Chapter 1

Continuous and discrete reproducing systems that arise from translations. Theory and applications of composite wavelets.

Demetrio Labate and Guido Weiss

Abstract Reproducing systems of functions such as the wavelet and Gabor systems have been particularly successful in a variety of applications from both mathematics and engineering. In this chapter, we review a number of recent results in the study of such systems and their generalizations developed by the authors and their collaborators. We first describe the unified theory of reproducing systems. This is a simple and flexible mathematical framework to characterize and analyze wavelets, Gabor systems and other reproducing systems in a unified manner. The systems of interest to us are obtained by applying families of translations, modulations and dilations to a countable set of functions. As the reader will see, we can rewrite such systems as a countable family of translations applied to a countable collection of functions. Building in part on this approach, we define the wavelets with composite dilations, a novel class of reproducing systems which provide truly multidimensional generalizations of traditional wavelets. For example, in dimension two, the elements of such systems are defined not only at various scales and locations, as traditional wavelet systems, but also at various orientations. The shearlet system is a special case of a composite wavelet system which provides optimally sparse representation for a large class of bivariate functions. This is useful for a number of applications in image processing, such as image denoising and edge detection. Finally, we discuss some related issues about the continuous wavelet transform and the continuous analogues of composite wavelets.

Demetrio Labate
North Carolina State University, Raleigh, NC 27695, USA e-mail: dlabate@math.ncsu.edu

Guido Weiss
Washington University in St. Louis, St. Louis, MO 63130, USA e-mail: guido@math.wustl.edu

1.1 Introduction

These lectures present an overview of a program of research developed by the authors and their collaborators at Washington University in St.Louis during the past 10 years, which is devoted to the study of reproducing systems of functions. By *reproducing systems of functions*, we refer to those families of functions $\{\psi_i : i \in \mathscr{I}\}$ in $L^2(\mathbb{R}^n)$ which are obtained by applying a countable collections of operators to a countable set of "generating" functions and have the property that any function $f \in L^2(\mathbb{R}^n)$ can be recovered from the reproducing formula

$$f = \sum_{i \in \mathscr{I}} \langle f, \psi_i \rangle \, \psi_i,$$

with convergence in the L^2–norm. The *wavelet systems*, for example, have received a great deal of attention in the last 20 years, since their applications in mathematics and engineering have been especially successful. In dimension $n = 1$, they are defined as those collections of the form

$$\Psi = \{\psi_{j,k} = 2^{j/2}\, \psi(2^j \cdot - k) : \; j,k \in \mathbb{Z}\}, \tag{1.1}$$

where ψ is a fixed function in $L^2(\mathbb{R})$. As the expression above shows, Ψ is obtained by applying dyadic dilations and integer translations to the generating function ψ. For particular choices of the generator ψ, the wavelet system Ψ is an orthonormal basis or a Parseval frame for $L^2(\mathbb{R})$, in which case any $f \in L^2(\mathbb{R}^n)$ can be recovered as

$$f = \sum_{j,k \in \mathbb{Z}} \langle f, \psi_{j,k} \rangle \, \psi_{j,k}, \tag{1.2}$$

with convergence in the L^2–norm. Other important classes of reproducing systems are the Gabor systems, which are obtained by applying translations and modulations to a fixed generator, and the wave packet systems, which involve translations, dilations and modulations.

One main theme developed in these lectures is that there is a general framework which allows us to describe and analyze wavelet systems, Gabor systems and many other reproducing systems by using a unified approach. Indeed, for a large class of reproducing systems of the form

$$\{g_p(\cdot - C_p k) : k \in \mathbb{Z}^n, p \in \mathscr{P}\}, \tag{1.3}$$

where $\mathscr{P}$ is countable and $\{C_p\}$ is a set of invertible matrices, there is a relatively simple set of equations which characterizes those generating functions $\{g_p\}_{p \in \mathscr{P}}$ such that the corresponding system (1.3) is an orthonormal basis or, more generally, a Parseval frame for $L^2(\mathbb{R}^n)$. For example, it was discovered by Gripenberg [14] and Wang [41] independently, in 1995, that a function $\psi \in L^2(\mathbb{R})$ is the generator of an orthonormal wavelet system if and only if $\|\psi\|_2 = 1$,

$$\sum_{j\in\mathbb{Z}} |\hat{\psi}(2^j\xi)|^2 = 1 \qquad \text{for a. e. } \xi \in \mathbb{R}, \tag{1.4}$$

and

$$t_q(\xi) = \sum_{j\geq 0} \hat{\psi}(2^j\xi)\overline{\hat{\psi}(2^j(\xi+q))} = 0 \qquad \text{for a.e. } \xi \in \mathbb{R}, \tag{1.5}$$

whenever q is an odd integer. It is remarkable that a similar set of characterization equations holds not only for wavelet systems in higher dimensions, but also for many other reproducing systems. This topic, and the corresponding *unified theory of reproducing systems* will be presented in Section 1.2.

Parallel to the unified theory mentioned above, there is another "unifying" perspective to the study of reproducing systems which are provided by representation theory and, more specifically, by the study of the continuous wavelet transform and its generalization. In Section 1.3, we introduce the continuous analogues of the wavelet systems (1.1), which are obtained by applying dilations (with respect to a *dilation* group) and continuous translations to a function $\psi \in L^2(\mathbb{R}^n)$. For example, in dimension $n=1$, the continuous wavelet system is a system of the form

$$\{\psi_{at} = a^{-1/2}\psi(a^{-1}(\cdot - t)) : a > 0, t \in \mathbb{R}\},$$

and the (one-dimensional) *continuous wavelet transform* is the mapping

$$f \mapsto \left\{\langle f, \psi_{at}\rangle = a^{-1/2}\int_0^\infty f(y)\overline{\psi(a^{-1}(y-t))}\,dy : (a,t) \in \mathbb{R}^+ \times \mathbb{R}\right\}.$$

Then, provided that ψ satisfies a certain admissibility condition, any $f \in L^2(\mathbb{R})$ can be expressed using the *Calderòn reproducing formula*:

$$f = \int_{\mathbb{R}}\int_0^\infty \langle f, \psi_{at}\rangle\, \psi_{at}\, \frac{da}{a}\, dt. \tag{1.6}$$

The close relationship between the discrete and continuous framework is apparent by comparing the last expression with formula (1.2). A number of observations concerning this relationship, as well as several multidimensional extensions of the continuous wavelet transform are discussed in Section 1.3.

Traditional multidimensional wavelet systems are obtained by taking tensor products of one-dimensional ones and, as a result, they have a very limited capability to deal effectively with those directional features which typically occur in images and other multidimensional data. To overcome such limitations, several extensions and generalizations have been proposed in applied harmonic analysis during the last 10 years. One such approach is the *theory of wavelets with composite dilations*, which was originally introduced by the authors and their collaborators, and provides a very flexible and powerful framework to construct "truly" multidimensional extensions of the wavelet approach.

An example of a composite wavelet system, in dimension $n=2$, is the collection:

$$\{\psi_{ijk} = |\det A|^{i/2}\, \psi(B^jA^i \cdot -k) : i, j \in \mathbb{Z}, k \in \mathbb{Z}^2\}, \tag{1.7}$$

where $A = \begin{pmatrix} 2 & 0 \\ 0 & \sqrt{2} \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. The elements of such systems are defined not only at various scales and locations, as traditional wavelet systems, but also at various orientations, associated with the powers of the *shearing matrix* B. In additions, for appropriate choices of ψ, the elements ψ_{ijk} have the ability to provide very efficient representations for data containing directional and anisotropic features (see Section 1.5). There are a variety of systems of the form (1.7) forming Parseval frames or even orthonormal bases, for many choices of matrices A and B. Indeed, the theory of wavelets with composite dilations encompasses the theory of wavelets, and there is a generalized Multiresolution Analysis associated with this theory. As in the case of the classical MRA, this framework allows one to obtain a variety of constructions with many different geometric and analytic properties. An outline of this theory is presented in Section 1.4.

In Section 1.5, we examine a generalization of the wavelet transform associated with the affine group

$$G = \{(M,t) : M \in \mathscr{D}_\alpha, t \in \mathbb{R}^2\},$$

where, for each $0 < \alpha < 1$, $\mathscr{D}_\alpha \subset GL_2(\mathbb{R})$ is the set of matrices:

$$\mathscr{D}_\alpha = \left\{ M = M_{as} = \begin{pmatrix} a & -a^\alpha s \\ & \\ 0 & a^\alpha \end{pmatrix}, \quad a > 0, s \in \mathbb{R} \right\}.$$

Associated with this is the *continuous shearlet transform* $\mathscr{S}_\psi^\alpha$, defined by

$$f \to \{\mathscr{S}_\psi^\alpha f(a,s,t) = \langle f, \psi_{ast} \rangle : a > 0, s \in \mathbb{R}, t \in \mathbb{R}^2\},$$

which is mapping $f \in L^2(\mathbb{R}^2)$ into a transform domain dependent on the scale a, the shearing parameter s and the location t. The analyzing elements ψ_{ast}, forming a *continuous shearlet system*, are the functions

$$\psi_{ast}(x) = |\det M_{as}|^{-\frac{1}{2}} \psi(M_{as}^{-1}(x-t)), \tag{1.8}$$

with $M_{as} \in \mathscr{D}_\alpha$. One remarkable property is that the continuous shearlet transform of a function f has the ability to completely characterize both the location and the geometry of the set of singularities of f.

A *discrete shearlet system* is obtained by appropriately discretizing the functions (1.8). Indeed, such a discrete system can be designed so that it forms a Parseval frame and it provides us with a special case of wavelets with composite dilations (1.7). In addition, the generator ψ can be chosen to be a well-localized function; that is, ψ has fast decay both in the space and the frequency domains (see [19, 21]). As a result, the elements of the discrete shearlet systsm form a collection of well-localized waveforms at various scales, locations and orientations and provide optimally sparse representations for a large class of bivariate functions with distributed discontinuities. Only the curvelets introduced by Candès and Donoho have been proved to have similar properties; however, the curvelets do not share the simple affine-like structure of wavelets with composite dilations. To illustrate the advantages of the shearlet framework with respect to wavelets and other traditional representations, we describe a number of useful applications of shearlets to the analysis and processing of images, including some representative applications of feature extraction and edge detection.

1.2 Unified Theory of Reproducing Systems

In order to describe the types of reproducing systems that we will consider in this study, it will be useful to introduce the following definitions. We adopt the convention that $x \in \mathbb{R}^n$ is a column vector, i.e., $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, and that $\xi \in \widehat{\mathbb{R}}^n$ is a row vector, i.e., $\xi = (\xi_1, \ldots, \xi_n)$. A vector x multiplying a matrix $M \in GL_n(\mathbb{R})$ on the right is understood to be a column vector, while a vector ξ multiplying M on the left is a row vector. Thus, $Mx \in \mathbb{R}^n$ and $\xi M \in \widehat{\mathbb{R}}^n$.

Let $f \in L^2(\mathbb{R}^n)$. For $y \in \mathbb{R}^n$, the *translation operator* T_y is defined by $T_y f(x) = f(x-y)$; for $M \in GL_n(\mathbb{R})$, the *dilation operator* D_M is defined by $D_M f(x) = |\det M|^{-1/2} f(M^{-1}x)$; for $v \in \mathbb{R}^n$, the *modulation operator* M_v is defined by $(M_v f)(x) = e^{2\pi i v x} f(x)$.

We will use the Fourier transform in the form

$$\hat{f}(\xi) = \int_{\mathbb{R}^n} f(x)\, e^{-2\pi i \xi x}\, dx,$$

for $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$. Thus the inverse Fourier transform is given by

$$\check{f}(x) = \int_{\widehat{\mathbb{R}}^n} f(\xi)\, e^{2\pi i \xi x}\, d\xi.$$

We remark that $(T_y f)^\wedge(\xi) = (M_y \hat{f})(\xi)$ and $(D_M f)^\wedge(\xi) = (\widehat{D}_M \hat{f})(\xi)$, where $(D_M f)^\wedge(\xi) = (\hat{D}_M \hat{f})(\xi) = |\det M|^{1/2}\, \hat{f}(\xi M)$.

Virtually all systems of functions which are used in harmonic analysis to generate subspaces of $L^2(R^n)$ are obtained by applying a certain combination of translations, dilations and modulations to a finite family of functions in $L^2(\mathbb{R}^n)$. Let us start by recalling the definitions of the systems commonly used in many harmonic analysis applications.

• *Gabor Systems.* Let $\Psi = \{\psi^1, \ldots, \psi^L\} \subset L^2(\mathbb{R}^n)$, and $B, C \in GL_n(\mathbb{R})$. The *Gabor systems* are the collections

$$\mathscr{G} = \mathscr{G}_{B,C}(\Psi) = \{M_{Bm}\, T_{Ck}\, \psi^\ell : m, k \in \mathbb{Z}^n, \ell = 1, \ldots, L\}$$

or

$$\tilde{\mathscr{G}} = \tilde{\mathscr{G}}_{B,C}(\Psi) = \{T_{Ck}\, M_{Bm}\, \psi^\ell : m,k \in \mathbb{Z}^n, \ell = 1,\dots,L\}.$$

Notice that $\tilde{\mathscr{G}}$ is obtained from $\mathscr{G}$ by interchanging the order of the translation and modulation operators. Also, it is easy to see that

$$M_{Bm}\, T_{Ck}\, \psi^\ell = e^{-2\pi i BmCk}\, T_{Ck}\, M_{Bm}\, \psi^\ell.$$

• *Affine Systems.* Given $\Psi = \{\psi^1,\dots,\psi^L\} \subset L^2(\mathbb{R}^n)$, $A \subset GL_n(\mathbb{R})$ and $\Gamma \subset \mathbb{R}^n$, the *affine systems* are the collections

$$\mathscr{F} = \mathscr{F}_{A,\Gamma}(\Psi) = \{D_a\, T_\gamma\, \psi^\ell : a \in A, \gamma \in \Gamma, \ell = 1,\dots,L\}.$$

Very often we use the notation $\mathscr{D} = \{M^j : j \in \mathbb{Z}\}$, where $M \in GL_n(\mathbb{R})$ is expanding (i.e., each proper value λ of M satisfies $|\lambda| > 1$), and Γ is the lattice $C\mathbb{Z}^n$, where $C \in GL_n(\mathbb{R})$.

• *Wave Packet Systems.* These include the above two systems. For $\Psi = \{\psi^1,\dots,\psi^L\}$, they consist of those functions

$$\mathscr{WP}_{\Gamma,A,S}(\Psi) = \{T_\gamma D_a M_y\, \psi^\ell : \gamma \in \Gamma, a \in \mathscr{D}, y \in S, \ell = 1,\dots,L\},$$

where Γ, S are countable (or finite) subsets of $\mathbb{R}^n$, $A \subset GL_n(\mathbb{R})$. As will be discussed below, the order of the three operators T_γ, D_a, M_y can be permuted.

It is easy to see that each of the above systems can be expressed in the following form.

Let $\mathscr{P}$ be a countable indexing set, $\{g_p : p \in \mathscr{P}\}$ a family of functions in $L^2(R^n)$ and $\{C_p : p \in \mathscr{P}\}$ a corresponding collection of matrices in $GL_n(\mathbb{R})$. Then each of the systems we just described has the form:

$$\left\{T_{C_p k}\, g_p : k \in \mathbb{Z}^n, p \in \mathscr{P}\right\}. \tag{1.9}$$

Indeed, in order to write down the general wave packet system into the form (1.9), one needs just to use the "commutativity relations" $D_M\, T_k = T_{Mk}\, D_M$ and $M_y\, T_k = e^{2\pi i yk} T_k\, M_y$ (notice that $e^{2\pi i yk}$ is a constant of absolute value 1).

