter and Blum [50] and others [51–53].

**Definition 1.** *Let $\{\delta_\alpha\}$ be a sequence of kernel functions on $(-\infty, \infty)$ which are integrable over every bounded interval. We call $\{\delta_\alpha\}$ a delta sequence kernel of positive type if:*

*(1) $\int_{-a}^{a} \delta_\alpha \to 1$ as $\alpha \to \alpha_0$ for some finite constant $a$.*
*(2) For every constant $\gamma > 0$, $(\int_{-\infty}^{-\gamma} + \int_{\gamma}^{\infty})\delta_\alpha \to 0$ as $\alpha \to \alpha_0$.*
*(3) $\delta_\alpha(x) \geqslant 0$ for all $x$ and $\alpha$.*

Useful examples of delta sequence kernels of the positive type include the following:

*The impulse functions*

$$\delta_\alpha(x) = \begin{cases} \alpha & \text{for} \quad 0 < x < 1/\alpha \qquad \alpha = 1, 2, \ldots \\ 0 & \text{otherwise.} \end{cases} \tag{10}$$

*Gauss's delta sequence kernel*

$$\delta_\alpha(x) = \frac{1}{\sqrt{2\pi}\,\alpha} \mathrm{e}^{-x^2/2\alpha^2}. \tag{11}$$

*Lorentz's delta sequence kernel*

$$\delta_\alpha(x) = \frac{1}{\pi} \frac{\alpha}{x^2 + \alpha^2} \tag{12}$$

and its generalized form

$$\delta_{\alpha,n}(x) = \frac{1}{\pi} \frac{\alpha^n x^{n-1}}{x^{2n} + \alpha^{2n}} \qquad \text{for} \quad n \geqslant 1. \tag{13}$$

*Landau's delta sequence kernel*

$$\delta_n(x) = \begin{cases} L_n(x) & \text{for} \quad |x| \leqslant a \\ 0 & \text{otherwise} \end{cases} \tag{14}$$

where

$$L_n(x) = \frac{(a^2 - x^2)^n}{\int_{-a}^{a} (a^2 - y^2)^n \,\mathrm{d}y} \qquad \text{for} \quad n = 0, 1, 2, \ldots \qquad \text{and} \qquad a > 0. \tag{15}$$

*Poisson's delta sequence kernel family*

$$\delta_\alpha(x) = \begin{cases} P_\alpha(x) & \text{for} \quad |x| \leqslant \pi \\ 0 & \text{otherwise} \end{cases} \tag{16}$$

where

$$\begin{aligned} P_\alpha(x) &= \frac{1}{\pi}\left[\frac{1}{2} + \alpha\cos(x) + \alpha^2\cos(2x) + \cdots\right] \\ &= \frac{1 - \alpha^2}{2\pi(1 - 2\alpha\cos(x) + \alpha^2)} \qquad (0 \leqslant \alpha < 1). \end{aligned} \tag{17}$$

*Fejér's delta sequence kernel*

$$\delta_\alpha(x) = \begin{cases} F_\alpha(x) & \text{for} \quad |x| \leqslant \pi \qquad \text{for} \quad \alpha = 0, 1, 2, \ldots \\ \\ 0 & \text{otherwise} \end{cases} \tag{18}$$

where

$$\begin{aligned} F_k(x) &= \frac{1}{k}[D_0(x) + D_1(x) + \cdots + D_{k-1}(x)] \\ &= \frac{\sin^2(\frac{1}{2}kx)}{2\pi k \sin^2(\frac{1}{2}x)} \qquad -\infty < x < \infty \end{aligned} \tag{19}$$

and $D_k$ are given by the partial sum of the discrete Fourier series:

$$\begin{aligned} D_k(x) &= \frac{1}{\pi}\left[\frac{1}{2} + \cos(x) + \cos(2x) + \cdots + \cos(kx)\right] \\ &= \frac{\sin[(k+\frac{1}{2})x]}{2\pi \sin\frac{1}{2}x} \qquad k = 0, 1, 2, \ldots. \end{aligned} \tag{20}$$

*Generalized Fejér's delta sequence kernel*

$$\delta_\alpha(x) = \frac{2}{\pi}\frac{\sin^2(\alpha x)}{\alpha x^2} \qquad \forall x \in R. \tag{21}$$

*Delta sequence kernels generated by dilation*

Let $\rho \in L^1(R)$ be a non-negative function with $\int \rho(x)\,\mathrm{d}x = 1$, then dilation of $\rho$ given by

$$\rho_\alpha(x) = \frac{1}{\alpha}\rho\left(\frac{x}{\alpha}\right) \qquad (\alpha > 0) \tag{22}$$

leads to a delta sequence kernel, $\rho_\alpha \to \delta$, as $\alpha \to 0$.

A common feature of the above-mentioned delta sequence kernels is that they are all non-negative functions. In fact, what are used in most of our numerical computations are the delta sequence kernels of Dirichlet type.

**Definition 2.** *Let $\{\delta_\alpha\}$ be a sequence of functions on $(-\infty, \infty)$ which are integrable over every bounded interval. We call $\{\delta_\alpha\}$ a delta sequence kernel of the Dirichlet type if:*

*(1) $\int_{-a}^{a} \delta_\alpha \to 1$ as $\alpha \to \alpha_0$ for some finite constant $a$.*
*(2) For every constant $\gamma > 0$, $(\int_{-\infty}^{-\gamma} + \int_{\gamma}^{\infty})\delta_\alpha \to 0$ as $\alpha \to \alpha_0$.*
*(3) There are positive constants $C_1$ and $C_2$ such that*

$$|\delta_\alpha(x)| \leqslant \frac{C_1}{|x|} + C_2$$

*for all $x$ and $\alpha$.*

We list a few useful examples of delta sequence kernels of the Dirichlet type:
*Dirichlet kernel*

$$\delta_\alpha(x) = \begin{cases} D_\alpha(x) & \text{for} \quad |x| \leqslant \pi \qquad \text{for} \quad \alpha = 0, 1, 2, \ldots \\ 0 & \text{otherwise} \end{cases} \tag{23}$$

where $D_\alpha$ is the Dirichlet kernel given by equation (20).
*Modified Dirichlet kernel*

$$\delta_\alpha(x) = \begin{cases} D^*_\alpha(x) & \text{for} \quad |x| \leqslant \pi \qquad \text{for} \quad \alpha = 0, 1, 2, \ldots \\ 0 & \text{otherwise} \end{cases} \tag{24}$$

where

$$D^*_\alpha(x) = D_\alpha - \frac{1}{2\pi}\cos(\alpha x)$$
$$= \frac{\sin(\alpha x)}{2\pi \tan(\frac{1}{2}x)} \qquad \alpha = 0, 1, 2, \ldots. \tag{25}$$

*Lagrange kernel*

$$\delta_{M,k}(x) = \begin{cases} L_{M,k}(x) & \text{for} \quad a \leqslant x \leqslant b \qquad \text{for} \quad M = 1, 2, \ldots \\ 0 & \text{otherwise} \end{cases} \tag{26}$$

where the Lagrange interpolation formula

$$L_{M,k}(x) = \prod_{i=k-M, i\neq k}^{i=k+M} \frac{x - x_i}{x_k - x_i} \qquad (M \geqslant 1) \tag{27}$$

is defined on an interval $(a, b)$ with a set of $2M + 1$ ordered discrete points:

$$\{x_i\}_{i=k-M}^{k+M}: \qquad x_{k-M} = a < x_{k-M+1} < \cdots < x_k < \cdots < x_{k+M} = b. \tag{28}$$