1.2.1 Unified Theorem for Reproducing Systems

In the theory of wavelets and, more generally, in Harmonic Analysis, it is of paramount importance to construct such systems that form a reproducing set for the space $L^2(\mathbb{R}^n)$ (or more general function spaces). For example, it is of particular interest to know when a system $\{\phi_j : j \in \mathbb{Z}\}$ of functions in $L^2(\mathbb{R}^n)$ is an orthonormal basis or, more generally, a frame. Many characterizations of systems that are Parseval frames have been given in the literature; most often these results concern themselves with affine systems [14, 24, 25, 32, 38, 41].

We shall now give necessary and sufficient conditions for the system (1.3) to be a Parseval frame for $L^2(\mathbb{R}^n)$. For simplicity, we are letting the lattice Γ to be $\mathbb{Z}^n$; our arguments below can be easily extended to a more general Γ.

Recall that a countable collection $\{\phi_i\}_{i\in I}$ in a (separable) Hilbert space $\mathscr{H}$ is a *Parseval frame* (sometimes called a *tight frame* with constant 1) for $\mathscr{H}$ if

$$\sum_{i\in I} |\langle f,\phi_i\rangle|^2 = \|f\|^2, \quad \text{for all } f \in \mathscr{H}.$$

This is equivalent to the reproducing formula $f = \sum_i \langle f,\phi_i\rangle\, \phi_i$, for all $f \in \mathscr{H}$, where the series converges unconditionally in the norm of $\mathscr{H}$. This shows that a Parseval frame provides a basis-like representation even though a Parseval frame need not be a basis in general. We refer the reader to [4, 6] for more details about frames.

We refer to the following result as the "Unifying Theorem for reproducing systems" [25]:

Theorem 1. *Let $\mathscr{P}$ be a countable indexing set, $\{g_p\}_{p\in\mathscr{P}}$ a collection of functions in $L^2(\mathbb{R}^n)$ and $\{C_p\}_{p\in\mathscr{P}} \subset GL_n(\mathbb{R})$. Let*

$$\mathscr{E} = \left\{ f \in L^2(\mathbb{R}^n) \colon \hat{f} \in L^\infty(\mathbb{R}^n) \text{ and supp } \hat{f} \text{ is compact} \right\},$$

and suppose that

$$\mathscr{L}(f) = \sum_{p\in\mathscr{P}} \sum_{m\in\mathbb{Z}^n} \int_{supp\ \hat{f}} |\hat{f}(\xi + mC_p^{-1})|^2 \frac{1}{|\det C_p|} |\hat{g}_p(\xi)|^2 \, d\xi < \infty \tag{1.10}$$

for all $f \in \mathscr{E}$. Then the system (1.3) is a Parseval frame for $L^2(\mathbb{R}^n)$ if and only if

$$\sum_{p\in\mathscr{P}_\alpha} \frac{1}{|\det C_p|} \overline{\hat{g}_p(\xi)}\, \hat{g}_p(\xi+\alpha) = \delta_{\alpha,0} \quad \text{for a.e. } \xi \in \mathbb{R}^n, \tag{1.11}$$

for each $\alpha \in \Lambda = \bigcup_{p\in\mathscr{P}} \mathbb{Z}^n C_p^{-1}$, where $\mathscr{P}_\alpha = \{p \in \mathscr{P} \colon\ \alpha C_p \in \mathbb{Z}^n\}$ and δ is the Kronecker delta for $\mathbb{R}^n$.

Before discussing the proof of this theorem, it will be useful to make a few comments about this result, in order to elucidate its context and its impact.

Remark 1. It is relatively well known that if $\psi \in L^2(\mathbb{R})$, then $\{\psi_{jk} = D_{2^j} T_k \psi \colon\ j,k \in \mathbb{Z}\}$ is an orthonormal basis for $L^2(\mathbb{R})$ (i.e., ψ is an *orthonormal wavelet*) if and only if equations (1.4) and (1.5) hold. As we mentioned above, this result was obtained independently by G. Gripenberg [14] and X. Wang [41]. As it will be discussed below, these equations are a simple consequence of Theorem 1 (see exercise 1, at the end of this section).

Remark 2. The assumption (1.10) is referred to as the Local Integrability Condition (LIC). At first sight, it might appear as a rather formidable technical hypothesis. In some cases, however, it can be shown that it is a simple consequence of the system being considered. For example, let us consider the Gabor system $\tilde{\mathscr{G}}_{B,C}(G)$, where $G = \{g^1, \dots, g^L\}$, and let us write it in the form (1.3). Namely, let $\mathscr{P} = \mathbb{Z}^n \times \{1,2,\dots,L\}$, $g_p = g_{j,\ell} = M_{Bj}\, g^\ell$, and $C_p = C$, so that

$$T_{C_p k}\, g_p = T_{Ck}\, M_{Bj}\, g^\ell.$$

Without loss of generality, we can assume that $L = 1$. Thus, the expression of (1.10) is

$$\begin{aligned}\mathscr{L}(f) &= \sum_{p \in \mathscr{P}} \sum_{m \in \mathbb{Z}^n} \int_K |\hat{f}(\xi + m\,(C_p)^{-1}|^2\, |\hat{g}_p(\xi)|^2 \,\frac{d\xi}{|\det C_p|} \\ &= \frac{1}{|\det C|} \sum_{p \in \mathbb{Z}^n} \sum_{m \in \mathbb{Z}^n} \int_K |\hat{f}(\xi + m\,(C)^{-1}|^2\, |\hat{g}(\xi - B\,p)|^2\, d\xi,\end{aligned}$$

for $f \in \mathscr{E}$ and $K = \operatorname{supp} \hat{f}$ is compact. Since $\xi \in K$, only a finite number of terms in the sum $\sum_{m \in \mathbb{Z}^n}$ are non-zero. Moreover, if $\mathbb{T}^n$ is the n-torus, for each $j \in \mathbb{Z}^n$, the set $\{B(\mathbb{T}^n + j - p) : p \in \mathbb{Z}^n\}$ is a partition of $\mathbb{R}^n$. Thus,

$$\|g\|_2^2 = \int_{\bigcup_{p \in \mathbb{Z}^n} B(\mathbb{T}^n + j - p)} |\hat{g}(\eta)|^2\, d\eta = \sum_{p \in \mathbb{Z}^n} \int_{\bigcup_{p \in \mathbb{Z}^n} B(\mathbb{T}^n + j)} |\hat{g}(\xi - Bp)|^2\, d\xi.$$

Now observe that a finite union of the sets $\{B(\mathbb{T}^n + j) : j \in \mathbb{Z}^n\}$ covers K. Using this fact and the fact that $\|\hat{f}\|_\infty \le \infty$ (since $f \in \mathscr{E}$), it is not difficult to show that

$$\mathscr{L}(f) \le C\, \|g\|_2^2,$$

where C is a positive constant. As a result, the characterization theorem for the Gabor systems can be stated explicitly as:

Theorem 2. *The system $\mathscr{G}_{B,C}(G)$ (or the system $\widetilde{\mathscr{G}}_{B,C}(G)$) is a Parseval frame for $L^2(\mathbb{R}^n)$ if and only if*

$$\sum_{\ell=1}^{L} \sum_{k \in \mathbb{Z}^n} \frac{1}{|\det C|}\, \hat{g}^\ell(\xi - Bk)\, \overline{\hat{g}^\ell(\xi - Bk + mC^{-1})} = \delta_{m,0}$$

for a.e. $\xi \in \mathbb{R}^n$, all $m \in \mathbb{Z}^n$.

This result is well known, and can be found, for example, in [28, 39, 7, 32].

The situation for the "usual" affine systems is somewhat more subtle. Here, by the word "usual" we mean the case where $A = \{a^j : j \in \mathbb{Z}\}$ where $a \in GL_n(\mathbb{R})$ is expanding, and $\Gamma = \mathbb{Z}^n$. In this case one can show that, if the conditions (1.11) are true, then the LIC is valid and, conversely, if the system (1.3) is a Parseval frame, then the LIC also holds. Thus, in the characterization of Parseval frames given by Theorem 1 it is not needed to assume the LIC. The characterization theorem for these systems can be written down explicitly as:

Theorem 3. *Let $\Psi = \{\psi^1, \dots, \psi^L\} \subset L^2(\mathbb{R}^n)$ and $a \in GL_n(\mathbb{R})$ be expanding. Then the system $\mathscr{F}_{A,\Gamma}(\Psi) = \{D_{a^j}\, T_k\, \psi^\ell : j \in \mathbb{Z}, k \in \mathbb{Z}^n, \ell = 1, \dots, L\}$ is a Parseval frame for $L^2(\mathbb{R}^n)$ if and only if*

$$\sum_{\ell=1}^{L} \sum_{j \in \mathscr{P}_\alpha} \hat{\psi}^\ell(\xi\, a^{-j})\, \overline{\hat{\psi}^\ell((\xi+\alpha)\, a^{-j})} = \delta_{\alpha,0}, \text{ for a.e. } \xi \in \mathbb{R}^n, \tag{1.12}$$

for all $\alpha \in \Lambda = \bigcup_{j \in \mathbb{Z}} \mathbb{Z}^n\, a^j$, where $\mathscr{P}_\alpha = \{j \in \mathbb{Z} :\, \alpha\, a^{-j} \in \mathbb{Z}^n\}$.

Apart from the argument needed to establish the validity of the LIC which we mentioned above, this last theorem is a simple consequence of Theorem 1 once the system $\mathscr{F}_{A,\Gamma}(\Psi)$ is expressed in the form (1.3). Notice that there is a redundancy in the condition (1.12). Indeed an elementary argument shows that (1.12) can be simplified to

$$\sum_{\ell=1}^{L} \sum_{j \in \mathscr{P}_m} \hat{\psi}^\ell(\xi\, a^{-j})\, \overline{\hat{\psi}^\ell((\xi+m)\, a^{-j})} = \delta_{m,0}, \text{ for a.e. } \xi \in \mathbb{R}^n, \tag{1.13}$$

for all $m \in \mathbb{Z}^n$, where $\mathscr{P}_m = \{j \in \mathbb{Z} :\, m a^{-j} \in \mathbb{Z}^n\}$. It follows easily from this form of Theorem 3 that the result of Gripenberg and Wang (given in Remark 1) holds for $n = 1$ and $a = 2$.

In order to present the ideas involved in the proof of Theorem 1, it is useful to introduce the *C-bracket product* of $f, g \in L^2(\mathbb{R}^n)$, which, for $C \in GL_n(\mathbb{R})$, is defined by

$$[f,g](x;C) \,=\, \sum_{k \in \mathbb{Z}^n} f(x - Ck)\, \overline{g(x - Ck)}.$$

It is clear that $[f,g]$ is CZ^n- periodic; that is, $[f,g](x + Cm; C) = [f,g](x;C)$ for each $m \in \mathbb{Z}^n$.

That the system (1.3) is a Parseval frame for $L^2(\mathbb{R}^n)$ is equivalent to

$$N^2(f) = \sum_{p \in \mathscr{P}} \sum_{k \in \mathbb{Z}^n} |\langle f, T_{C_p k}\, g_p \rangle|^2 = \|f\|_2^2, \tag{1.14}$$

for all $f \in \mathscr{E}$ (recall that $\mathscr{E}$ is dense in $L^2(\mathbb{R}^n)$).

Using the fact that $\mathbb{R}^n = \bigcup_{l \in \mathbb{Z}^n} \{(\mathbb{T}^n - l)C^{-1}\}$ is a disjoint union, it follows easily that

$$\begin{aligned}
\sum_{p \in \mathscr{P}} \sum_{k \in \mathbb{Z}^n} |\langle f, T_{C_p k}\, g_p \rangle|^2 &= \sum_{k \in \mathbb{Z}^n} \Big| \int_{\mathbb{R}^n} \hat{f}(\xi)\, \overline{\hat{g}(\xi)}\, e^{2\pi i Ck\cdot\xi}\, d\xi \Big|^2 \\
&= \sum_{l \in \mathbb{Z}^n} \int_{C^I(\mathbb{T}^n)} \hat{f}(\xi - C^I l)\, \overline{\hat{g}(\xi - C^I l)}\, e^{2\pi i Ck\cdot\xi}\, d\xi \\
&= \int_{C^I(\mathbb{T}^n)} [\hat{f}, \hat{g}](\xi; C^I)\, e^{2\pi i Ck\cdot\xi}\, d\xi\,.
\end{aligned}$$

Under all these assumptions, let us consider the function

$$H(x) = \sum_{k\in\mathbb{Z}^n} |\langle T_x f, T_{Ck}\, g\rangle|^2,$$

where $C \in GL_n(\mathbb{R})$. Indeed, it is clear that the function H is CZ^n- periodic. Using the fact that $\hat{f}$ has compact support, one can show that

Lemma 1. *The function $H(x)$ is the trigonometric polynomial where*

$$H(x) = \sum_{m\in\mathbb{Z}^n} \hat{H}(m)\, e^{2\pi i (C^I m)\cdot x},$$

where

$$\hat{H}(m) = \frac{1}{|\det C|} \int_{\mathbb{R}^n} \hat{f}(\xi)\, \overline{\hat{f}(\xi + C^I m)}\, \overline{\hat{g}(\xi)}\, \hat{g}(\xi + C^I m)\, d\xi,$$

and only a finite number of these expressions is non-zero.

The fact that $\hat{H}(m) \neq 0$ for finitely many m at most follows from the fact that $\hat{f}$ has compact support.

To show that equality (1.14) holds for all $f \in \mathscr{E}$, consider now the function

$$w(x) = N^2(T_x f) = \sum_{p\in\mathscr{P}} H_p(x),$$

where $H_p(x) = |\langle T_x f, T_{C_p k}\, g_p\rangle|^2$. By Lemma 1, for each $p \in \mathscr{P}$,

$$H_p(x) = \sum_{m\in\mathbb{Z}^n} \hat{H}_p(m)\, e^{2\pi i (C_p^I m)\cdot x},$$

where

$$\hat{H}_p(m) = \frac{1}{|\det C_p|} \int_{\mathbb{R}^n} \hat{f}(\xi)\, \overline{\hat{f}(\xi + C_p^I m)}\, \overline{\hat{g}_p(\xi)}\, \hat{g}_p(\xi + C_p^I m)\, d\xi.$$

Thus, using the assumptions of Theorem 1, from the observations we made above, we have the expression

$$w(x) = N^2(T_x f) = \sum_{\alpha\in\Lambda} \hat{w}(\alpha)\, e^{2\pi i \alpha\cdot x}, \tag{1.15}$$

where

$$\hat{w}(\alpha) = \int_{\mathbb{R}^n} \hat{f}(\xi)\, \overline{\hat{f}(\xi+\alpha)} \sum_{p\in\mathscr{P}} \frac{1}{|\det C_p|}\, \overline{\hat{g}_p(\xi)}\, \hat{g}_p(\xi+\alpha)\, d\xi. \tag{1.16}$$

This integral is absolutely convergent, and the series defining $w(x)$ is absolutely and uniformly convergent. Notice that the LIC plays an important role to establish these convergence properties and the various uses of Fubini's theorem needed for the formulae developed here.

To complete the proof of Theorem 1 we argue as follows. Let us assume (1.11). Then, by equation (1.16),

$$\hat{w}(\alpha) = \delta_{\alpha,0} \int_{\mathbb{R}^n} \hat{f}(\xi)\, \overline{\hat{f}(\xi+\alpha)}\, d\xi.$$

By equation (1.15), this implies

$$w(x) = N^2(T_x f) = \sum_{\alpha\in\Lambda} \hat{w}(\alpha)\, e^{2\pi i\alpha\cdot x} = \hat{w}(0) = \|f\|_2^2.$$

Hence, the system (1.3) is a Parseval frame for $L^2(\mathbb{R}^n)$.

Conversely, let us now assume that the system (1.3) is a Parseval frame for $L^2(\mathbb{R}^n)$. Hence, by our assumptions, we know that

$$N^2(T_x f) = w(x) = \sum_{\alpha\in\Lambda} \hat{w}(\alpha)\, e^{2\pi i\alpha\cdot x} = \|T_x f\|_2^2 = \|f\|_2^2,$$

for all $f \in \mathscr{E}$.