*The de la Vallée Poussin delta sequence kernel*

$$\delta_{n,p}(x) = \begin{cases} P_{n,p}(x) & \text{for} \quad |x| \leqslant \pi \qquad \text{for} \quad p = 0, \ldots, n \qquad n = 0, 1, \ldots \\ 0 & \text{otherwise} \end{cases} \tag{29}$$

where

$$P_{n,p}(x) = \frac{1}{p+1}\sum_{k=n-p}^{n} D_k(x)$$
$$= \frac{1}{2\pi} + \frac{1}{\pi}\sum_{k=1}^{n-p}\cos kx + \frac{1}{\pi}\sum_{k=1}^{p}\left[1 - \frac{k}{p+1}\right]\cos[(n-p+k)x]$$
$$= \frac{\sin[(2n+1-p)\frac{x}{2}]\sin[(p+1)\frac{x}{2}]}{2\pi(p+1)\sin^2(\frac{x}{2})} \qquad p = 0, \ldots, n \qquad n = 0, 1, \ldots. \tag{30}$$

Note that the de la Vallée Poussin kernel reduces to the positively defined Ferér's kernel $F_{n+1}(x)$ when $p = n$. A simplified de la Vallée Poussin kernel given by

$$\delta_\alpha(x) = \frac{1}{\pi\alpha}\frac{\cos(\alpha x) - \cos(2\alpha x)}{x^2} \tag{31}$$

is found to be very useful numerically [22].

*Shannon's delta sequence*

$$\delta_\alpha(x) = \frac{\sin(\alpha x)}{\pi x}. \tag{32}$$

*DSC kernels constructed by orthogonal basis expansions*

Let $\{\psi_i\}$ be a complete orthonormal $L^2(a, b)$ basis, then

$$\delta_n(x, y) = \sum_{i=0}^{n}\psi_i(x)\psi_i(y) \qquad x, y \in (a, b) \tag{33}$$

are DSC delta sequence kernels. In the case of trigonometric functions, we again obtain the Dirichlet kernels. Hermite function expansion, evaluated at $x = 0$, is given by

$$\delta_n(x) = \exp(-x^2)\sum_{k=0}^{n}\left(\frac{-1}{4}\right)^k \frac{1}{\sqrt{\pi}k!}H_{2k}(x) \qquad \forall x \in R \tag{34}$$

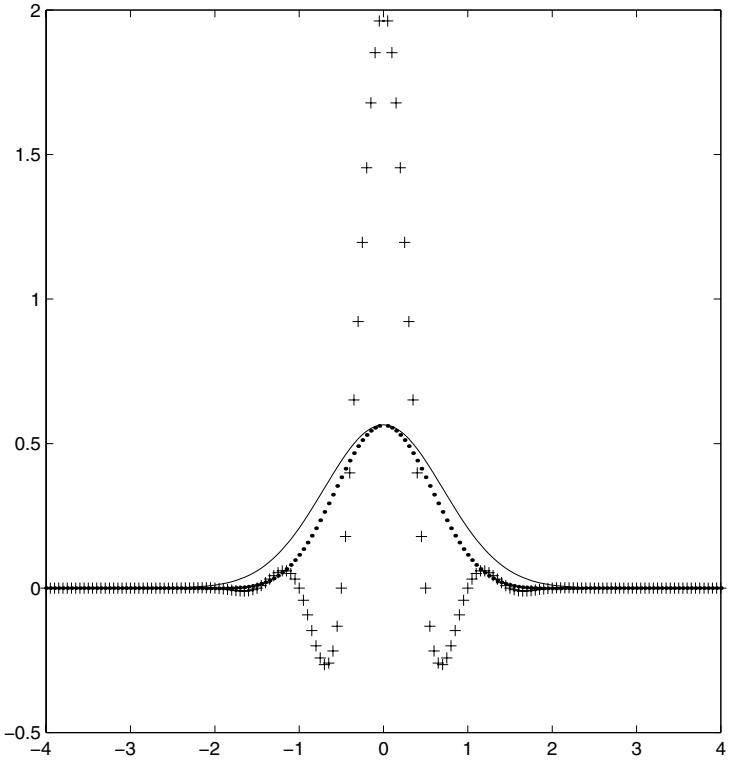


**Figure 1.** A comparison of delta sequence kernels of the positive type and Dirichlet type. Full curve: $\frac{1}{\sqrt{\pi}}e^{-x^2}$; dots: $\frac{1}{\pi}\frac{\sin(x)}{x}e^{-x^2}$; crosses: $\frac{1}{\pi}\frac{\sin(2\pi x)}{x}e^{-x^2}$.

where $H_{2k}(x)$ is the usual Hermite polynomial.

The regularity of a kernel approximation can be improved by using a regularized form

$$\delta_\alpha(x)R_\sigma(x) \tag{35}$$

where $R_\sigma(x)$ is a regularizer:

$$\lim_{\sigma\to\infty} R_\sigma(x) = 1 \tag{36}$$

$$R_\sigma(0) = 1. \tag{37}$$

A commonly used delta-kernel regularizer is the Gaussian, $e^{-x^2/2\sigma^2}$, which is a Schwartz class function. Except for the delta sequence kernels generated by the Hermite functions, all of the above-mentioned kernels of the Dirichlet type are found to perform better in many practical applications in their regularized form as prescribed by equation (35) [22, 25]. A comparison of delta sequence kernels of positive type and Dirichlet type is given in figure 1.

## 3. Wavelets generated by using DSC kernels

Wavelets have been widely used as an analysis tool for various applications. The essential reason for this lies in the fact that both orthogonal and nonorthogonal wavelets can provide decomposition of a function at a variety of different scales. In other words, wavelets form special $L^2(R)$ bases or frames for representing a function at various levels of detail, leading to

so-called mathematical microscopes. This turns out to be very efficient for approximating and analysing functions in many applications. Orthogonal wavelets and multiresolution analysis have been successfully used in a variety of engineering fields. They play a special role in those applications where orthogonality is strongly required. In many other applications, nonorthogonal wavelets, or frames, are also very useful. Generating new types of wavelets has been of great importance in wavelet theory. Most of the DSC delta sequence kernels described in section 2 can be regarded as scaling functions (father wavelets). Physically, they are low-pass filters. In particular, Shannon's scaling function is an ideal low-pass filter in the frequency space with a finite bandwidth, which allows a perfect representation for a band-limited function by a discrete but infinite set of samplings. Hence, these scaling functions can also be systematically transformed into mother wavelets which are high-pass filters from the physics point of view. For an orthogonal system, the combination of a low-pass filter and all high-pass filters will give rise to a wavelet basis representation for the $L^2(R)$. The orthogonal wavelets are briefly reviewed in the first subsection. The connection between DSC delta sequence kernels and wavelets is made in the second subsection. The construction of mother wavelets from various DSC delta sequence kernels is discussed in the last three subsections.