Since Λ is countable and the "Fourier coefficients" $\hat{w}(\alpha)$ of this generalized Fourier series are unique, we must have $\hat{w}(\alpha) = 0$ if $\alpha \neq 0$ and $\hat{w}(0) = 1$. We can then use (1.16) and appropriate choices of $f \in \mathscr{E}$ to show that the equalities (1.11) must hold. For example, by letting f to be such that $\hat{f}(\xi) = \hat{f}_\varepsilon(\xi) = \frac{1}{\sqrt{|B(\varepsilon)|}} \chi_{B(\varepsilon)}(\xi - \xi_0)$, where $B(\varepsilon)$ is a ball of radius ε about the origin, $\varepsilon > 0$ and ξ_0 is a point of differentiability of the integral of $h(\xi) = \sum_{p\in\mathscr{P}} \frac{1}{|\det C_p|} |\hat{g}_p(\xi)|^2$, one obtains easily from (1.16) that $h(\xi_0) = 1$. This gives (1.11) when $\alpha = 0$.

This is, to conclude, the basic idea of the proof of Theorem 1. The role played by these generalized Fourier series is arrived at naturally; it arises from the importance of the notion of shift invariance which is essentially related to the structure of these families of reproducing systems.

Theorem 1 has many applications and several of them are described in [25, 26]. As mentioned above, they include Gabor, affine and wave packet systems. Theorem 1 applies also to the *quasi-affine systems*. In dimension $n = 1$, these are the systems $\{\tilde{\psi}_{jk} :\, j,k \in \mathbb{Z}\}$ obtained from $\psi \in L^2(\mathbb{R})$ by setting

$$\widetilde{\psi}_{j,k} = \begin{cases} 2^{j/2}\, T_k D_{2^{-j}}\, \psi^\ell, & j > 0 \\ D_{2^{-j}}\, T_k\, \psi^\ell, & j \le 0. \end{cases}$$

These systems (as well as their higher dimensional versions) were introduced by Ron and Shen in [37]. They pointed out that, unlike the affine systems, these systems are shift-invariant. Furthermore, the quasi-affine system $\{\widetilde{\psi}_{j,k}\}$ is a Parseval frame if and only if the corresponding affine system $\{\psi_{j,k}\}$ is a Parseval frame.

Recall that, in higher dimensions, affine and quasi-affine systems are typically defined using dilations of the form D_{M^j}, where M is an expanding matrix, that is, each proper value λ of M satisfies $|\lambda| > 1$. Notice that this condition is equivalent to the existence of constants k and γ, satisfying $0 < k \le 1 < \gamma < \infty$, such that

$$|M^j x| \ge k\, \gamma^j\, |x| \tag{1.17}$$

when $x \in \mathbb{R}^n$, $j \in \mathbb{Z}$, $j \geq 0$, and

$$|M^j x| \leq \frac{1}{k} \gamma^j |x| \tag{1.18}$$

when $x \in \mathbb{R}^n$, $j \in \mathbb{Z}$, $j \leq 0$. One remarkable property of Theorem 1 is that it applies not only to the case of expanding dilations matrices, but also to a more general class of dilations which are *expanding on a subspace* [25], and are defined as follows.

Definition 1. Given $M \in GL_n(\mathbb{R})$ and a non-zero linear subspace F of $\mathbb{R}^n$, we say that *M is expanding on F* if there exists a complementary (not necessarily orthogonal) linear subspace E of $\mathbb{R}^n$ with the following properties[1]:

(i) $\mathbb{R}^n = F + E$ and $F \cap E = \{0\}$; that is, for any $x \in \mathbb{R}^n$, there exist unique $x_F \in F$ and $x_E \in E$ such that $x = x_F + x_E$;
(ii) $M(F) = F$ and $M(E) = E$, that is, F and E are invariant under M;
(iii) conditions (1.17) and (1.18) hold for all $x \in F$;
(iv) For any $j \geq 0$, there exists $k_1 = k_1(M) > 0$ such that, $|x_E| \leq k_1 |M^j x_E|$.

It is clear that if a matrix M is expanding, then it is also expanding on a subspace. However, there are several examples of matrices which satisfy Definition 1 and are not expanding. For example, the following matrices are all expanding on a subspace:

- $M = \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix}$, where $a \in \mathbb{R}$, $|a| > 1$;
- $M = \begin{pmatrix} a & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$, where $a \in \mathbb{R}$, $|a| > 1$.

It is shown in [16, 25] that, for affine systems where the dilation matrix M is expanding on a subspace, according to the definition above, then the LIC is "automatically" satisfied. Hence, Theorem 3 applies to this class of affine systems as well.

The examples seem to suggest that Theorem 3 applies whenever the dilation matrix M has all eigenvalues $|\lambda_k| \geq 1$ and at least one eigenvalue $|\lambda_1| > 1$. However, this is not the case. In [16] there is an example of a 3×3 dilation matrix having eigenvalues $\lambda_1 = a > 1$ and $\lambda_2 = \lambda_3 = 1$, for which the LIC fails. Indeed it turns out that the information about the eigenvalues of M alone is not sufficient to determine the LIC or even the existence of corresponding affine systems. We refer to [27, 40] for additional results and observations about this topic.

Exercises

1. Show that equation (1.12) in Theorem 3 can be simplified to obtain (1.13). Next, show that, for $n = 1$, when the dilation matrix a is replaced by the dyadic factor

[1] This is the revised definition from [16]. It turned out that the definition initially proposed in [25], with a different condition (iv), was not sufficient to guarantee that the LIC was satisfied

2, equation (1.13) yields the "classical" Gripenberg-Wang equations (1.4) and (1.5).

2. Show that the matrices $M = \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix}$ and $M = \begin{pmatrix} a & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$, where $a > 1$, are expanding on a subspace (that is, they satisfy Definition 1).

1.3 Continuous Wavelet Transform

The *full affine group of motions on* $\mathbb{R}^n$, denoted by $\mathbf{A}_n$, consists of all pairs $(M,t) \in GL_n(\mathbb{R}) \times \mathbb{R}^n$ (endowed with the product topology) together with the group operation

$$(M,t) \cdot (M',t') = (MM', t' + (M')^{-1}t).$$

This operation is associated with the action $x \to M(x+t)$ on $\mathbb{R}^n$. The subgroup $\mathscr{N} = \{(M,t) \in \mathbf{A}_n : M = I, t \in \mathbb{R}^n\}$ is clearly a normal subgroup of $\mathbf{A}_n$.

We consider a class of subgroups $\{G\}$ of $\mathbf{A}_n$ of the form

$$G = \{(M,t) \in \mathbf{A}_n : M \in \mathscr{D}, t \in \mathbb{R}^n\},$$

where $\mathscr{D}$ is a closed subgroup of $GL_n(\mathbb{R})$. We can identify $\mathscr{D}$ with the subgroup $\{(M,t) \in G : M \in \mathscr{D}, t = 0\}$. Hence we refer to $\mathscr{D}$ as the *dilation subgroup* and to $\mathscr{N}$ as the *translation subgroup* of G. If μ is the left Haar measure for $\mathscr{D}$, then $d\lambda(M,t) = d\mu(M)\,dt$ is the element of the left Haar measure for G.

Let U be the unitary representation of G acting on $L^2(\mathbb{R}^n)$ defined by

$$\left(U_{(M,t)}\,\psi\right)(x) = |\det M|^{-1/2}\,\psi(M^{-1}x - t) := \psi_{M,t}(x), \tag{1.19}$$

for $(M,t) \in G$ and $\psi \in L^2(\mathbb{R}^n)$. The elements $\{\psi_{M,t} : (M,t) \in G\}$ are the *continuous affine systems* with respect to G. The corresponding expression in the frequency domain is:

$$\left(U_{(M,t)}\,\psi\right)^{\wedge}(\xi) = |\det M|^{1/2}\,\hat{\psi}(\xi M)\,e^{-2\pi i M t}.$$

For a fixed $\psi \in L^2(\mathbb{R}^n)$, the *wavelet transform* associated with G is the mapping

$$f \to (\mathscr{W}_\psi f)(M,t) = \langle f, \psi_{M,t} \rangle = |\det M|^{-1/2} \int_{\mathbb{R}^n} f(y)\,\overline{\psi(M^{-1}y - t)}\,dy,$$

where $f \in L^2(\mathbb{R}^n)$ and $(M,t) \in G$. If there exists a function $\psi \in L^2(\mathbb{R}^n)$ such that, for all $f \in L^2(\mathbb{R}^n)$, the reproducing formula

$$f = \int_G \langle f, \psi_{M,t} \rangle\, \psi_{M,t}\, d\lambda(M,t) \tag{1.20}$$

holds, then ψ is a *continuous wavelet* with respect to G. Expression (1.20) is a generalized version of the Calderòn reproducing formula (1.6) presented in Chapter 1.1. Notice that equality (1.20) is understood in the weak sense (see the proof of Theorem 4 below); the pointwise result is much more subtle.

The following theorem establishes an admissibility condition for ψ that guarantees that (1.20) is satisfied:

Theorem 4. *Equality (1.20) is valid for all $f \in L^2(\mathbb{R}^n)$ if and only if, for a.e. $\xi \in \mathbb{R}^n \setminus \{0\}$,*

$$\Delta_\psi(\xi) = \int_{\mathscr{D}} |\hat{\psi}(\xi M)|^2 \, d\mu(M) = 1. \tag{1.21}$$

Proof. Suppose that (1.21) is satisfied. Then, by direct computation we have that

$$\begin{aligned}
\|\mathscr{W}_\psi f\|^2_{L^2(G,\lambda)} &= \int_{\mathscr{D}} \int_{\mathbb{R}^n} |\langle f, \psi_{M,t} \rangle|^2 \, dt \, d\mu(M) \\
&= \int_{\mathscr{D}} \int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} \hat{f}(\xi) \, \hat{\psi}(\xi M) \, e^{2\pi i \xi M t} \, d\xi \right|^2 |\det M| \, dt \, d\mu(M) \\
&= \int_{\mathscr{D}} \left(\int_{\mathbb{R}^n} \left| \left(\hat{f}(\xi) \overline{\hat{\psi}(\cdot M)} \right)^{\vee} (Mt) \right|^2 |\det M| \, dt \right) d\mu(M) \\
&= \int_{\mathscr{D}} \int_{\mathbb{R}^n} |\hat{f}(\xi)|^2 |\hat{\psi}(\xi M)|^2 \, d\xi \, d\mu(M) \\
&= \int_{\mathbb{R}^n} |\hat{f}(\xi)|^2 \Delta_\psi(\xi) \, d\xi \\
&= \|f\|^2_{L^2(\mathbb{R}^n)}.
\end{aligned}$$

This shows that the mapping $\mathscr{W}_\psi : L^2(\mathbb{R}^n) \to L^2(G, \lambda)$ is an isometry. By polarization we then obtain

$$\langle \mathscr{W}_\psi f, \mathscr{W}_\psi g \rangle_{L^2(G)} = \langle f, g \rangle_{L^2(\mathbb{R}^n)}, \tag{1.22}$$

for all $f, g \in L^2(\mathbb{R}^n)$.

Conversely, suppose that equality (1.20) holds in the weak sense (i.e., (1.22) holds). Consider the expression

$$\int_{\mathbb{R}^n} |\hat{f}(\xi)|^2 \Delta_\psi(\xi) \, d\xi,$$

with f satisfying $|\hat{f}(\xi)|^2 = |\beta(r, \xi_0)|^{-1} \chi_{\beta(r,\xi_0)}(\xi)$, where $\beta(r, \xi_0)$ is a ball of radius r and center ξ_0, and ξ_0 is a point of differentiability of Δ_ψ. Then, by reversing the chain of equalities above, we obtain that

$$|\beta(r, \xi_0)|^{-1} \int_{\beta(r,\xi_0)} \Delta_\psi(\xi) \, d\xi = 1,$$

for all $r > 0$. By taking $\lim_{r \to 0^+}$, we conclude that $\Delta_\psi(\xi_0) = 1$. Thus, $\Delta_\psi(\xi) = 1$ for a.e. $\xi \in \mathbb{R}^n$. □

Theorem 4 can easily be extended to the case where G is not a subgroup of $GL_n(\mathbb{R})$, but simply a subset of $GL_n(\mathbb{R})$. Furthermore, Theorem 4 extends to functions on subspaces of $L^2(\mathbb{R}^n)$ of the form

$$L^2(V)^\vee = \{f \in L^2(\mathbb{R}^n) : \operatorname{supp} \hat{f} \subset V\}.$$

The proof of this fact is left as an exercise.

In the special case of Theorem 4 where $n = 1$ and $\mathscr{D} = \{2^j : j \in \mathbb{Z}\}$, equation (1.21) is $\sum_{j\in\mathbb{Z}} |\hat{\psi}(2^j\xi)|^2 = 1$ for a.e. $\xi \in \mathbb{R}$ (this is the classical Calderòn equation), and equation (1.20) is

$$f = \sum_{j\in\mathbb{Z}} \int_{\mathbb{R}} \langle f, \psi_{j,t}\rangle\, \psi_{j,t}\, dt, \tag{1.23}$$

where $\psi_{j,t}(x) := 2^{-j/2}\psi(2^{-j}x - t)$, $j \in \mathbb{Z}$, $t \in \mathbb{R}$. Thus, the classical orthonormal wavelet expansion

$$f = \sum_{j\in\mathbb{Z}} \sum_{k\in\mathbb{Z}} \langle f, \psi_{jk}\rangle\, \psi_{jk}$$

is a "discretization" of (1.23). This shows, by equation (1.4), that an orthonormal wavelet (in this classical case) is always a continuous wavelet satisfying property (1.23) for all $f \in L^2(\mathbb{R})$. This raises the question of how to "discretize" continuous wavelets associated with general dilations groups $\mathscr{D}$. We refer to [42] for more observations about this topic.

A variant of the affine group $\mathbf{A}_n$ (and the corresponding affine systems (1.19)) is obtained by considering the group G^* consisting of all pairs $(M,t) \in GL_n(\mathbb{R}) \times \mathbb{R}^n$ (endowed with the product topology) together with the group operation

$$(M,t)\cdot(M',t') = (MM', t + M't').$$

This operation is associated with the action $x \to Mx + t$ on $\mathbb{R}^n$. The *co-affine systems* associated with G^* are then defined as the elements

$$\left(V_{(M,t)}\,\psi\right)(x) = |\det M|^{-1/2}\psi(M^{-1}(x-t)) := \psi^*_{M,t}(x),$$

for $(M,t) \in G^*$ and $\psi \in L^2(\mathbb{R}^n)$. The corresponding expression in the frequency domain is:

$$\left(U_{(M,t)}\,\psi\right)^\wedge(\xi) = |\det M|^{1/2}\,\hat{\psi}(\xi M)\, e^{-2\pi i t}.$$

The left Haar measure, λ^*, for G^* is easily seen to satisfy $d\lambda^*(M,t) = |\det M|^{-1} d\mu(M)\, dt$, where μ is the left Haar measure for $\mathscr{D}$. Then the "co-affine" reproducing formula is

$$f = \int_G \langle f, \psi^*_{M,t}\rangle\, \psi^*_{M,t}\, d\lambda^*(M,t). \tag{1.24}$$

A straightforward calculation shows that (1.24) holds if and only if ψ satisfies condition (1.21). Thus, *ψ is a continuous affine wavelet if and only if it is a continuous co-affine wavelet.*

Notice that the situation observed above is different from the discrete case. In fact, consider the systems $\Psi = \{\psi_{j,k} = 2^{-j/2}\psi(2^{-j}\cdot - k) : j,k \in \mathbb{Z}\}$ and $\Psi^* =$

$\{\psi^*_{j,k} = 2^{-j/2}\psi(2^{-j}(\cdot - k)) :\ j,k \in \mathbb{Z}\}$. A simple calculation shows that

$$\langle \psi^*_{j,k}, \psi_{-1,-1}\rangle = \langle \psi_{j,0}, \psi_{-1,2k-1}\rangle.$$

This shows that the co-affine systems cannot generate the space $L^2(\mathbb{R})$ if the corresponding affine system Ψ is an orthonormal basis for $L^2(\mathbb{R})$. In fact, the affine system Ψ is an orthonormal basis for $L^2(\mathbb{R})$ (in which case the right hand side of the above expression is zero) if and only if the co-affine system Ψ^* has a non-empty orthogonal complement.

1.3.1 Admissible groups

It is not difficult to show that there are dilation groups $\mathscr{D}$ for which one can find no functions ψ satisfying equation (1.21). In particular, if $\mathscr{D}$ is compact, there are no associated functions ψ that satisfy this condition. For example, let $\mathscr{D} = SO(2)$ and suppose that there is a function $\psi \in L^2(\mathbb{R}^2)$ satisfying (1.21). Notice that, in this case, using polar coordinates equation (1.21) can be expressed as

$$\int_0^{2\pi} |\hat{\psi}(re^{i\phi}e^{i\theta})|^2 \,\frac{d\theta}{2\pi} = 1,$$

for a.e. $\xi = r\,e^{i\phi}$. Multiplying both sides of the equality by $r > 0$ and integrating with respect to $r \in [0,\infty)$, we obtain:

$$\begin{aligned}
\infty &= \int_0^\infty r\,dr \\
&= \int_0^\infty r \int_0^{2\pi} |\hat{\psi}(re^{i\phi}e^{i\theta})|^2 \,\frac{d\theta}{2\pi}\,dr \\
&= \int_0^\infty r \int_0^{2\pi} |\hat{\psi}(re^{i\theta})|^2 \,\frac{d\theta}{2\pi}\,dr \\
&= \|\psi\|^2 < \infty.
\end{aligned}$$

This is clearly a contradiction and, thus, there is no ψ satisfying (1.21). In this situation, we say that the group $SO(2)$ is not *admissible*. That a general compact $\mathscr{D} \subset GL_n(\mathbb{R})$ is not admissible is not much harder to prove.

The observation above leads to the question: what are the groups $\mathscr{D}$ that are admissible? Our result on admissibility involves the notion of *ε–stabilizer* of $x \in \mathbb{R}^n$, which is defined as the set

$$\mathscr{D}_x^\varepsilon = \{M \in \mathscr{D} : |xM - x| \le \varepsilon\},$$

for each $\varepsilon > 0$. The set $\mathscr{D}_x := D_x^0 = \{M \in \mathscr{D} : xM = x\}$ is called the *stabilizer* of x. The *modular function* Δ, on $\mathscr{D}$, defined by the property

$$\mu(EM) = \Delta(M)\,\mu(E)$$

for all μ–measurable $E \subset \mathscr{D}$ and $M \in \mathscr{D}$, also plays an important role in the following basic result about admissible dilation groups.

Theorem 5. *(a) If $\mathscr{D}$ is admissible, then $\Delta \neq |\det M|$ and the stabilizer of x is compact for a.e. $x \in \mathbb{R}^n$.*

(b) If $\Delta \neq |\det M|$ and for a.e. $x \in \mathbb{R}^n$ there exists an $\varepsilon > 0$ such that the ε–stabilizer of x is compact, then $\mathscr{D}$ is admissible.

The proof of Theorem 5 is rather involved and can be found in [33]. Even though Theorem 5 "just fails" to be a characterization of admissibility, still it is quite useful for determining admissibility or non-admissibility of particular groups $\mathscr{D}$. For example, if $\mathscr{D}$ is compact, then $\Delta = |\det M| = 1$ and, thus, it cannot be admissible. Another example where Theorem 5 can be used effectively is the case where $\mathscr{D}$ is a one-parameter group. Namely, let $\mathscr{D} = \{M_t = e^{tL} : t \in \mathbb{R}\}$, where L is a real $n \times n$ matrix. Then $\mathscr{D}$ is admissible if and only if $\mathrm{trace}(L) \neq 0$. Indeed, since $\det M_t = e^{t\,\mathrm{trace}(L)}$ and $\mathscr{D}$ is Abelian, it follows that the modular function, Δ, is identically 1. Thus, when $\mathrm{trace}(L) \neq 0$, we have that $\det M_t \neq 1 = \Delta$ and $\mathscr{D}$ is admissible.

1.3.2 Wave Packet Systems

In [5], Córdoba and Fefferman introduced "wave packets" as those families of functions obtained by applying certain collections of dilations, modulations and translations to the Gaussian function. More generally, we will describe as "wave packet systems" any collections of functions which are obtained by applying a combination of dilations, modulations and translations to a finite family of function in $L^2(\mathbb{R}^n)$. For $\Psi = \{\psi^\ell : 1 \leq \ell \leq L\} \subset L^2(\mathbb{R}^n)$, where $L \in \mathbb{N}$, and $S \subset GL_n(\mathbb{R}) \times \mathbb{R}^n$, the *continuous wave packet system* with respect to S that is generated by Ψ is the collection

$$\mathscr{W}\mathscr{P}_S(\Psi) = \{D_A M_\nu T_y \psi^\ell : (A,\nu) \in S, y \in \mathbb{R}^n, 1 \leq \ell \leq L\}, \tag{1.25}$$

where M_ν is the modulation operator defined at the beginning of Section 1.2.

Let

$$G = \{U = c D_A M_\nu T_y : c \in \mathbb{C}, |c| = 1, (A,\nu,y) \in GL_n(\mathbb{R}) \times \mathbb{R}^n \times \mathbb{R}^n\}.$$

G is a subgroup of the unitary operators on $L^2(\mathbb{R}^n)$ which is preserved by the action of the mapping $U \to \widehat{U}$, where $\widehat{U} f = (U f)^\wedge$.

In the definition (1.25), we considered the map $(A,\nu,y) \to U^{(0)}_{(A,\nu,y)} = D_A M_\nu T_y$, which is a one-to-one mapping from $S \times \mathbb{R}^n$ into the group G. By changing the order of the operators, we can also define the following one-to-one mappings from $S \times \mathbb{R}^n$ into G:

$$U^{(1)}_{(A,\nu,y)} = D_A T_y M_\nu$$

$$
\begin{aligned}
U^{(2)}_{(A,v,y)} &= T_y D_A M_v \\
U^{(3)}_{(A,v,y)} &= M_v D_A T_y \\
U^{(4)}_{(A,v,y)} &= T_y M_v D_A \\
U^{(5)}_{(A,v,y)} &= M_v T_y D_A.
\end{aligned}
$$

Hence, we can generate alternate continuous wave packet systems, $\mathscr{W}\mathscr{P}^{(i)}_S(\Psi)$, by replacing $U^{(0)}_{(A,v,y)}$ with $U^{(i)}_{(A,v,y)}$, for $1 \le i \le 5$. The systems $\mathscr{W}\mathscr{P}^{(0)}_S(\Psi)$ and $\mathscr{W}\mathscr{P}^{(1)}_S(\Psi)$ are equivalent in the sense that one is a Parseval frame if and only if the other one is a Parseval frame (in fact, by the commutativity relations of translations and modulations, they only differ by a unimodular scalar factor). The same is true for $\mathscr{W}\mathscr{P}^{(4)}_S(\Psi)$ and $\mathscr{W}\mathscr{P}^{(5)}_S(\Psi)$. The other systems, on the other hand, have substantial differences.

Each subgroup $U^{(i)}_{(A,v,y)}$, , $i = 0, \ldots, 5$, is associated to a continuous wave packet system generated by $\Psi \subset L^2(\mathbb{R}^n)$. We can characterize those Ψ for which we have Parseval frames:

$$
\sum_{\ell=1}^{L} \int_{S\times\mathbb{R}^n} \left| \langle f, U^{(i)}_{(A,v,y)\psi^\ell} \rangle \right|^2 d\lambda(A,v)\, dy = \|f\|_2^2
$$

for all $f \in L^2(\mathbb{R}^n)$, where λ is a measure on S. Such a characterization is an extension of Theorem 4, and is given by an analog of equality (1.21). Explicitly, we have the result

Theorem 6. *Let $\Psi = \{\psi^\ell : 1 \le \ell \le L\} \subset L^2(\mathbb{R}^n)$. The systems $\mathscr{W}\mathscr{P}^{(i)}_S(\Psi)$, $i = 0, \ldots, 5$ are continuous Parseval frame wave packet systems with respect to (S, λ), for $L^2(\mathbb{R}^n)$, if and only if*

$$
\Delta^{(i)}_{\Psi}(\xi) = 1, \text{ for a.e. } \xi \in \mathbb{R}^n,
$$

where

$$
\begin{aligned}
\Delta^{(1)}_{\Psi}(\xi) = \Delta^{(0)}_{\Psi}(\xi) &= \sum_{\ell=1}^{L} \int_S |\hat{\psi}^\ell(\xi A^{-1} - v)|^2\, d\lambda(A,v); \\
\Delta^{(2)}_{\Psi}(\xi) &= \sum_{\ell=1}^{L} \int_S |\hat{\psi}^\ell(\xi A^{-1} - v)|^2\, |\det A|^{-1}\, d\lambda(A,v); \\
\Delta^{(3)}_{\Psi}(\xi) &= \sum_{\ell=1}^{L} \int_S |\hat{\psi}^\ell((\xi - v)A^{-1})|^2\, d\lambda(A,v); \\
\Delta^{(4)}_{\Psi}(\xi) = \Delta^{(5)}_{\Psi}(\xi) &= \sum_{\ell=1}^{L} \int_S |\hat{\psi}^\ell((\xi - v)A^{-1})|^2\, |\det A|^{-1}\, d\lambda(A,v).
\end{aligned}
$$

Exercises

1. Show that Theorem 4 is valid for functions on subspaces of $L^2(\mathbb{R}^n)$ of the form

$$L^2(V)^\vee = \{f \in L^2(\mathbb{R}^n) : \operatorname{supp} \hat{f} \subset V\}.$$

1.4 Affine Systems with Composite Dilations

To describe the class of systems which will be considered in this section, it will be useful to begin with one example in $L^2(\mathbb{R}^2)$.

Let $A = \begin{pmatrix} 2 & 0 \\ 0 & \varepsilon \end{pmatrix}$, where $\varepsilon \neq 0$, $B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $G = \{(B^j, k) : j \in \mathbb{Z}, k \in \mathbb{Z}^2\}$. Then G is a group with group multiplication:

$$(B^\ell, m)(B^j, k) = (B^{\ell+j}, k + B^{-j} m). \tag{1.26}$$

In particular, we have $(B^j, k)^{-1} = (B^{-j}, -B^j k)$. The multiplication (1.26) is consistent with the operation that maps $x \to B^j(x+k)$ of $\mathbb{R}^2$ into $\mathbb{R}^2$. Let π be the unitary representation of G, acting on $L^2(\mathbb{R}^2)$ which is defined by

$$\Big(\pi(B^j, k) f\Big)(x) = f((B^j, k)^{-1} x) = f(B^{-j} x - k) = \Big(D_B^j T_k f\Big)(x), \tag{1.27}$$

for $f \in L^2(\mathbb{R}^2)$. Notice that $\det B^j = 1$. The observation that

$$(D_B^\ell T_m)(D_B^j T_k) = (D_B^{\ell+j} T_{k+B^{-j} m}),$$

where $\ell, j \in \mathbb{Z}$, $k, m \in \mathbb{Z}^2$, shows how the group operation (1.26) is associated with the unitary representation (1.27).

Let $S_0 = \{\xi = (\xi_1, \xi_2) \in \widehat{\mathbb{R}}^2 : |\xi_1| \leq 1\}$ and define

$$V_0 = L^2(S_0)^\vee = \{f \in L^2(\mathbb{R}^n) : \operatorname{supp} \hat{f} \subset S_0\}.$$

Since, for all $j \in \mathbb{Z}$ and $k \in \mathbb{Z}^2$, we have[2]

$$\Big(\pi(B^j, k) f\Big)^\wedge(\xi) = \Big(D_B^j T_k f\Big)^\wedge(\xi) = e^{-2\pi i \xi B^j k} \hat{f}(\xi B^j),$$

and $\xi B^j = (\xi_1, \xi_2) B^j = (\xi_1, \xi_2 + j\xi_1)$, then the action of B^j maps the vertical strip domain S_0 into itself and, thus, the space V_0 is invariant under the action of $\pi(B^j, k)$. The same invariance property holds for the vertical strips

$$S_i = S_0 A^i = \{\xi = (\xi_1, \xi_2) \in \widehat{\mathbb{R}}^2 : |\xi_1| \leq 2^i\},$$

[2] Recall that, according to the notation introduced in Section 1.2, in the frequency domain, the matrices B^j multiply row vectors on the right.

$i \in \mathbb{Z}$, and, as a consequence, the spaces $V_i = L^2(S_i)^\vee$ are also invariant under the action of the operators $\pi(B^j,k)$.

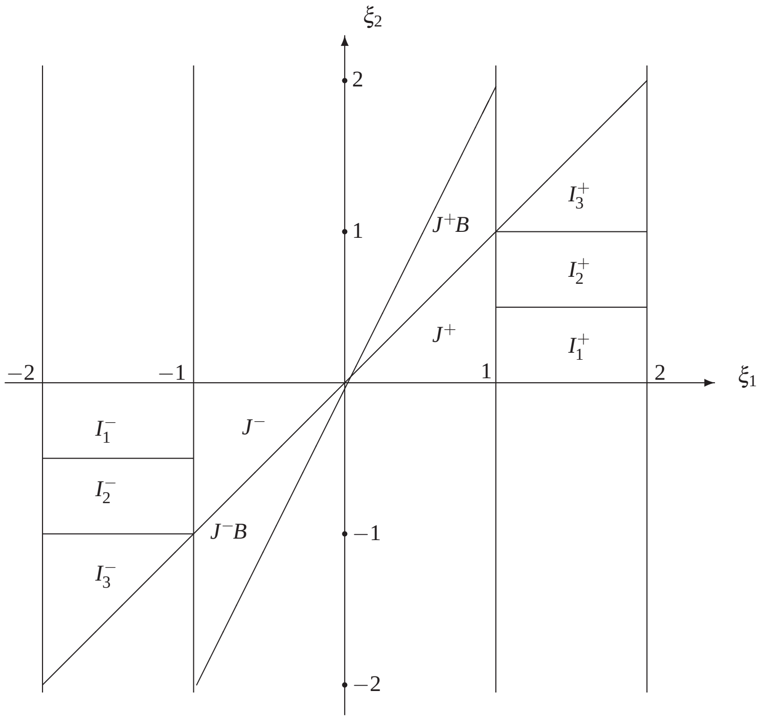


Fig. 1.1 Example of ON *AB*–MRA. The sets $\{J^+B^j, J^-B^j : j \in \mathbb{Z}\}$ form a disjoint partition of S_0.

The spaces $\{V_i\}_{i\in\mathbb{Z}}$ defined above satisfy the basic MRA properties:

1. $V_i \subset V_{i+1}$, $i \in \mathbb{Z}$;
2. $D_A^{-i} V_0 = V_i$;
3. $\bigcap_{i\in\mathbb{Z}} V_i = \{0\}$;
4. $\overline{\bigcup_{i\in\mathbb{Z}} V_i} = L^2(\mathbb{R}^n)$.

The complete definition of an MRA includes the assumption that V_0 is generated by the integer translates of a $\phi \in V_0$, called the *scaling function*, and that these translates $\{T_k\phi : k \in \mathbb{Z}^2\}$ are an orthonormal basis of V_0. In some cases, there are more than one scaling function.