### *3.1. Orthogonal wavelets*

The formal theory of orthogonal wavelets on $L^2(R)$ has been presented in many books [35, 36, 38]. An orthogonal wavelet system is usually generated from a single function, either a scaling function (father wavelet) $\phi$ or a mother wavelet $\psi$, by standard translation and dilation techniques:

$$\phi_{mn}(x) = 2^{-\frac{m}{2}}\phi\left(\frac{x}{2^m} - n\right) \qquad m, n \in Z \tag{38}$$

$$\psi_{mn}(x) = 2^{-\frac{m}{2}}\psi\left(\frac{x}{2^m} - n\right) \qquad m, n \in Z \tag{39}$$

where the symbol $Z$ denotes the set of all integers. This can be phrased rigorously in terms of a multiresolution analysis, i.e. a nested sequence of subspaces $\{V_m\}$, $m \in Z$ such that:

(1) $\{\phi(x-n)\}$ is an orthogonal basis of $V_0$;
(2) $\cdots \subset V_1 \subset V_0 \subset V_{-1} \subset \cdots \subset L^2(R)$;
(3) $f(x) \in V_m \leftrightarrow f(2x) \in V_{m-1}$;
(4) $\cap_m V_m = \{0\}$ and $\overline{\cup_m V_m} = L^2(R)$.

Since $\phi \in V_0 \subset V_{-1}$, it can be expressed as a superposition of $\{\phi_{-1,n}\}$, which spans an orthogonal basis for $V_{-1}$

$$\phi(x) = \sum_n a_n \phi_{-1,n} \tag{40}$$

where $\{a_n\}$ is a set of finite coefficients.

For an orthogonal system, the subspace $V_{-1}$ can be further split into its orthogonal projection on $V_0$ and a complementary $W_0$:

$$V_{-1} = V_0 \oplus W_0 \tag{41}$$

where $W_0$ is a subspace spanned by orthogonal mother wavelets $\{\psi\}$. In general, $\psi_{mn}$, $n \in Z$ is an orthogonal basis of $W_{-m}$ and

$$\bigoplus_{m\in Z} W_m = L^2(R). \tag{42}$$

It follows that $\psi_{mn}$ $(m, n \in Z)$ is an orthogonal basis of $L^2(R)$. Similar to equation (40), the mother wavelet can also be expressed as a superposition of $\{\phi_{-1,n}\}$

$$\psi(x) = \sum_n b_n \phi_{-1,n} \tag{43}$$

where $b_n$ are expansion coefficients.

Perhaps the simplest example is the Haar's wavelet system [38] which is given by $\phi(x) = \chi_{[0,1)}(x)$, the characteristic function of interval [0, 1). It obviously has orthogonal translation. The dilation of $\phi(x)$ results in characteristic functions for a smaller (or larger) interval and each of them spans a subspace $V_m$ by translations. Haar's wavelets play a unique role in the wavelet theory and application, by their simplicity.

It is not obvious whether a multiresolution analysis exists for $\phi$ other than the Haar system. The construction of the first few orthogonal wavelet bases is more or less an art rather than a procedure; it requires ingenuity, special tricks and subtle computations. One procedure used by Meyer [35] is to start with a spline function $\theta(x) = (1 - |x - 1|)\chi_{[0,2)}$ which, by translations, spans a nonorthogonal Riesz basis (a frame of the least redundancy possible). The corresponding orthogonal basis $\{\phi\}$ is represented in Fourier space as $\hat{\phi}(\xi) = \alpha(\xi)\hat{\theta}(\xi)$, where $\alpha(\xi)$ is the function to be determined. Using both the orthonormality requirement:

$$\delta_{0,n} = \langle \phi_{0,0}, \phi_{0,n} \rangle \tag{44}$$

and the periodicity, $\hat{\phi}$, can be resolved as

$$\hat{\phi}(\xi) = \frac{\sin^2(\xi/2)}{(\xi/2)^2} \left(1 - \frac{2}{3}\sin^2\frac{\xi}{2}\right)^{-\frac{1}{2}}. \tag{45}$$

The space $V_0$ is spanned by piecewise linear functions.

Daubechies presented another approach for constructing orthogonal wavelets [36]. In the Fourier representation, the dilation equation can be written as

$$\frac{\hat{\phi}(\xi)}{\hat{\phi}(\xi/2)} = m_0(\xi/2) \tag{46}$$

where $m_0$ is a $2\pi$-periodic function. The orthonormality condition then requires

$$|m_0(\xi/2)|^2 + |m_0(\xi/2 + \pi)|^2 = 1. \tag{47}$$

It turns out that, if the set of expansion coefficients $a_n$ of equation (40) is chosen as $a_0 = \nu(\nu-1)/(\nu+1)\sqrt{2}, a_1 = -(\nu-1)/(\nu+1)\sqrt{2}, a_2 = (\nu-1)/(\nu+1)\sqrt{2}, a_3 = \nu(\nu+1)/(\nu+1)\sqrt{2}$, $(\nu \in R)$, then equation (47) will be satisfied and consequently $\phi$ can be found recursively.

### *3.2. DSC kernels as scaling functions*

Let $\{\delta_\alpha\}$ (where $\alpha \to \alpha_0$) be a sequence of $C^m$ functions on $(-\infty, \infty)$ which are integrable over every bounded interval, and

(1) $\hat{\delta}_\alpha(0) = 1$ for each $\alpha$;
(2) $\lim_{\alpha\to\alpha_0} \hat{\delta}_\alpha(\xi) \to 1$ for all $\xi$;
(3) for every constant $\gamma > 0$, $(\int_{-\infty}^{-\gamma} + \int_{\gamma}^{\infty})\delta_\alpha \to 0$ as $\alpha \to \alpha_0$;
(4) and $\| x\delta_\alpha(x) \|_\infty < \infty$ for all $x$ and $\alpha$.

Then $\{\delta_\alpha\}$ refers to the DSC delta sequence kernels (or DSC kernels in short) and each function can be admitted as a scaling function, $\int \delta_\alpha = 1$, for all $\alpha \neq \alpha_0$ and the integration stands even for $\alpha = \alpha_0$ from the point of view of distributions. It is noted that most DSC kernels given in the previous section satisfy these conditions. We call this class of scaling functions 'DSC

kernel-generated scaling functions $\phi_\alpha$'. The corresponding 'DSC kernel-generated mother wavelets $\psi_\alpha$' have the Fourier transform property

$$\hat{\psi}_\alpha(0) = \int \psi_\alpha(x)\,\mathrm{d}x = 0. \tag{48}$$

It is natural to view DSC delta sequence kernels as scaling functions. In particular, if a DSC delta sequence kernel has the structure that $\delta_\alpha = \frac{1}{\alpha}\rho(\frac{x}{\alpha})$ and $\int \rho(x)\,\mathrm{d}x = 1$ (this is the case for many examples given in section 2), then $\delta_\alpha$ is a sequence of scaling functions at different scales. In contrast to the delta distribution which only has a point support, a function in a DSC delta sequence kernel can have an arbitrary support, depending on the scale. In the limit $\alpha \to \alpha_0$, the DSC delta sequence kernel converges to the delta distribution and the support shrinks down to a point. The resulting delta distribution actually helps to furnish the wavelet multiresolution analysis

$$\{\delta\} \oplus \bigoplus_{m\in Z} W_m = L^2(R) \tag{49}$$

where $\{\delta\}$ is the space spanned by the delta distribution, which, however, has only a point support. Clearly, if a DSC delta sequence kernel is an orthogonal system, such as Dirichlet's continuous delta sequence kernel, for a fixed $\alpha \neq \alpha_0$, we have

$$\{\delta_\alpha\} \oplus \bigoplus_{-\infty}^{m=0} W_m = \{\delta\} \oplus \bigoplus_{m\in Z} W_m = L^2(R) \tag{50}$$

where $\{\delta_\alpha\}$ spans the wavelet subspace $V_0$. Hence the orthogonal DSC delta sequence kernel spans the wavelet subspace $\{\delta\} \oplus \bigoplus_1^\infty W_m$ for an appropriate choice of $\alpha$.