The situation here is a bit different and the *scaling* property is replaced by an analogous property. Namely, consider $\widehat{V}_0 = L^2(S_0)$ and let $\hat{\phi} = \chi_J$, where $J = J^+ \cup J^-$, J^+ is the triangle with vertices $(0,0)$, $(1,0)$, $(1,1)$ and J^- is the triangle with vertices $(0,0)$, $(-1,0)$, $(-1,-1)$. The sets JB^j, $j \in \mathbb{Z}$, form a partition of S_0; that is, $S_0 = \bigcup_{j\in\mathbb{Z}} JB^j$, except for the set of points $\{(0,\xi_2) : \xi_2 \neq 0\}$, which is, however, a set of measure 0. The set J has measure 1 and the collection $\{e^{-2\pi ik\xi}\chi_J : k \in \mathbb{Z}^2\}$ is easily seen to be an ON basis of $L^2(J)$. Since

$$\left(e^{-2\pi ik\cdot}\,\chi_J(\cdot):\right)^{\vee}(x) = (T_k\phi)(x) = \phi(x-k),$$

these last functions form an ON basis of $L^2(J)^{\vee}$. It follows that $\{D_{B^j}\,T_k\,\phi : k\in\mathbb{Z}^2\}$ is an ON basis of $L^2(J\,B^j)^{\vee}$, for each $j\in\mathbb{Z}^2$. Hence, the set

$$\left\{D_{B^j}\,T_k\,\phi : j\in\mathbb{Z}, k\in\mathbb{Z}^2\right\} = \left\{T_k\,D_{B^j}\,\phi : j\in\mathbb{Z}, k\in\mathbb{Z}^2\right\}$$

is an ON basis of V_0. The sets J^+, J^-, as well as the other sets used in this construction are illustrated in Figure 1.1.

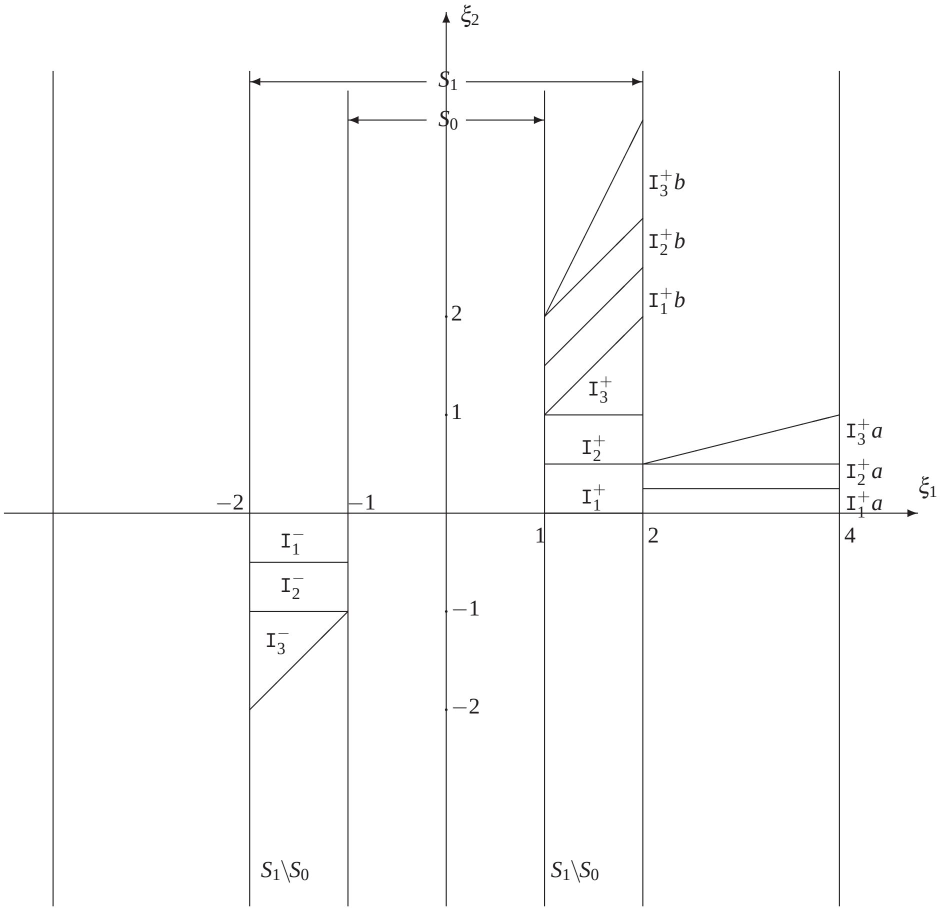


Fig. 1.2 Example of orthonormal *AB* wavelet.

Thus, the “complete” definition of the MRA, introduced above, adds to (1)–(4) the property:

(5) V_0 is generated by a "scaling function" ϕ, in the sense that $\{T_k D_{B^j} \phi : j \in \mathbb{Z}, k \in \mathbb{Z}^2\}$ is an ON basis of V_0.

Let G_B be the group $\{B^j : j \in \mathbb{Z}\}$; this is equivalent to the dilation group $\{D_{B^j} : j \in \mathbb{Z}\}$. Then $G = \{(B^j, k) : j \in \mathbb{Z}, k \in \mathbb{Z}^2\}$ is the semidirect product of G_B and $\mathbb{Z}^2$, denoted by $G_B \ltimes \mathbb{Z}^2$. This shows that the *shift–invariance* of the traditional MRA is replaced by a notion of $G_B \ltimes \mathbb{Z}^n$*–invariance*, that is, the space V_0 is invariant with respect to both integer translations and G_B dilations.

We shall now show how the MRA we just introduced can be used to construct a wavelet–like basis of $L^2(\mathbb{R}^2)$. We begin by constructing an ON basis of W_0, defined to be the orthogonal complement of V_0 in V_1, that is, $V_1 = V_0 \oplus W_0$. It will be convenient to work in the frequency domain. We have that $\widehat{V}_1 = \widehat{V}_0 \oplus \widehat{W}_0$ and, consequently, $\widehat{W}_0 = L^2(R_0)$, where $R_0 = S_1 \setminus S_0 = \{\xi = (\xi_1, \xi_2) \in \widehat{\mathbb{R}}^2 : 1 < |\xi_1| \le 2\}$. We define the following subsets of $R_0 = S_1 \setminus S_0$:

$$I_1 = I_1^+ \cup I_1^-, I_2 = I_2^+ \cup I_2^-, I_3 = I_3^+ \cup I_3^-,$$

where

$$\begin{aligned} I_1^+ &= \{\xi = (\xi_1, \xi_2) \in \widehat{\mathbb{R}}^2 : 1 < \xi_1 \le 2, 0 \le \xi_2 < 1/2\}, \\ I_2^+ &= \{\xi = (\xi_1, \xi_2) \in \widehat{\mathbb{R}}^2 : 1 < \xi_1 \le 2, 1/2 \le \xi_2 < 1\}, \\ I_3^+ &= \{\xi = (\xi_1, \xi_2) \in \widehat{\mathbb{R}}^2 : 1 < \xi_1 \le 2, 1 \le \xi_2 < \xi_1\}, \end{aligned}$$

and $I_\ell^- = \{\xi \in \widehat{\mathbb{R}}^2 : -\xi \in I_\ell^+\}$, $\ell = 1, 2, 3$. These sets are illustrated in Figures 1.1 and 1.2. Observe that each set I_ℓ is a *fundamental domain* for $\mathbb{Z}^2$: the functions $\{e^{2\pi i \xi k} : k \in \mathbb{Z}^2\}$, restricted to I_ℓ, form an ON basis for $L^2(I_\ell)$, $\ell = 1, 2, 3$. We then define ψ^ℓ, $\ell = 1, 2, 3$ by setting $\hat{\psi}^\ell = \chi_{I_\ell}$, $\ell = 1, 2, 3$. It follows from the observations about the sets $\{I_\ell\}$ that the collection

$$\{e^{2\pi i \xi k} \, \hat{\psi}^\ell(\xi) : k \in \mathbb{Z}^2\}$$

is an orthonormal basis of $L^2(I_\ell)$, $\ell = 1, 2, 3$. A simple direct calculation shows that the sets $\{I_\ell \, b^j : j \in \mathbb{Z}, \ell = 1, 2, 3\}$ are a partition of R_0, that is,

$$\bigcup_{\ell=1}^{3} \bigcup_{j \in \mathbb{Z}} I_\ell B^j = R_0,$$

where the union is disjoint. As a consequence, the collection

$$\{e^{2\pi i \xi k} \, \hat{\psi}^\ell(\xi B^j) : k \in \mathbb{Z}^2, j \in \mathbb{Z}, \ell = 1, 2, 3\} \tag{1.28}$$

is an orthonormal basis of $L^2(R_0)$ and, thus, by taking the inverse Fourier transform of (1.28), we have that

$$\{\pi(B^j, k) \, \psi^\ell : k \in \mathbb{Z}^2, j \in \mathbb{Z}, \ell = 1, 2, 3\} \tag{1.29}$$

is an orthonormal basis of $W_0 = L^2(R_0)^\vee$. Notice that, since, for each $j \in \mathbb{Z}$ fixed, B^j maps $\mathbb{Z}^2$ into itself, the collection $\{e^{2\pi i \xi B^j k} : k \in \mathbb{Z}^2\}$ is equal to the collection $\{e^{2\pi i \xi k} : k \in \mathbb{Z}^2\}$.

It is clear that, by applying the dilations D_{A^i}, $i \in \mathbb{Z}$, to the system (1.29), we obtain an ON basis of $L^2(R_i)^\vee$, where

$$R_i = R_0 A^i = \{\xi = (\xi_1, \xi_2) \in \widehat{\mathbb{R}}^2 : 2^i < |\xi_1| \le 2^{1+i}\}.$$

Furthermore, we have that $\bigcup_{i\in\mathbb{Z}} R_i = \widehat{\mathbb{R}}^2$, where the union is disjoint, and, hence we can write $L^2(\mathbb{R}^2) = \bigoplus_{i\in\mathbb{Z}} W_i$. Hence, by combining the observations above, it follows that the collection

$$\{D_{A^i} D_{B^j} T_k \psi^\ell : k \in \mathbb{Z}^2, i, j \in \mathbb{Z}, \ell = 1,2,3\} \tag{1.30}$$

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

1.4.1 Affine System with Composite Dilations

The construction given above is a particular example of a general class of affine-like systems called *affine system with composite dilations*, which have the form:

$$\mathscr{A}_{AB}(\Psi) = \{D_A D_B T_k \psi^\ell : A \in G_A, B \in G_B, k \in \mathbb{Z}^n, \ell = 1, \ldots, L\}, \tag{1.31}$$

where $\Psi \subset \{\psi^1, \ldots, \psi^\ell\} \in L^2(\mathbb{R}^n)$, $G_A \subset GL_n(\mathbb{R})$ (usually, $G_A = \{A^i : i \in \mathbb{Z}\}$, with A expanding or having some "expanding" property), and $G_B \subset GL_n(\mathbb{R})$ with $|\det B| = 1$. Later on, we will show that there are several examples of such systems that form ON bases of $L^2(\mathbb{R}^n)$ or, more generally, Parseval frames of $L^2(\mathbb{R}^n)$.

The roles played by the two families of dilations, G_A and G_B, in definition (1.31), are very different. The elements $A \in G_A$ dilate (at least in some direction), while the elements of G_B affect the geometry of the reproducing system $\mathscr{A}_{AB}(\Psi)$. In the example we worked out, $G_B = \{\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)^j : j \in \mathbb{Z}\}$ is the *shear group* and exhibits a "shear geometry", in which objects in the plane are stretched vertically without increasing their size (like the trapezoids in Figure 1.2). In Section 1.5, we will use this group and a construction similar to the one above to obtain the *shearlets*, whose geometrical properties are similar to the example above, and are, in addition, well-localized functions (i.e., they have rapid decay both in the space and frequency domains). They have similarities to the *curvelets* introduced by Candès and Donoho [2] and to the *contourlets* of Do and Vetterli [10]. However, their mathematical construction is simpler, since it derived from the structure of affine systems and, as a result, their development and applications are "more systematic" [22, 23].

As indicated by the example above, there is a special multiresolution analysis associated with the affine systems with composite dilations which is useful for constructing "composite wavelets". Let us give a proper definition of this new framework. Let G_B be a countable subset of $\widetilde{SL}_n(\mathbb{Z}) = \{B \in GL_n(\mathbb{R}) : |\det B| = 1\}$ and

$G_A = \{A^i : i \in \mathbb{Z}\}$, where $A \in GL_n(\mathbb{Z})$ (notice that A is an *integral* matrix). Also assume that A *normalizes* G_B, that is, $ABA^{-1} \in G_B$ for every $B \in G_B$, and that the quotient space $B/(ABA^{-1})$ is finite. Then the sequence $\{V_i\}_{i\in\mathbb{Z}}$ of closed subspaces of $L^2(\mathbb{R}^n)$ is an *AB-multiresolution analysis (AB-MRA)* if the following holds:

(i) $D_B T_k V_0 = V_0$, *for any* $B \in G_B$, $k \in \mathbb{Z}^n$,
(ii) *for each* $i \in \mathbb{Z}$, $V_i \subset V_{i+1}$,, *where* $V_i = D_A^{-i} V_0$,
(iii) $\bigcap V_i = \{0\}$ *and* $\overline{\bigcup V_i} = L^2(\mathbb{R}^n)$,
(iv) *there exists* $\phi \in L^2(\mathbb{R}^n)$ *such that* $\Phi_B = \{D_B T_k \phi : B \in G_B, k \in \mathbb{Z}^n\}$ *is a semi-orthogonal Parseval frame for* V_0, *that is,* Φ_B *is a Parseval frame for* V_0 *and, in addition,* $D_B T_k \phi \perp D_{B'} T_{k'} \phi$ *for any* $B \neq B'$, $B, B' \in G_B$, $k, k' \in \mathbb{Z}^n$.

The space V_0 is called an *AB scaling space* and the function ϕ is an *AB scaling function* for V_0. In addition, if Φ_B is an orthonormal basis for V_0, then ϕ is an *orthonormal AB scaling function.*

The number of generators L of an orthonormal MRA *AB*-wavelet is completely determined by the group $G = \{(B^j, k) : j \in \mathbb{Z}, k \in \mathbb{Z}^n\}$. Indeed we have the following simple fact:

Proposition 1. *Let G be a countable group and $u \to T_u$ be a unitary representation of G acting on a (separable) Hilbert space $\mathscr{H}$. Suppose $\Phi = \{\phi^1, \ldots, \phi^N\}$, $\Psi = \{\psi^1, \ldots, \psi^M\} \subset \mathscr{H}$, where $N, M \in \mathbb{N} \bigcup \{\infty\}$. If $\{T_u \phi^k : u \in G, 1 \le k \le N\}$ and $\{T_u \psi^i : u \in G, 1 \le i \le M\}$ are each orthonormal bases for $\mathscr{H}$, then $N = M$.*

Proof. It follows from the assumptions that, for each $1 \le k \le N$:

$$\|\phi^k\|^2 = \sum_{u\in G} \sum_{i=1}^{M} |\langle \phi^k, T_u \psi^i \rangle|^2.$$

Thus, by the properties of T_u, we have:

$$\begin{aligned} N = \sum_{k=1}^{N} \|\phi^k\|^2 &= \sum_{k=1}^{N} \sum_{u\in G} \sum_{i=1}^{M} |\langle \phi^k, T_u \psi^i \rangle|^2 \\ &= \sum_{i=1}^{M} \sum_{u\in G} \sum_{k=1}^{N} |\langle T_{u^{-1}} \phi^k, \psi^i \rangle|^2 \\ &= \sum_{i=1}^{M} \|\psi^i\|^2 = M. \quad \square \end{aligned}$$

Using Proposition 1, one obtains the following result which establishes the number of generators needed to obtain an orthonormal MRA *AB*-wavelet. □

Theorem 7. *Let $\Psi = \{\psi^1, \ldots, \psi^L\}$ be an orthonormal MRA AB-multiwavelet for $L^2(\mathbb{R}^n)$, and let $N = |B/ABA^{-1}|$ (= the order of the quotient group B/ABA^{-1}). Assume that $|\det A| \in \mathbf{N}$. Then $L = N|\det A| - 1$.*

The composite wavelet system $\mathscr{A}_{AB}(\Psi)$ have associated continuous multiwavelets. The simplest case is the one in which the translations are $\{T_y : y \in \mathbb{R}^n\}$. In this case, we have the reproducing formula corresponding to (1.20):

$$f = \sum_{\ell=1}^{L} \sum_{i,j\in\mathbb{Z}} \int_{\mathbb{R}^n} \langle f, D_{A^i} D_{B^j} T_y \psi^\ell \rangle D_{A^i} D_{B^j} T_y \psi^\ell \, dy, \tag{1.32}$$

for $f \in L^2(\mathbb{R}^n)$. As in Section 1.3, one can show that $\Psi = \{\psi^1, \dots, \psi^\ell\}$ satisfies (1.32) if and only if it satisfies the Calderòn equation

$$\sum_{\ell=1}^{L} \sum_{i,j\in\mathbb{Z}} |\hat{\psi}^\ell(\xi A^i B^j)| = 1 \quad \text{for a.e. } \xi \in \mathbb{R}^n.$$

Some more general examples of continuous composite wavelet system will be examined in Section 1.5.