DSC kernel-generated mother wavelets can be constructed in many different ways. We discuss two of them in the following subsections.

### *3.3. Wavelets generated by differentiation pairs*

For a given $C^m$ symmetric scaling function $\phi$, we define a family of wavelet generators

$$G^n = x\frac{\partial^n}{\partial x^n} + n\frac{\partial^{n-1}}{\partial x^{n-1}} \qquad n = 0, 1, 2, \ldots, m \tag{51}$$

for generating a family of $m + 1$ mother wavelets

$$\psi_{\alpha,n}(x) = G^n\phi_\alpha(x) \qquad \text{for} \quad \phi \in C^m \quad \text{and} \quad n = 0, 1, 2, \ldots, m. \tag{52}$$

Obviously, the wavelet generator given by Holschneider *et al* [40, 41] in equation (3) is a special case of $G^n$, corresponding to $G^1$. It is noted that this approach is not restricted to the DSC kernel-generated wavelet and is actually a very general and efficient way for creating wavelets from a given $C^m$ symmetric scaling function.

The present wavelet generators are closely related to the transformation Lie group of translation and dilation. This is because the Fourier image of distributions

$$G^n = x\frac{\partial^n}{\partial x^n} + n\frac{\partial^{n-1}}{\partial x^{n-1}} \qquad n = 0, 1, 2, \ldots \tag{53}$$

forms an infinite-dimensional wavelet Lie algebra with elements $\{X_n = \xi^{n-1}\partial_\xi | n = 0, 1, 2, \ldots\}$ (here we follow the convention that statements concerning the structure of a Lie algebra are made only on the basis of the real Lie algebra). The whole Lie algebra structure of $\{X_n\}$ is simply given by

$$[X_n, X_m] = (m - n)X_{m+n-2} \qquad n, m = 0, 1, 2, \ldots. \tag{54}$$

$X_1$ by itself generates a one-parameter noncompact Abelian group which is obviously the translation group in momentum space. $X_2$ is one of the generators for a dilation group. There are two nontrivial invariant subalgebras:

$$[X_1, X_2] = X_1 \tag{55}$$

and

$$[X_1, X_2] = X_1 \qquad [X_1, X_3] = 2X_2 \qquad [X_2, X_3] = X_3. \tag{56}$$

$X_1, X_2$ are the infinitesimal generators of a two-dimensional translation and dilation group. The third element $X_3$ is a quadratic dilation (superdilation), which allows us to generate another invariant subalgebra, equation (56). This result indicates that the present method of systematically generating wavelets is a very general one. It is noted that the present wavelet algebra is closely related to the mathematical structure (Virasoro algebra) in the superstring theory. More details about the general connection between the present wavelet generators and quantum field theory will be discussed elsewhere.

It is very easy to construct various DSC kernel-generated wavelets by applying examples given in section 2 to the right-hand side of equation (52). A few examples are given below.

**Example 1 (Mexican hat wavelet and generalizations).** If we take Gauss's delta sequence kernel as a scaling function, $\phi_\alpha(x) = \frac{1}{\sqrt{2\pi}\alpha} e^{-x^2/2\alpha^2}$, then the expression given by $n = 1$

$$\psi_{\alpha,1}(x) = \frac{1}{\sqrt{2\pi}\alpha}\left(1 - \frac{x^2}{\alpha^2}\right) e^{-x^2/2\alpha^2} \tag{57}$$

is the well known Mexican hat wavelet [36]. Taking $n = 3$ yields

$$\psi_{\alpha,3}(x) = \frac{-1}{\sqrt{2\pi}\alpha}\left(\frac{x^4}{\alpha^6} - \frac{6x^2}{\alpha^4} + \frac{3}{\alpha^2}\right) e^{-x^2/2\alpha^2} \tag{58}$$

which is an interesting 'Mexican superhat wavelet'. This wavelet is expected to perform better than the Mexican hat for some applications. Since Gauss's kernels are $C^\infty$ functions, there are infinitely many Gauss kernel-generated wavelets given by

$$\psi_n(x) = G^n \frac{1}{\sqrt{\pi}} e^{-x^2} = \frac{1}{\sqrt{\pi}} \frac{(-1)^n}{2} H_{n+1}(x) e^{-x^2} \qquad n = 0, 1, 2, \ldots. \tag{59}$$

It is seen that the celebrated Mexican hat wavelet [36] is just a special case of equation (59). From the quantum mechanics point of view, the operator $G^n$ creates an $n + 1$ excited mode from the ground state $H_0$. This is true only for the system described by the Hermite functions. In general, the operator $G^n$ creates different patterns for different systems.

The first few examples of Gauss kernel-generated wavelets are plotted in figure 2 (full curves). It is seen that $\psi_0(x)$ looks like the Haar wavelet and is very useful in image processing and edge detection (known as the Canny operator). The Mexican hat wavelet, $\psi_1(x)$, has two nontrivial zeros and, as a wavelet filter, it should have better frequency response to higher frequency components. In general, Gauss kernel-generated wavelets involve the Hermite polynomials of degree $n + 1$ and have $n + 1$ nontrivial zeros. For applications in image processing, particularly edge detection, the first few low-order Gauss kernel-generated wavelets are sufficient because most of the energy of an image edge is concentrated in the low-frequency region. However, for accurate scientific computations, such as solving a partial differential equation, higher-order Gauss kernel-generated wavelets are required.

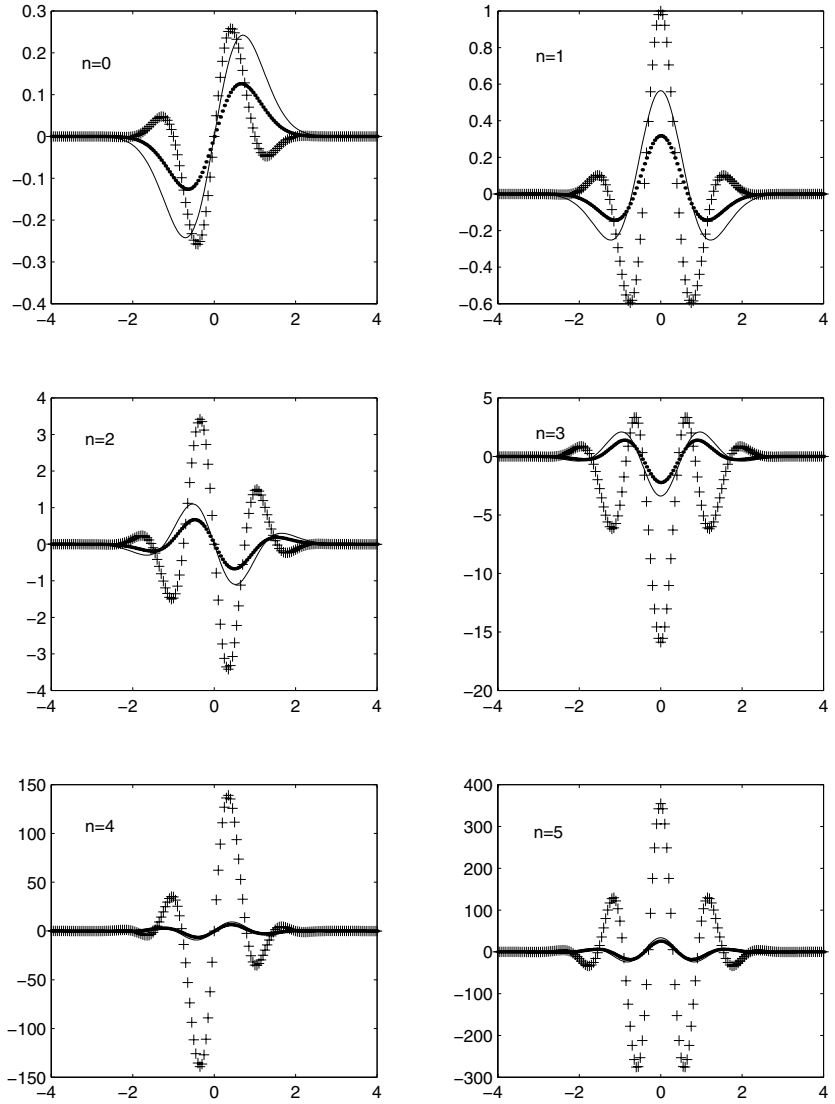


**Figure 2.** A comparison of wavelets generated by $G^n$. Full curve: Gauss kernel-generated wavelets $G^n \frac{1}{\sqrt{\pi}} e^{-x^2}$; dots: regularized Shannon's wavelets $G^n \frac{1}{\pi} \frac{\sin(x)}{x} e^{-x^2}$; crosses: $G^n \frac{1}{\pi} \frac{\sin(\pi x)}{x} e^{-x^2}$.

It is interesting to note that all higher-order Hermite functions ($n \neq 0$), together with the Gaussian weight, are mother wavelets, while the lowest-order Hermite function is a scaling function. This can be naturally seen from the orthonormality condition

$$\int \frac{1}{\sqrt{\pi 2^n n!}} H_n(x) H_m(x) e^{-x^2} \, dx = \delta_{nm} \qquad n, m = 0, 1, 2, \ldots. \tag{60}$$

Here, if the second polynomial is fixed as a constant when $m = 0$, then the only case for $H_n$ to give a nonzero integration is $n = 0$, which determines a scaling function. All other $H_n$ ($n \neq 0$) give rise to mother wavelets. (Note that, when $m = 0$, the integration becomes $L^1$ and is related to Fourier transform at the origin.) We point out that this result is *not* limited

to Hermite polynomials. It is generally true that *all* polynomial systems form an orthogonal $L^2(a, b)$ with respect to appropriate weights. A systematic study of the general connection between wavelets and conventional Hilbert space bases is discussed elsewhere.

**Example 2 (Hermite's kernel generated wavelets).** In the case of Hermite's kernels, equation (34), we have

$$\begin{aligned}\Psi_H(x|m,n) &= \left(x\frac{\partial^m}{\partial x^m} + m\frac{\partial^{m-1}}{\partial x^{m-1}}\right)\delta_n(x) \\ &= \exp(-x^2)\sum_{k=0}^{n}(-1)^{k+m}\frac{1}{\sqrt{\pi}2^{2k+1}k!}[H_{2k+m+1}(x) + 4kH_{2k+m-1}(x)] \end{aligned} \tag{61}$$

where $m, n = 0, 1, 2, \ldots$. Here, some simple properties of Hermite functions have been used for simplifying the results. If one chooses the scaling as $\alpha = \frac{1}{\sqrt{2}}$, Mexican hat wavelet, $\frac{1}{\sqrt{\pi}}(1 - 2x^2)\mathrm{e}^{-x^2}$, and the Mexican superhat wavelet, $\frac{-2}{\sqrt{\pi}}(4x^4 - 12x^2 + 3)\mathrm{e}^{-x^2}$, are given by $\Psi_H(x|1, 0)$ and $\Psi_H(x|3, 0)$, respectively. In general, the whole series of Gauss kernel-generated wavelets, equation (59), are given as special cases, by $\Psi_H(x|m, 0)$, $m = 0, 1, 2, \ldots$.

**Example 3 (Shannon's wavelet family).** Dirichlet's continuous kernel is related to the well known Shannon's scaling function [36] or the sinc function, $\phi_\alpha(x) = \frac{1}{\pi}\frac{\sin(\alpha x)}{x}$. The latter is known for generating an orthogonal basis for a reproducing kernel Hilbert space. The corresponding reproducing kernels give rise to a sampling basis for certain band-limited $L^2$ functions. A family of (mother) wavelets can be generated by using the present wavelet generators, equation (51),

$$\begin{aligned}\psi_{\alpha,n}(x) &= \left(x\frac{\partial^n}{\partial x^n} + n\frac{\partial^{n-1}}{\partial x^{n-1}}\right)\frac{1}{\pi}\frac{\sin(\alpha x)}{x} \\ &= \frac{1}{\pi}\sin(\alpha x) && \text{for} \quad n = 0 \\ &= \frac{\alpha}{\pi}\cos(\alpha x) && \text{for} \quad n = 1 \\ &= \frac{-\alpha^2}{\pi}\sin(\alpha x) && \text{for} \quad n = 2 \\ &\vdots \\ &= \begin{cases} \dfrac{(-1)^q\alpha^{2q}}{\pi}\sin(\alpha x) & \text{for} \quad n = 2q \\ \dfrac{(-1)^q\alpha^{2q+1}}{\pi}\cos(\alpha x) & \text{for} \quad n = 2q+1 \quad (q \in Z^+). \end{cases}\end{aligned} \tag{62}$$

These are the basis functions for the well known sine and cosine transforms. Thus, the commonly used sine and cosine transforms can be regarded as special wavelets generated from the Shannon's scaling function (or the sinc function). These results are in contrast to Shannon's wavelet, $\frac{1}{\pi x}[\sin(2\pi x) - \sin(\pi x)]$. Obviously, all these wavelets can be used to generate orthogonal wavelet bases by the standard method of translation and dilation. It seems to us that if the starting scaling function generates an orthogonal system, then the corresponding wavelets created by the present wavelet generators, equation (51), can also be orthogonal systems.

**Example 4 (Regularized Shannon's wavelet family).** Obviously the expressions given in the last example do not decay asymptotically. The regularized Shannon's kernel given in equation (35)

$$\phi_{\alpha,\sigma}(x) = \frac{1}{\pi}\frac{\sin(\alpha x)}{x}\mathrm{e}^{-x^2/2\sigma^2} \tag{63}$$

can be used to generate a family of wavelets with the desired asymptotic properties:

$$\begin{aligned}
\psi_n(x) &= \left(x\frac{\partial^n}{\partial x^n} + n\frac{\partial^{n-1}}{\partial x^{n-1}}\right)\frac{1}{\pi}\frac{\sin(\alpha x)}{x}\mathrm{e}^{-x^2} \\
&= \frac{1}{\pi}\sin(\alpha x)\mathrm{e}^{-x^2} && \text{for} \quad n = 0 \\
&= \frac{1}{\pi}[\alpha\cos(\alpha x) - 2x\sin(\alpha x)]\mathrm{e}^{-x^2} && \text{for} \quad n = 1 \\
&= \frac{1}{\pi}[-\alpha^2\sin(\alpha x) - 4\alpha x\cos(\alpha x) - 2\sin(\alpha x) \\
&\quad +4x^2\sin(\alpha x)]\mathrm{e}^{-x^2} && \text{for} \quad n = 2 \\
&= \frac{1}{\pi}[-\alpha^3\cos(\alpha x) + 6\alpha^2 x\sin(\alpha x) - 6\alpha\cos(\alpha x) \\
&\quad +12\alpha x^2\cos(\alpha x) - 8x^3\sin(\alpha x) + 12x\sin(\alpha x)]\mathrm{e}^{-x^2} && \text{for} \quad n = 3 \\
&= \frac{1}{\pi}[12\sin(\alpha x) + 12\alpha^2\sin(\alpha x) - 48x^2\sin(\alpha x) + 8\alpha^3 x\cos(\alpha x) + 16x^4\sin(\alpha x) \\
&\quad -24\alpha^2 x^2\sin(\alpha x) - 32\alpha x^3\cos(\alpha x) + \alpha^4\sin(\alpha x) \\
&\quad +48\alpha x\cos(\alpha x)]\mathrm{e}^{-x^2} && \text{for} \quad n = 4\cdots.
\end{aligned} \tag{64}$$