1.4.2 Other Examples

There are several other examples of affine systems with composite dilations $\mathscr{A}_{AB}(\Psi)$ which form ON bases or Parseval frames.

In particular, the construction presented above in dimension $n = 2$ extends to the general n-dimensional setting. In this case, the shear group is given by $G_B = \{B^i : i \in \mathbb{Z}\}$, where $B \in GL_n(\mathbb{R})$ is characterized by the equality $(B - I_n)^2 = 0$, and I_n is the $n \times n$ identity matrix. We refer to [23] for more detail about these systems.

A different type of affine systems with composite dilations arises when G_B is a finite group. For example, let $G_B = \{\pm B_0, \pm B_1, \pm B_2, \pm B_3\}$ be the 8-element group consisting of the isometries of the square $[-1,1]^2$. Specifically: $B_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $B_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $B_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, $B_3 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$. Let U be the parallelogram with vertices $(0,0), (1,0), (2,1)$ and $(1,1)$ and $S_0 = \bigcup_{b\in B} U\, b$ (see the snowflake region in Figure 1.3). It is easy to verify that S_0 is B-invariant.

Let A be the quincunx matrix $\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$, and $S_i = S_0 A^i$, $i \in \mathbb{Z}$. Observe that A is expanding, $ABA^{-1} = B$ and $S_0 \subseteq S_0 A = S_1$. In particular, the region $S_1 \setminus S_0$ is the disjoint union $\bigcup_{b\in B} RB$, where the region R is the parallelogram illustrated in Figure 1.3. Thus, as in the case of the shear composite wavelet that we have described above, it follows that the system

$$\{D_A^i D_B T_k \psi : i \in \mathbb{Z}, B \in G_B, k \in \mathbb{Z}^2\}, \tag{1.33}$$

where $\hat{\psi} = \chi_R$, is an orthonormal basis for $L^2(\mathbb{R}^2)$.

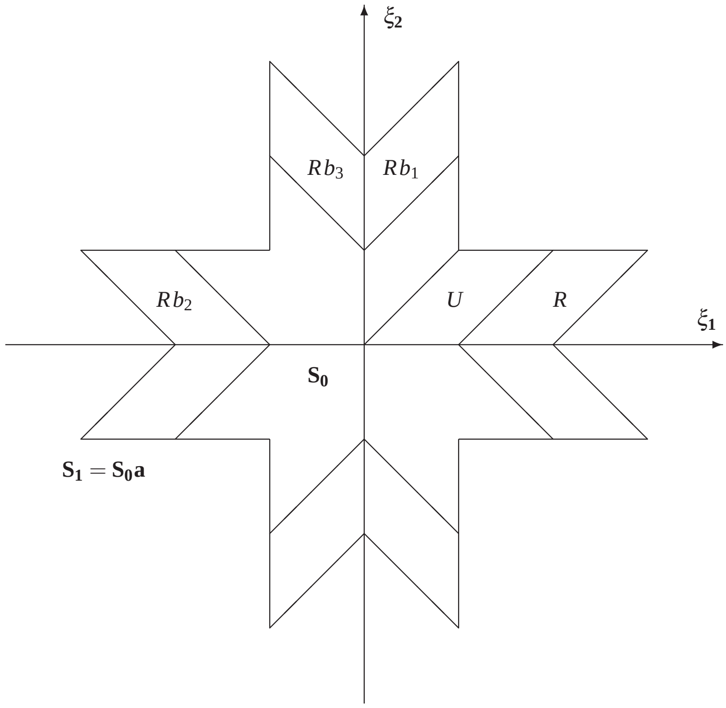


Fig. 1.3 Example of composite wavelet with finite group. $G_A = \{A^i : i \in \mathbb{Z}\}$, where A is the quincunx matrix, and G_B is the group of isometries of the square $[-1,1]^2$.

If the quincunx matrix A is replaced by the matrix $\tilde{A} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$, we obtain a different ON basis. Let B, U and S_i, $i \in \mathbb{Z}$, be defined as above. Also in this case, $\tilde{A}$ is expanding, $\tilde{A}B\tilde{A}^{-1} = B$ and $S_1 = S_0 a \supset S_0$. A direct computation shows that the region $S_1 \setminus S_0$ is the disjoint union $\bigcup_{B \in G_B} Rb$, where $R = R_1 \bigcup R_2 \bigcup R_3$ and the regions R_1, R_2, R_3 are illustrated in Figure 1.4. Observe that each of the regions R_1, R_2, R_3 is a fundamental domain. Thus, the system

$$\{D^i_{\tilde{A}} D_B T_k \psi^\ell : i \in \mathbb{Z}, B \in G_B, k \in \mathbb{Z}^2, \ell = 1, \ldots, 3\}, \tag{1.34}$$

where $\hat{\psi}^\ell = \chi_{R_\ell}$, $\ell = 1,2,3$, is an orthonormal basis for $L^2(\mathbb{R}^2)$.

Note that the system in the first example (equation (1.33)) was generated by a single function, while the second system (equation (1.34)) is generated by three functions ψ^1, ψ^2, ψ^3. This is consistent with Theorem 7. In fact, if B is a finite group, then $N = |B/ABA^{-1}| = 1$, and so, in this situation, the number of generators is $L = |\det A| - 1$. Thus, by Theorem 7, in the first example we obtain that the number of generators is $L = 1$ since A is the quincunx matrix and $\det A = 2$. In the second example, the number of generators is $L = 3$ since $\tilde{A} = 2I$ and $\det \tilde{A} = 4$. Fi-

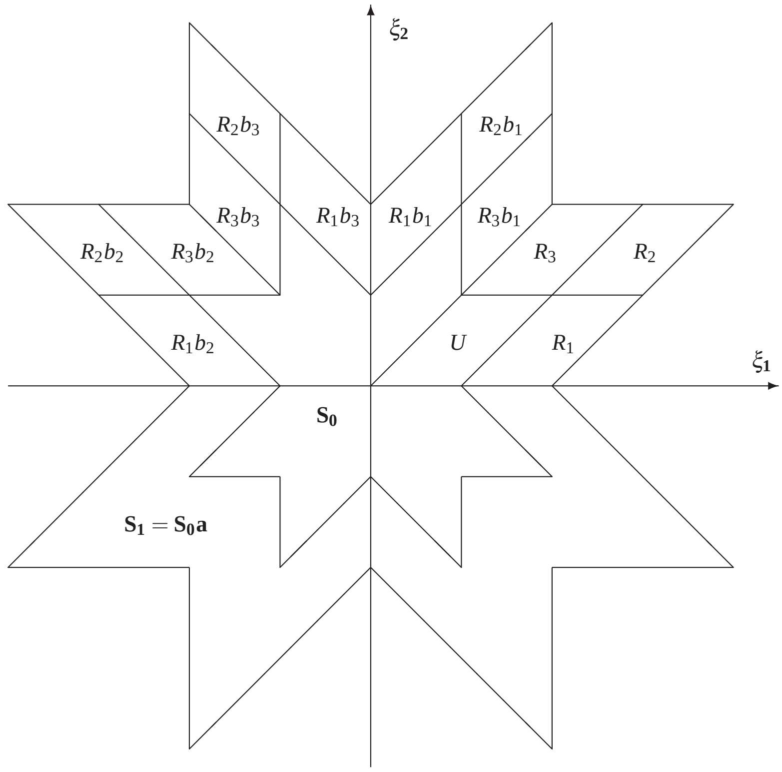


Fig. 1.4 Example of composite wavelet with finite group. $G_A = \{A^i : i \in \mathbb{Z}\}$, where $A = 2I$, and G_B is the group of isometries of the square $[-1,1]^2$.

nally, in the example at beginning of this Section, where G_B is the two-dimensional group of shear matrices and $G_A = \{A^i : i \in \mathbb{Z}\}$, with $A = \begin{pmatrix} 2 & A_{1,2} \\ 0 & A_{2,2} \end{pmatrix} \in GL_2(\mathbb{Z})$, a calculation shows that $|B/ABA^{-1}| = 2|A_{2,2}|^{-1}$ and, thus, the number of generators is $L = 2|A_{2,2}|^{-1}\, 2|A_{2,2}| - 1 = 3$.

In higher dimensions, the type of constructions we have just described extend by using the Coxeter group. These are finite groups (hence, their elements have determinant 1 in magnitude) generated by reflections through hyperplanes.

Other examples of composite wavelets, in dimension $n = 2$, are obtained, for each $\lambda > 1$ fixed, by considering the group

$$G_B = \{B_j = \begin{pmatrix} \lambda^j & 0 \\ 0 & \lambda^{-j} \end{pmatrix} : j \in \mathbb{Z}\},$$

and choosing G_A to be a group of expanding matrices; for example $G_A = \{A^i : i \in \mathbb{Z}\}$, where A is diagonal and $|\det A| > 1$. We refer to [23] for more detail about this construction.

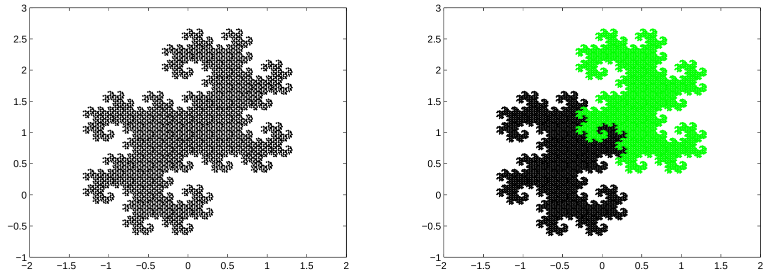


Fig. 1.5 On the left: the fractal set known as “Twin Dragon”. On the right: support of the 2-dimensional Haar wavelet ψ; $\psi = 1$ on the darker set, $\psi = -1$ on the lighter set.

All examples of composite wavelets presented so far are “direct” constructions in the frequency domain. let us now discuss a different class of composite wavelets in the “time domain”.

Perhaps the simplest dyadic-dilation wavelet in dimension $n = 1$ is the Haar wavelet. It is produced by the scaling function $\phi = \chi_{[0,1)}$ and is generated by the Haar function $\psi = \chi_{[0,1/2)} - \chi_{[1/2,1)}$. The *Haar ON basis* of $L^2(\mathbb{R})$ is the affine system

$$\{\psi_{i,k} = D_{2^i}\, T_k\, \psi : i,k \in \mathbb{Z}\}.$$

It is a natural question to ask what are the extensions of this compactly supported wavelet ψ in higher dimensions. For example, in dimension $n = 2$, consider the quincunx matrix $A_q = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ and the associated affine system

$$\{\psi_{i,k} = D_{A_q^i}\, T_k\, \psi : i \in \mathbb{Z}, k \in \mathbb{Z}^2\}. \tag{1.35}$$

Then, similarly to the one-dimensional Haar wavelet, one can find an MRA wavelet ψ produced by a scaling function ϕ that is the characteristic function of a compact set $Q \subset \mathbb{R}^2$ of area 1. However, the functions ϕ and ψ are not that simple. In fact, the scaling function ϕ is the characteristic function of a rather complicated fractal set known at the “twin dragon” and ψ is the difference of two similar characteristic functions (see Figure 1.5).

We can construct an affine system with composite dilations having the same expanding dilation group $G_A = \{A_q^i : i \in \mathbb{Z}\}$ and the same translations that does, however, generate a very simple Haar-type wavelet. For the group G_B, let us choose again the group of symmetries of the unit square given at the beginning of this section. Let R_0 be the triangle with vertices $(0,0)$, $(1/2,0)$, $(1/2,1/2)$ and $R_\ell = B_\ell R_0$, $\ell = 1,\ldots,7$ (see Figure 1.6). Then, for $\phi = 2\sqrt{2}\,\chi_{R_0}$, it follows that the system

$$\{D_{B_\ell}\, T_k\, \phi : \ell = 0,\ldots,7, k \in \mathbb{Z}^2\}$$

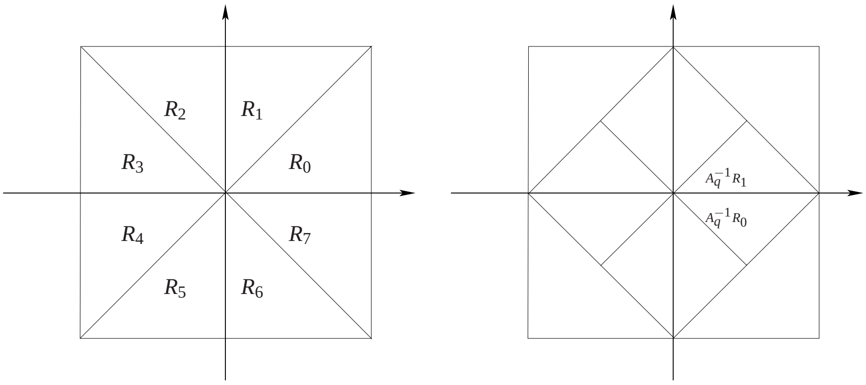


Fig. 1.6 Example of a composite wavelet with finite support.

is an ON basis for the space V_0, which is the closed linear span of the subspace of $L^2(\mathbb{R}^2)$ consisting of the functions which are constant on each $\mathbb{Z}^2$-translate of the triangles R_ℓ, $\ell = 0,1,\dots,7$. Let us now consider the spaces $V_i = D_{A_q^{-i}} V_0$, $i \in \mathbb{Z}$. Then one can verify that each space V_i is the closed linear span of the subspace of $L^2(\mathbb{R}^2)$ consisting of the functions which are constant on each $A_q^{-i}\mathbb{Z}^2$-translate of the triangles $A_q^{-i}R_\ell$, $\ell = 0,1,\dots,7$. Thus, $V_i \subset V_{i+1}$ for each $i \in \mathbb{Z}$, and the spaces $\{V_i\}$ form an *AB*–MRA, with ϕ as an *AB*–scaling function. We can now construct a simple Haar-like wavelet obtained from this *AB*–MRA. Specifically, let

$$R_0 = A_q^{-1}R_1 \cup \left[A_q^{-1}R_6 + \begin{pmatrix}1/2\\1/2\end{pmatrix}\right] = A_q^{-1}R_1 \cup A_q^{-1}\left[R_6 + \begin{pmatrix}0\\1\end{pmatrix}\right].$$

Thus, $\chi_{R_0} = \chi_{A_q^{-1}R_1} + \chi_{A_q^{-1}R_6 + \left(\begin{smallmatrix}1/2\\1/2\end{smallmatrix}\right)}$ or, equivalently,

$$\phi^{(0)}(x) = \phi^{(1)}(A_q x) + \phi^{(6)}(A_q x - \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)), \tag{1.36}$$

where $\phi^{(\ell)} = D_{B_\ell}\phi$, for $\ell = 0,1,\dots,7$. It is now easy to see that $\psi = \phi^{(1)}(A_q x) - \phi^{(6)}(A_q x - \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right))$ is the desired Haar–like *AB*–wavelet. The space V_0 is generated by applying the translations T_k, $k \in \mathbb{Z}^2$, to the scaling functions $\phi^{(\ell)} = D_{B_\ell}\phi$, $\ell = 0,1,\dots,7$. We see that this is the case by applying D_{B_ℓ} in equality (1.36); we we obtain:

$$\begin{aligned}
\phi^{(0)} &= \phi^{(1)}(A_q x) + \phi^{(6)}(A_q x - \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right))\\
\phi^{(1)} &= \phi^{(2)}(A_q x) + \phi^{(5)}(A_q x - \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right))\\
\phi^{(2)} &= \phi^{(3)}(A_q x) + \phi^{(0)}(A_q x - \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right))\\
\phi^{(3)} &= \phi^{(4)}(A_q x) + \phi^{(7)}(A_q x - \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right))\\
\phi^{(4)} &= \phi^{(5)}(A_q x) + \phi^{(2)}(A_q x - \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right))
\end{aligned}$$

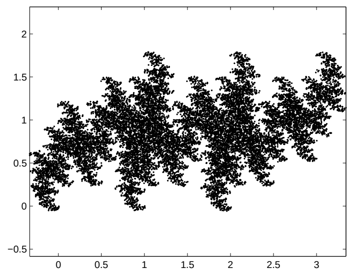
$$\begin{aligned}\phi^{(5)} &= \phi^{(6)}(A_q x) + \phi^{(1)}(A_q x - \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right))\\ \phi^{(6)} &= \phi^{(7)}(A_q x) + \phi^{(4)}(A_q x - \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right))\\ \phi^{(7)} &= \phi^{(0)}(A_q x) + \phi^{(3)}(A_q x - \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)).\end{aligned}$$

It follows that

$$\{D_{A_q^i} D_{B_\ell} T_k \psi : i \in \mathbb{Z}, \ell = 0,1,\ldots,7, k \in \mathbb{Z}^2\}$$

is an ON basis for $L^2(\mathbb{R}^2)$. This Haar-type *AB*-wavelet is clearly simpler that the twin dragon wavelet obtained above. We refer to [1, 29] for more information about this type of constructions.

Other complicated fractal wavelets appear in many situations. For example, if the dilation matrix A_q in the affine system (1.35) is replaced by $A_{q1} = \begin{pmatrix} 1 & \sqrt{3} \\ -\sqrt{3} & 1 \end{pmatrix}$ or $A_{q2} = \begin{pmatrix} 3/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 3/2 \end{pmatrix}$, then also in this case there is a compactly supported MRA wavelet generated by a (compactly supported) scaling function ϕ which is the characteristic function of a fractal set (see Figure 1.7).



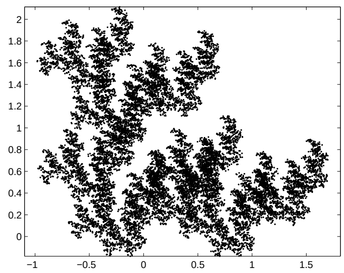


Fig. 1.7 The fractal sets associated with the MRA generated by the dilation matrices A_{q1} (on the left) and A_{q2} (on the right).

The construction given above, suggests that also in these cases one should be able to find an *AB*–MRA such that the associated compactly supported *AB*–wavelet has a simpler “non-fractal” support.

Exercises

1. Let $\psi_1 \in L^2(\mathbb{R})$ be a dyadic wavelet with supp $\hat{\psi}_1 \subset [-\frac{1}{2}, \frac{1}{2}]$ and $\psi_2 \in L^2(\mathbb{R})$ be such that supp $\hat{\psi}_1 \subset [-1, 1]$ and

$$\sum_{k\in\mathbb{Z}} |\hat{\psi}_2(\omega+k)|^2 = 1 \text{ for a.e. } \omega \in \mathbb{R}.$$

For $\xi = (\xi_1, \xi_2) \in \mathbb{R}^2$, let ψ be defined by $\hat{\psi}(\xi) = \hat{\psi}_1(\xi_1)\,\hat{\psi}_2(\frac{\xi_2}{\xi_1})$. Show that the affine system $\{D_A^i D_B^j T_k \psi : i, j \in \mathbb{Z}, k \in \mathbb{Z}^2\}$, where $A = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, is a Parseval frame for $L^2(\mathbb{R}^2)$.

1.5 Continuous Shearlet Transform

An important class of subgroups of the affine group $\mathbf{A}_2$ (which was described in Section 1.3) is obtained by considering

$$G = \{(M,t) : M \in \mathscr{D}_\alpha, t \in \mathbb{R}^2\}, \tag{1.37}$$

where, for each $0 < \alpha < 1$, $\mathscr{D}_\alpha \subset GL_2(\mathbb{R})$ is the set of matrices:

$$\mathscr{D}_\alpha = \left\{ M = M_{as} = \begin{pmatrix} a & -a^\alpha s \\ & \\ 0 & a^\alpha \end{pmatrix}, \quad a > 0, s \in \mathbb{R} \right\}.$$

The matrices M_{as} can be factorized as $M_{as} = B_s A_a$, where

$$B_s = \begin{pmatrix} 1 & -s \\ 0 & 1 \end{pmatrix} \quad A_a = \begin{pmatrix} a & 0 \\ 0 & a^\alpha \end{pmatrix}. \tag{1.38}$$

The matrix B_s is called a *shear matrix* and, for each $s \in \mathbb{R}$, is a non-expanding matrix ($\det B_s = 1$, for each s). The matrix A_a is an anisotropic dilation matrix, that is, the dilation rate is different in the x and y directions. In particular, if $\alpha = 1/2$ the matrix A_a produces *parabolic scaling* since $f(A_a x) = f\left(A_a \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right)$ leaves invariant the parabola $x_1 = x_2^2$. Thus, the action associated with the dilation group $\mathscr{D}_\alpha$ can be interpreted as the superposition of anisotropic dilation and shear transformation.

Using Theorem 4 from Section 1.3, we can establish simple conditions on the function ψ so that it will satisfy the Calderón reproducing formula (1.20) with respect to G. This is done in the following proposition.

Proposition 2. *Let G be given by (1.37) and, for $\xi = (\xi_1, \xi_2) \in \mathbb{R}^2$, $\xi_2 \neq 0$, let ψ be given by*

$$\hat{\psi}(\xi) = \hat{\psi}(\xi_1, \xi_2) = \hat{\psi}_1(\xi_1)\,\hat{\psi}_2(\tfrac{\xi_2}{\xi_1}).$$

Suppose that:

(i) $\psi_1 \in L^2(\mathbb{R})$ *satisfies*

$$\int_0^\infty |\hat{\psi}_1(a\xi)|^2 \frac{da}{a^{2\alpha}} = 1 \quad \textit{for a.e. } \xi \in \mathbb{R};$$

(ii) $\|\psi_2\|_{L^2} = 1$.

Then ψ satisfies (1.20) and, hence, is a continuous wavelet with respect to G.

Proof. A direct computation shows that $(\xi_1,\xi_2)M = (a\,\xi_1, a^\alpha(\xi_2 - s\xi_1))$. Also notice the element of the left Haar measure for $\mathscr{D}$ is $d\mu(M_{as}) = \frac{da}{|\det M_{as}|}\,ds$. Hence the admissibility condition (1.21) for ψ is

$$\Delta(\psi)(\xi) = \int_{\mathbb{R}}\int_{\mathbb{R}^+} |\hat{\psi}_1(a\,\xi_1)|^2\,|\hat{\psi}_2(a^{\alpha-1}(\tfrac{\xi_2}{\xi_1} - s))|^2\,\frac{da}{a^{1+\alpha}}\,ds = 1. \tag{1.39}$$

for a.e. $(\xi_1,\xi_2) \in \mathbb{R}^2$. Thus, by Theorem 4, to show that ψ is a continuous wavelet with respect to G, it is sufficient to show that (1.39) is satisfied. Using the assumption on ψ_1 and ψ_2, we have:

$$\begin{aligned}
\Delta(\psi)(\xi) &= \int_{\mathbb{R}}\int_{\mathbb{R}^+} |\hat{\psi}_1(a\,\xi_1)|^2\,|\hat{\psi}_2(a^{\alpha-1}(\tfrac{\xi_2}{\xi_1} - s))|^2\,\frac{da}{a^{1+\alpha}}\,ds\\
&= \int_{\mathbb{R}^+} |\hat{\psi}_1(a\,\xi_1)|^2 \left(\int_{\mathbb{R}} |\hat{\psi}_2(a^{\alpha-1}\,\tfrac{\xi_2}{\xi_1} - s)|^2\,ds\right)\frac{da}{a^{2\alpha}}\\
&= \int_{\mathbb{R}^+} |\hat{\psi}_1(a\,\xi_1)|^2\,\frac{da}{a^{2\alpha}} = 1,
\end{aligned}$$

for a.e. $\xi = (\xi_1,\xi_2) \in \mathbb{R}^2$. This shows that equality (1.39) is satisfied. □

In the following, to distinguish a continuous wavelets ψ associated with this particular group G from other continuous wavelets, we will refer to such a function as a *continuous shearlet*. Hence, for each $0 < \alpha < 1$, the *continuous shearlet transform* is the mapping

$$f \to \{\mathscr{S}^\alpha_\psi f(a,s,t) = \langle f, \psi_{ast}\rangle : a > 0, s \in \mathbb{R}, t \in \mathbb{R}^2\},$$

where the analyzing elements:

$$\{\psi_{ast}(x) = |\det M_{as}|^{-\frac{1}{2}}\,\psi(M_{as}^{-1}(x-t)) :, a > 0, s \in \mathbb{R}, t \in \mathbb{R}^2\},$$

with $M_{as} \in \mathscr{D}_\alpha$, form a *continuous shearlet system*. Notice that, according to the terminology introduced in Section 1.3, the elements $\{\psi_{ast}\}$ are co-affine functions.

An useful variant of the continuous shearlet transform is obtained by restricting the range of the shear variable s associated with the shearing matrices B_s to a finite interval. Namely, for $0 < \alpha < 1$, let us redefine

$$\mathscr{D}^{(h)}_\alpha = \left\{ M_{as} = \begin{pmatrix} a & -a^\alpha s \\ 0 & a^\alpha \end{pmatrix}, \quad 0 < a \le \frac{1}{4}, -\frac{3}{2} \le s \le \frac{3}{2}, \right\},$$

and

$$G^{(h)} = \{(M,t) : M \in \mathscr{D}^{(h)}_\alpha, t \in \mathbb{R}^2\}.$$

Also, consider the subspace of $L^2(\mathbb{R}^2)$ given by $L^2(C_h)^\vee = \{f \in L^2(\mathbb{R}^2) : \operatorname{supp} \hat{f} \subset C_h\}$, where C_h is the “horizontal cone” in the frequency plane:

$$C_h = \{(\xi_1, \xi_2) \in \mathbb{R}^2 : |\xi_1| \geq 1 \text{ and } |\tfrac{\xi_2}{\xi_1}| \leq 1\}.$$

Hence we can show that, by slightly modifying the assumptions of Proposition 2, the function ψ is be a continuous shearlet for the subspace $L^2(C_h)^\vee$.

Proposition 3. *For $\xi = (\xi_1, \xi_2) \in \mathbb{R}^2$, $\xi_2 \neq 0$, let ψ be given by*

$$\hat{\psi}(\xi) = \hat{\psi}(\xi_1, \xi_2) = \hat{\psi}_1(\xi_1)\, \hat{\psi}_2(\tfrac{\xi_2}{\xi_1}),$$

where:

(i) $\psi_1 \in L^2(\mathbb{R})$ satisfies

$$\int_0^\infty |\hat{\psi}_1(a\xi)|^2 \frac{da}{a^{2\alpha}} = 1 \quad \textit{for a.e. } \xi \in \mathbb{R},$$

and supp $\hat{\psi}_1 \subset [-2, -\frac{1}{2}] \cup [\frac{1}{2}, 2]$;

(ii) $\|\psi_2\|_{L^2} = 1$ and supp $\hat{\psi}_2 \subset [-1, 1]$.

Then ψ satisfies (1.24). That is, for all $f \in L^2(C_h)^\vee$,

$$f(x) = \int_{\mathbb{R}^2} \int_{-\frac{3}{2}}^{\frac{3}{2}} \int_0^{\frac{1}{4}} \langle f, \psi_{ast} \rangle\, \psi_{ast}(x)\, \frac{da}{a^{2+2\alpha}}\, ds\, dt,$$

with convergence in the L^2 sense.

There are several examples of functions ψ_1 and ψ_2 satisfying the assumptions of Proposition 2 and Proposition 3. In particular, we can choose ψ_1, ψ_2 such that $\hat{\psi}_1$, $\hat{\psi}_2 \in C_0^\infty$ and we will make this assumption in the following. We refer to [15, 23] for the construction of these functions.

If the assumptions of Proposition 3 are satisfied, we say that the set

$$\Psi^{(h)} = \{\psi_{ast} : 0 < a \leq \tfrac{1}{4}, -\tfrac{3}{2} \leq s \leq \tfrac{3}{2}, t \in \mathbb{R}^2\}$$

is a *continuous shearlet system* for $L^2(C_h)^\vee$ and that the corresponding mapping from $f \in L^2(C_h)^\vee$ into $\mathscr{S}_\psi^{(h),\alpha} f(a,s,t) = \langle f, \psi_{ast} \rangle$ is the *continuous shearlet transform* on $L^2(C_h)^\vee$.

In the frequency domain, an element of the shearlet system ψ_{ast} has the form:

$$\hat{\psi}_{ast}(\xi_1, \xi_2) = a^{\frac{1+\alpha}{2}}\, \hat{\psi}_1(a\,\xi_1)\, \hat{\psi}_2(a^{\alpha-1}(\tfrac{\xi_2}{\xi_1} - s))\, e^{-2\pi i \xi t}.$$

As a result, each function $\hat{\psi}_{ast}$ has support:

$$\operatorname{supp} \hat{\psi}_{ast} \subset \{(\xi_1, \xi_2) : \xi_1 \in [-\tfrac{2}{a}, -\tfrac{1}{2a}] \cup [\tfrac{1}{2a}, \tfrac{2}{a}], |\tfrac{\xi_2}{\xi_1} - s| \leq a^{1-\alpha}\}.$$

As illustrated in Figure 1.8, the frequency support is a pair of trapezoids, symmetric with respect to the origin, oriented along a line of slope s. The support becomes increasingly elongated as $a \to 0$.

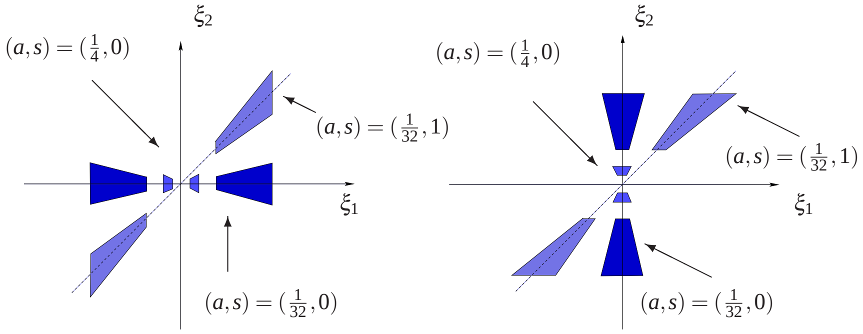


Fig. 1.8 Frequency support of the horizontal shearlets (left) and vertical shearlets (right) for different values of a and s.