For simplicity, we have set $\sigma = \frac{1}{\sqrt{2}}$. The first few terms of regularized Shannon's wavelets and a comparison with Gauss kernel-generated wavelets are given in figure 2. When $n$ is even, all wavelets are antisymmetric with respect to the origin and symmetric wavelets are obtained for odd $n$. For $\alpha = 1$ (dots), each regularized Shannon's wavelet resembles a Gauss kernel-generated wavelet of the same order. When a larger $\alpha$ is taken ($\alpha = \pi$, crosses in figure 2), regularized Shannon's wavelets become more oscillatory. Obviously, by an appropriate selection of the parameter $\alpha$, regularized Shannon's wavelets are expected to perform better than the Gauss kernel-generated wavelets in high-accuracy scientific computations.

### *3.4. More general wavelet generators*

For a given $C^m$ symmetric scaling function $\phi$, a generalized family of wavelet generators is defined as

$$\begin{aligned}
&G_l^n = \sum_{p=0}^{n}\frac{n!}{(n-p)!p!}[x^{2l+1}]^{(p)}\frac{\partial^{n-p}}{\partial x^{n-p}} \\
&n = 0, 1, 2, \ldots, m \qquad l = 0, 1, 2, \ldots, (m-1)/2 \qquad \text{or} \qquad (m-2)/2.
\end{aligned} \tag{65}$$

Obviously the wavelet generator discussed above is a special case of the present definition and corresponds to $l = 0$. In the case of $l = 1$, wavelet generators are given by

$$G_1^n = x^3\frac{\partial^n}{\partial x^n} + 3nx^2\frac{\partial^{n-1}}{\partial x^{n-1}} + 3n(n-1)x\frac{\partial^{n-2}}{\partial x^{n-2}} + n(n-1)(n-2)\frac{\partial^{n-3}}{\partial x^{n-3}}. \tag{66}$$

These can be used to generate families of wavelets.

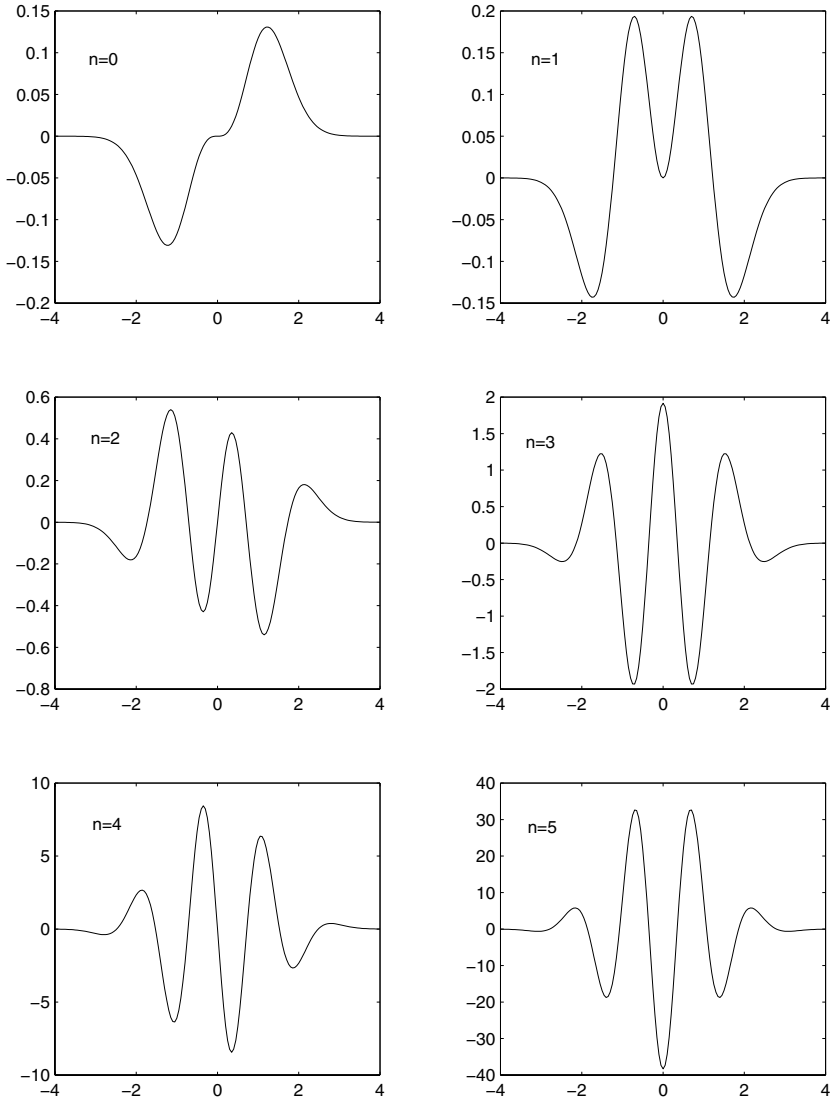


**Figure 3.** The first few terms of wavelets generated by $G_1^n \frac{1}{\sqrt{\pi}} e^{-x^2}$.

**Example 5 (Gauss's kernel-generated wavelets).** If the scaling function is given by the Gauss kernel, $\phi_\alpha(x) = \frac{1}{\sqrt{\pi\alpha}} e^{-x^2}$, we obtain a family of wavelets by using $G_1^n$

$$
\begin{aligned}
l\psi_{n,1}(x) &= G_1^n \frac{1}{\sqrt{\pi}} e^{-x^2} \\
&= \frac{(-1)^n}{2^3\sqrt{\pi}} [H_{n+3}(x) + 6H_{n+1}(x)] e^{-x^2} \qquad n = 0, 1, 2, \ldots . \qquad (67)
\end{aligned}
$$

The first few terms of these wavelets are plotted in figure 3.

**Example 6 (Shannon's wavelet family).** A new Shannon's wavelet family can be generated by applying generator $G_1^n$ to $\phi_\alpha(x) = \frac{1}{\pi}\frac{\sin(\alpha x)}{x}$. After some tedious derivations, we have

$$\psi_{n,1}(x) = G_1^n \frac{1}{\pi}\frac{\sin(\alpha x)}{x}$$
$$= \begin{cases} \dfrac{(-1)^q \alpha^{2q} x^2}{\pi}\sin(\alpha x) + \dfrac{(-1)^{q-1} 4q\alpha^{2q-1} x}{\pi}\cos(\alpha x) & \\ \quad + \dfrac{(-1)^{q-1} 2q(2q-1)\alpha^{2q-2}}{\pi}\sin(\alpha x) & \text{for} \quad n{=}2q \\ \dfrac{(-1)^q \alpha^{2q+1} x^2}{\pi}\cos(\alpha x) + \dfrac{(-1)^q (4q+2)\alpha^{2q} x}{\pi}\sin(\alpha x) & \\ \quad + \dfrac{(-1)^{q-1} 2q(2q+1)\alpha^{2q-1}}{\pi}\cos(\alpha x) & \text{for} \quad n{=}2q+1 \\ (n, = 0, 1, 2, \ldots). & \end{cases} \tag{68}$$

Obviously, this is an unbounded function for large $x$ and $\alpha > 1$. If the regularized DSC kernel, $\phi_{\alpha,\sigma}(x) = \frac{1}{\pi}\frac{\sin(\alpha x)}{x} e^{-x^2/2\sigma^2}$, is used, the resulting wavelet, $G_1^n \phi_{\alpha,\sigma}(x)$, is well behaved on the real line.

A more detailed discussion on general properties, including Lie algebraic aspects of these generators, is beyond the scope of this paper.

### *3.5. Wavelets generated by difference pairs*

Another simple and efficient way of generating wavelets from DSC kernels is to take the difference between two normalized DSC kernels:

$$\psi_{\alpha,\beta}(x) = \phi_\alpha - \phi_\beta. \tag{69}$$

**Example 7 (Hermite wavelets and the Mexican hat wavelet).** In the case of Hermite's kernel, equation (34), we have

$$\psi_{n,n'} = e^{-x^2/2} \sum_{k=0}^{n} \left(\frac{-1}{4}\right)^k \frac{1}{\sqrt{2\pi}k!} H_{2k}\left(\frac{x}{\sqrt{2}}\right) - e^{-x^2/2} \sum_{k=0}^{n'} \left(\frac{-1}{4}\right)^k \frac{1}{\sqrt{2\pi}k!} H_{2k}\left(\frac{x}{\sqrt{2}}\right)$$
$$= e^{-x^2/2} \sum_{k=n'+1}^{n} \left(\frac{-1}{4}\right)^k \frac{1}{\sqrt{2\pi}k!} H_{2k}\left(\frac{x}{\sqrt{2}}\right). \tag{70}$$

This is a general expression for a family of nonorthogonal wavelets. In particular, if $n = 1, n' = 0$, we obtain

$$\psi_{1,0}(x) = \frac{1}{2\sqrt{2\pi}}(1 - x^2)e^{-x^2/2}. \tag{71}$$

This is, once again, the well known Mexican hat wavelet [36]. The Hermite wavelets described in equation (59), which differs by a constant, can be easily obtained by appropriately choosing $n' = n - 1$ in equation (70).

**Example 8 (Shannon's wavelet).** We can use Dirichlet's continuous kernel as a scaling function, $\phi_\alpha(x) = \frac{1}{\pi}\frac{\sin(\alpha x)}{x}$. Then the corresponding mother wavelets generated by equation (69) are

$$\psi_{\alpha,\beta}(x) = \frac{1}{\pi}\left[\frac{\sin(\alpha x)}{x} - \frac{\sin(\beta x)}{x}\right] \qquad \text{for} \quad \alpha \neq \beta \neq 0. \tag{72}$$

This family includes the well known Shannon's wavelet [36] as a special case:

$$\psi_{2\pi,\pi}(x) = \frac{1}{\pi x}[\sin(2\pi x) - \sin(\pi x)]. \tag{73}$$

It is easy to check that this Shannon's wavelet generates an orthogonal system, which is in contrast to the other family of Shannon's wavelets, equation (62).

**Example 9 (Gauss's wavelets).** It is noted that this procedure of generating wavelets is also very general. For example, a wavelet can be constructed by combining a pair of functions from the Gauss kernel:

$$\psi_{\alpha,\beta}(x) = \frac{1}{\sqrt{2\pi}}\left[\frac{1}{\alpha}e^{-x^2/2\alpha^2} - \frac{1}{\beta}e^{-x^2/2\beta^2}\right] \qquad \text{for} \quad \alpha \neq \beta \neq 0. \tag{74}$$

Note that this is not a special case of example 1.

In the case where there is more than one parameter, the corresponding wavelets can be generated as differences of cross terms.

## 4. Discussion

The general connection between the delta sequence kernels used in the DSC algorithm and wavelets has been brought out in some detail. Qualitatively, DSC kernels can be regarded as wavelet scaling functions. If a DSC kernel is an orthogonal system, it is found to span the wavelet subspace $V_0$ in a multiresolution analysis. Many DSC kernels arising in mathematical, physical and engineering applications are discussed. Most DSC kernels of Dirichlet type have been found to be extremely efficient and robust for numerical solution of partial differential equations [22–25] and for image processing [33]. The use of DSC kernels of positive type for numerical computations is under consideration.

A family of wavelet generators is constructed for converting DSC kernels into mother wavelets. The second member of the family is exactly the wavelet generator given by Holschneider *et al* [40, 41]. The use of the wavelet generator given by Holschneider *et al* is restricted to non-negative functions, whereas many examples in this paper are constructed from DSC kernels of Dirichlet type. By an appropriate choice of the scale parameter, wavelets generated by DSC kernels of Dirichlet type are more oscillatory. Thus they have an advantage for numerical approximation of functions with high-frequency components.

The generators are connected with an infinite-dimensional Lie algebra which has an extremely simple algebraic structure and includes the algebra of translation and dilation as an invariant subalgebra. The corresponding Lie group provides a mathematical description of wavelets, which is more general than the usual translation and dilation group. A set of new orthogonal wavelets is found in the case of Dirichlet's continuous kernel, which enables us to make a connection between Shannon's wavelets and the standard sine and cosine transforms. The well known Mexican hat wavelet has been shown to be a special case of a variety of Hermite wavelets, derived by two distinct approaches.

The general connection between wavelet bases and conventional $L^2(a, b)$ polynomial bases is briefly discussed. Essentially, the lowest-order polynomial corresponds to a scaling function and all higher-order polynomials are related to mother wavelets, provided that the polynomials are orthogonal with respect to an appropriate weight.

The wavelet property of compact support is emphasized by the applied mathematics and computer science communities. However, the physics community seems to be more interested in other wavelet features, such as time–frequency localization and multiresolution analysis. In fact, non-compactly supported wavelets and Wigner functions have far more impact on the

physics community than the compactly supported ones. We note that DSC kernel-generated wavelets discussed in this paper are mostly non compact support. Our numerical experience indicates that compact supportness is not very important in digital computations. The important issue is, for a given algorithm, how to make the truncation error as small as one wishes.

In practical applications, the most attractive properties of wavelet methods are time–frequency localization and multiresolution analysis. Wavelet application to the numerical solution of PDEs can be formulated in a variety of ways, such as wavelet-Galerkin, wavelet-moment and wavelet-collocation. A common feature in most approaches is the use of wavelet multiresolution approximation of PDE operators. Such an approximation, when formulated on a nonuniform grid in each level of resolution, might have great potential for problems involving boundary-layer type of behaviour and/or problems associated with complex geometry where certain regions of the computational domain require a more dense grid. However, multiresolution approximation has to be carried out in cases where a multiresolution analysis is required for the purpose of understanding. For an ordinary task of solving a PDE, wavelet multiresolution approximation, particularly with a uniform grid in each level of resolution, provides no additional numerical advantage. In fact, technically, it creates an extra level of complication which reduces the numerical accuracy, stability and computational efficiency. In the author's opinion, such an extra level of complication hinders the use of wavelet methods for solving PDEs.

In the DSC algorithm, the usage is made of the wavelet feature of time–frequency localization. This is particularly true for regularized DSC kernels. In contrast to wavelet methods, the DSC algorithm performs computations only at the highest (single) level of resolution for a given mesh.

In our recent work [55], the DSC kernels and their derivatives have been used in image processing as low-pass filters and high-pass filters, respectively. These DSC filters have excellent performance in image edge detection and feature extraction. It is expected that the DSC kernel-generated wavelets discussed in this paper will be useful for image processing too. In fact, the Canny operator, a well known filter in edge detection, is identified with the first member of the Gauss kernel-generated wavelet family. The second member of the wavelet family happens to be the Mexican hat wavelet. The latter is widely used in image processing and pattern recognition. The higher-order terms generated by the present method are believed to have better frequency response to higher-frequency components and can be used to improve computational accuracy.

## Acknowledgments

This work was supported in part by the National University of Singapore and in part by the Natural Science and Engineering Research Council of Canada. The author is grateful to the referees for valuable comments and suggestions.

## References

[1] Lanczos C 1938 *J. Math. Phys.* **17** 123
[2] Harris D O, Engerholm G G and Gwinn W D 1965 *J. Chem. Phys.* **43** 1515
[3] Cooley J W and Tukey J W 1965 *Math. Comput.* **19** 297
[4] Finlayson B A and Scriven L E 1966 *Appl. Mech. Rev.* **19** 735
[5] Orszag S A *Stud. Appl. Math.* **51** 253
[6] Kreiss H O and Oliger J 1972 *Tellus* **24** 199
[7] Finlayson B A 1972 *The Method of Weighted Residuals and Variational Principles* (New York: Academic)
[8] Fornberg B 1973 *Math. Comput.* **27** 45

[9] Lill J V, Parker G A and Light L C 1984 *Chem. Phys. Lett.* **89** 483
[10] Kosloff D and Baysal E 1982 *Geophysics* **47** 1402
[11] Hutson J M and Le Roy R J 1985 *J. Phys. B: At. Mol. Phys.* **83** 1197
[12] Friesner R 1985 *Chem. Phys. Lett.* **116** 39
[13] Yang W and Peet A C 1988 *Chem. Phys. Lett.* **153** 98
[14] Hoffman D K, Nayar N, Sharafeddin O A and Kouri D J 1991 *J. Phys. Chem.* **95** 8299
[15] Colbert D T and Miller W H 1992 *J. Phys. Chem.* **96** 1982
[16] Fornberg B 1996 *A Practical Guide to Pseudospectral Methods* (Cambridge: Cambridge University Press)
[17] Wei G W, Zhang D S, Kouri D J and Hoffman D K 1997 *Phys. Rev. Lett.* **79** 775
[18] Wei G W 1998 *Chem. Phys. Lett.* **296** 215
[19] Hu X G, Ho T S and Rabitz H 1998 *Comput. Phys. Commun.* **113** 168
[20] Kadanoff L P 1997 *Phys. Today* **50** 11
[21] Ablowitz M J, Herbst B M and Schober C 1996 *J. Comput. Phys.* **126** 299
[22] Wei G W 1999 *J. Chem. Phys.* **110** 8930
[23] Wei G W 2000 *Physica* D **137** 247
[24] Wei G W 2000 *J. Phys. B: At. Mol. Opt. Phys.* **33** 343
[25] Wei G W 2000 *J. Phys. A: Math. Gen.* **33** 4935
[26] Schwartz L 1951 *Théore des Distributions* (Paris: Hermann)
[27] Wei G W 1999 *IEEE Proc. of Asia Pacific Microwave Conf. (Singapore)* p 562
[28] Wei G W 2000 *Comput. Methods Appl. Mech. Eng.* at press
[29] Wei G W 1999 *Computational Mechanics for the Next Millennium* ed C M Wang, K H Lee and K K Ang (New York: Elsevier) p 1049
[30] Wei G W 2000 *J. Sound Vib.* submitted
[31] Wan D C and Wei G W 1999 *Int. J. Numer. Methods Fluids* submitted
[32] Wan D C and Wei G W 1999 *Comput. Methods Appl. Mech. Eng.* submitted
[33] Wei G W 1999 *IEEE Signal Process. Lett.* **6** 165
[34] Guan S, Lai C H and Wei G W 2000 *Physica* D at press
[35] Meyer Y 1992 *Wavelets and Operators* (*Cambridge Stud. Adv. Math. vol 37*) (Cambridge: Cambridge University Press)
[36] Daubechies I 1988 *Commun. Pure Appl. Math.* **41** 909
Daubechies I 1992 *Ten Lectures on Wavelets (CBMS-NSF Series in Applied Mathematics)* (Philadephia: SIAM)
[37] Mallat S G 1989 *Trans. Am. Math. Soc.* **315** 69
[38] Chui C K 1992 *An Introduction to Wavelets* (San Diego: Academic)
[39] Kaiser G 1994 *A Friendly Guide To Wavelets* (Boston: Birkhäuser)
[40] Holschneider M, Kronland-Martinet R, Morlet J and Tachamitchian Ph 1988 *Wavelets, Time-Frequency Methods and Phase Space* ed J M Combes, Grossmann and Ph Tachamitchian (Berlin: Springer) pp 286–97
[41] Holschneider M 1995 *Wavelets, An Analysis Tool* (Oxford: Oxford Science Publications)
[42] Hernandez E and Weiss G 1996 *A First Course on Wavelets* (Boca Raton, FL: Chemical Rubber Company)
[43] Wickerhauser M V 1994 *Adapted Wavelet Analysis from Theory to Software* (Wellesley, MA: A. K. Peters)
[44] Vetterli M and Kovacevic J 1995 *Wavelets and Subband Coding* (Englewood Cliffs, NJ: Prentice-Hall)
[45] Meyer Y and Coifman R 1997 *Wavelets, Calderón-Zygmund and Multilinear Operators* (*Cambridge Stud. Adv. Math. vol 48*) (Cambridge: Cambridge University Press)
[46] See, for example, Szu H H 1997 *Wavelet Applications IV* (SPIE Proc.) (*Orlando*)
[47] John S 1998 *J. Math. Phys.* **39** 569
[48] Beylkin G and Keiser J M 1996 *J. Comput. Phys.* **132** 233
[49] Korevaar J 1955 *Ned. Akad. Westensh Proc. Ser.* A **58** 368
Korevaar J 1955 *Ned. Akad. Westensh Proc. Ser.* A **58** 483
Korevaar J 1955 *Ned. Akad. Westensh Proc. Ser.* A **58** 663
Korevaar J 1968 *Mathematical Methods* vol 1 (New York: Academic)
[50] Walter G G and Blum J 1977 *Ann. Stat.* **7** 328
[51] Winter B B 1975 *Ann. Stat.* **3** 759
[52] Wahba G 1975 *Ann. Stat.* **3** 15
[53] Kronmal R and Tarter M 1968 *J. Am. Stat. Assoc.* **63** 925
[54] Bohm A 1993 *Quantum Mechanics* 3rd edn (Berlin: Springer)
[55] Hou Z J and Wei G W 2000 *Pattern Recognition* submitted