As shown by Proposition 3, the continuous shearlet transform $\mathscr{S}_\psi^{(h),\alpha}$ provides a reproducing formula only for functions in a proper subspace of $L^2(\mathbb{R}^2)$. To extend the transform to all $f \in L^2(\mathbb{R}^2)$, we introduce a similar transform to deal with the functions supported on the "vertical cone":

$$C^{(v)} = \{(\xi_1, \xi_2) \in \mathbb{R}^2 : |\xi_2| \geq 1 \text{ and } |\tfrac{\xi_2}{\xi_1}| > 1\}.$$

Specifically, let

$$\hat{\psi}^{(v)}(\xi) = \hat{\psi}^{(v)}(\xi_1, \xi_2) = \hat{\psi}_1(\xi_2)\, \hat{\psi}_2(\tfrac{\xi_1}{\xi_2}),$$

where $\hat{\psi}_1$, $\hat{\psi}_2$ satisfy the same assumptions as in Proposition 3, and consider the dilation group

$$\mathscr{D}_\alpha^{(v)} = \{N_{as} = \begin{pmatrix} a^\alpha & 0 \\ -a^\alpha s & a \end{pmatrix} : 0 < a \leq 1, -\frac{3}{2} \leq s \leq \frac{3}{2}, t \in \mathbb{R}^2\}.$$

Then it is easy to verify that the set

$$\Psi^{(v)} = \{\psi_{ast}^{(v)} : 0 < a \leq \tfrac{1}{4}, -\tfrac{3}{2} \leq s \leq \tfrac{3}{2}, t \in \mathbb{R}^2\},$$

where $\psi_{ast}^{(v)} = |\det N_{as}|^{-\frac{1}{2}} \psi^{(v)}(N_{as}^{-1}(x-t))$, is a continuous shearlet system for $L^2(C^{(v)})^\vee$. The corresponding transform $\mathscr{S}_\psi^{(v),\alpha} f(a,s,t) = \langle f, \psi_{ast}^{(v)} \rangle$ is the continuous shearlet transform on $L^2(C^{(v)})^\vee$. Finally, by introducing an appropriate window function W, we can represent the functions with frequency support on the set $[-2,2]^2$ as

$$f = \int_{\mathbb{R}^2} \langle f, W_t \rangle W_t \, dt,$$

where $W_t(x) = W(x-t)$. As a result, any function $f \in L^2(\mathbb{R}^2)$ can be reproduced with respect of the full shearlet system, which consists of the horizontal shearlet system $\Psi^{(h)}$, the vertical shearlet system $\Psi^{(v)}$, and the collection of coarse-scale isotropic functions $\{W_t : t \in \mathbb{R}^2\}$. We refer to [31] for more details about this representation. For our purposes, it is only the behavior of the fine-scale shearlets that matters. Indeed, in the following, we will apply the continuous shearlet transforms $\mathscr{S}_\psi^{(h),\alpha}$ and $\mathscr{S}_\psi^{(v),\alpha}$, at fine scales ($a \to 0$), to resolve and precisely describe the boundaries of certain planar regions. Hence, it will be convenient to re-define shearlet transform, at "fine-scales", as follows. For $0 < a \leq 1/4$, $s \in \mathbb{R}$, $t \in \mathbb{R}^2$, the *(fine-scale) continuous shearlet transform* is the mapping from $f \in L^2(\mathbb{R}^2 \setminus [-2,2]^2)^\vee$ into $\mathscr{S}_\psi f$ which is defined by:

$$\mathscr{S}_\psi^\alpha f(a,s,t) = \begin{cases} \mathscr{S}_\psi^{(h),\alpha}(a,s,t) & \text{if } |s| \leq 1 \\ \mathscr{S}_\psi^{(v),\alpha}(a,\frac{1}{s},t) & \text{if } |s| > 1. \end{cases}$$

1.5.1 Edge Analysis using the Shearlet Tranform

One remarkable property of the continuous shearlet transform is its ability to provide a very precise characterization of the set of singularities of functions and distributions. Indeed, let f be a function on $\mathbb{R}^2$ consisting of several smooth regions Ω_n, $n = 1, \ldots, N$, separated by piecewise smooth boundaries $\gamma_n = \partial \Omega_n$:

$$f(x) = \sum_{n=1}^{N} f_n(x) \chi_{\Omega_n}(x),$$

where each function f_n is smooth. Then the continuous shearlet transform $\mathscr{S}_\psi^\alpha f(a,s,t)$ will signal both the location and orientation of the boundaries through its asymptotic decay at fine scales. In fact, $\mathscr{S}_\psi^\alpha f(a,s,t)$ will exhibit fast asymptotic decay $a \to 0$ for all (s,t), except for the values of t on the boundary curves γ_n and for the values of s associated with the normal orientation to the γ_n at t.

The study of these objects is motivated by image applications, where f is used to model an image, and the curves γ_n are the edges of the image f. We will show that the shearlet framework provides a very effective method for the detection and analysis of edges. This is a fundamental problem in many applications from computer vision image processing.

To illustrate how the shearlet transform can be employed to characterize the geometry of edges, let us consider the case where f is simply the characteristic function of a bounded subset of $\mathbb{R}^2$. Also, to simplify the presentation, we will only present the situation where $\alpha = 1/2$ and use the simplified notation $\mathscr{S}_\psi = \mathscr{S}_\psi^{\frac{1}{2}}$. The

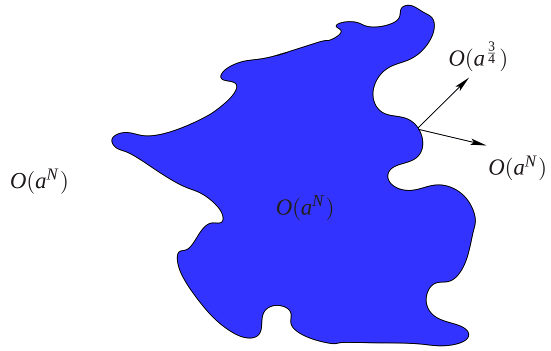


Fig. 1.9 Asymptotic decay of the continuous shearlet transform of the $B(x) = \chi_D(x)$. On the boundary ∂D, for normal orientation, the shearlet transform decays as $O(a^{\frac{3}{4}})$. For all other values of (t,s), the decay is as fast as $O(a^N)$, for any $N \in \mathbb{N}$.

more general case where $\alpha \in (0,1)$ the continuous shearlet transform $\mathscr{S}_\psi^\alpha$ is similar and details can be found in [21].

We then have the following result from [21].

Theorem 8. *Let $D \subset \mathbb{R}^2$ be a bounded region in $\mathbb{R}^2$ and suppose that the boundary curve $\gamma = \partial D$ is a simple C^3 regular curve. Denote $B = \chi_D$. If $t = t_0 \in \gamma$, and $s_0 = \tan\theta_0$ where θ_0 is the angle corresponding to the normal orientation to γ at t_0, then*

$$\lim_{a\to 0^+} a^{-\frac{3}{4}} \mathscr{S}_\psi B(a,s_0,t_0) \neq 0.$$

If $t = t_0 \in \gamma$ and $s \neq \tan\theta_0$, or if $t \notin \gamma$, then

$$\lim_{a\to 0^+} a^{-\beta} \mathscr{S}_\psi B(a,s,t) = 0, \quad \textit{for all } \beta > 0.$$

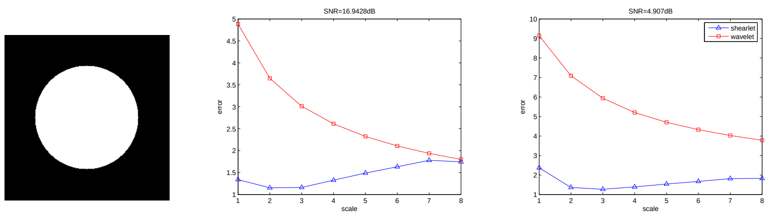
This shows that the continuous shearlet transform $\mathscr{S}_\psi B(a,s,t)$ has “slow” decay only for $t = t_0$ on γ when the value of the shear variable s corresponds exactly to the normal orientation to γ at t_0. For all other values of t and s the decay is fast. This behavior is illustrated in Figure 1.9.

Theorem 8 can be generalized to the situation where the boundary curve γ is piecewise smooth and contains finitely many corner points. Also in this case, the continuous shearlet transform provides a precise description of the geometry of the boundary curve through its asymptotic decay at fine scales. In particular, at the corner points, the asymptotic decay at fine scales is the slowest for values of s corresponding to the normal directions (notice that there are two of them). We refer the interested reader to [20] for a detailed discussion of the shearlet analysis of regions with piecewise smooth boundaries. We also refer to [21, 31] for other related results, including the situation where f is not simply the union of characteristic functions of sets.

Finally, we recall that the shearlet transform shares some of the features described above with *continuous curvelet transform*, another directional multiscale transform introduced by Candès and Donoho in [3]. Even if a result like Theorem 8 is not known for the curvelet transform, other results in [3] indicate that also the curvelet transform is able to capture the geometry of singularities in $\mathbb{R}^2$ through its asymptotic decay at fine scales Notice that, unlike the shearlet transform, the curvelet transform is not directly associated with an affine group.

1.5.2 A Shearlet Approach to Edge Analysis and Detection

Taking advantage of the properties of the continuous shearlet transform described above, an efficient numerical algorithm for edge detection was designed by one of the authors and his collaborators [43, 44]. The shearlet approach adapts several ideas from the well-known *Wavelet Modulus Maxima* method of Hwang, Mallat and Zhong [35, 36], where the edge points of an image f are identified as the locations corresponding to the local maxima of the magnitude of the continuous wavelet transform of f. Recall that, at a single scale, this wavelet-based method is indeed equivalent to the *Canny Edge Detector*, which is a standard edge detection algorithm [8].



(a) (b) (c)

Fig. 1.10 (a) Test image. (b-c) Comparison of the average error in angle estimation using the wavelet method versus the shearlet method, as a function of the scale, with different noise levels; (b) PSNR = 16.9 dB, (c) PSNR = 4.9 dB (Courtesy of Sheng Yi).

As shown above, one main feature of the continuous shearlet transform is its superior directional selectivity with respect to wavelets and other traditional methods. This property plays a very important role in the design of the edge detection algorithm. In fact, one major task in edge detection, is to accurately identify the edges of an image in the presence of noise and, to perform this task, both the location and the orientation of edge points have to be estimated from a noisy image. In the usual Wavelet Modulus Maxima approach, the edge orientation of an image f, at the location t, is estimated by looking at the ratio of the vertical over the horizontal components of $W_\psi f(a,t)$, the wavelet transform of f. However, this approach is

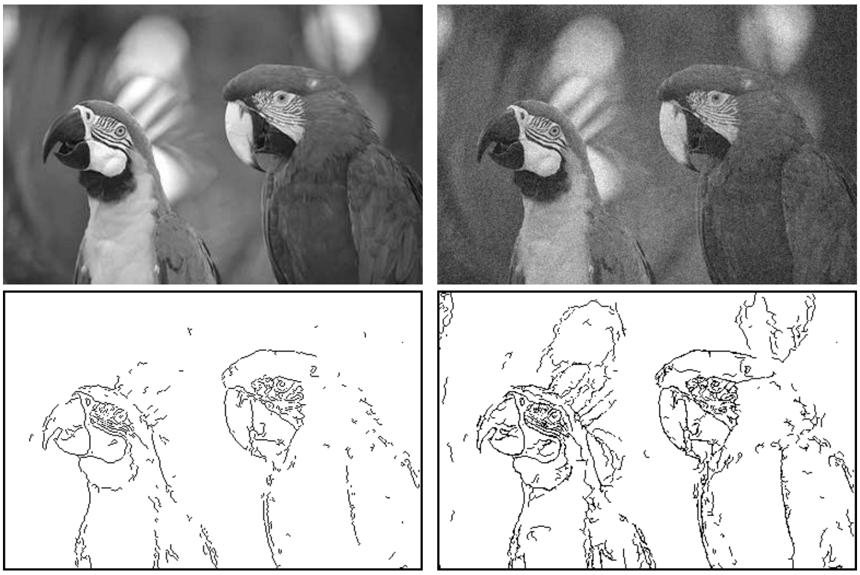


Fig. 1.11 Comparison of edge detection using shearlet-based method versus wavelet-based method. From top left, clockwise: Original image, noisy image (PSNR= 24.59 dB), shearlet result, and wavelet result (Courtesy of Glenn Easley).

not very accurate when dealing with discrete data. The advantage of the continuous shearlet transform is that, by representing the image as a function of scale, location and orientation, the directional information is directly available. A number of tests conducted in [43, 44] show indeed that a shearlet-based approach provides a very accurate estimate of the edge orientation of a noisy images; this method significantly outperforms the wavelet-based approach. A typical numerical experiment is illustrated in Figure 1.10, where the test image is the characteristic function of a disc. This figure displays the average angular error in the estimate of the edges orientation, as a function of the scale a. The average angle error is defined by

$$\frac{1}{|E|}\cdot\sum_{t\in E}|\hat{\theta}(t)-\theta(t)|,$$

where E is the set of edge points, θ is the exact angle and $\hat{\theta}$ the estimated angle. The average angle error is indicated for both shearlet- and wavelet-based methods, in presence of additive Gaussian noise. As the figure shows, the shearlet approach significantly outperforms the wavelet method, especially at finer scales, and is extremely robust to noise.

Using these properties, a very competitive algorithm for edge detection was developed in [44] and a representative numerical tests is illustrated in Figure 1.11. We refer to [43, 44] for details about these algorithms and for additional numerical demonstrations.

1.5.3 Discrete Shearlet System

By sampling the continuous shearlet transform

$$f \to \mathscr{S}_\psi f(a,s,t) = \langle f, \psi_{a,s,t} \rangle$$

on an appropriate discrete set of the scaling, shear, and translation parameters $(a,s,t) \in \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^2$, it is possible to obtain a frame or even a Parseval frame for $L^2(\mathbb{R}^2)$. Notice that, as above, we will only consider the case $\mathscr{S}_\psi = \mathscr{S}_\psi^{\frac{1}{2}}$.

To construct the discrete shearlet system (see [30] for more detail), we start by choosing a discrete set of scales $\{a_j\}_{j\in\mathbb{Z}} \subset \mathbb{R}^+$; next, for each fixed j, we choose the shear parameters $\{s_{j,\ell}\}_{\ell\in\mathbb{Z}} \subset \mathbb{R}$ so that the directionality of the representation is allowed to change with the scale. Finally, to provide a "uniform covering" of $\mathbb{R}^2$, we allow the location parameter to describe a different grid depending on j on ℓ; hence we let $t_{j,\ell,k} = B_{s_{j,\ell}} A_{a_j} k$, $k \in \mathbb{Z}^2$, where the matrices B_s, for $s \in \mathbb{R}$, and A_a, for $a > 0$ are given by 1.38). Observing that

$$T_{\{B_{s_{j,\ell}} A_{a_j} k\}} D_{B_{s_{j,\ell}} A_{a_j}} = D_{B_{s_{j,\ell}} A_{a_j}} T_k,$$

we obtain the discrete system

$$\{\psi_{j,\ell,k} = D_{B_{s_{j,\ell}} A_{a_j}} T_k \, \psi : \; j,\ell \in \mathbb{Z}, k \in \mathbb{Z}^2\}.$$

In particular, we will set $a_j = 2^{2j}$, $s_{j,\ell} = \ell \sqrt{a_j} = \ell \, 2^j$. Thus, observing that $B_{\ell 2^j} A_{2^{2j}} = A_{2^{2j}} B_\ell$, we finally obtain the *discrete shearlet system*

$$\{\psi_{j,\ell,k} = D_{A_{4^j}} \, D_{B_\ell} \, T_k \, \psi : \; j,\ell \in \mathbb{Z}, k \in \mathbb{Z}^2\}. \tag{1.40}$$

Notice that (1.40) is an example of the affine systems with composite dilations (1.31), described in Section 1.4. More specifically, the discrete shearlet system obtained above is similar to the "shearlet-like" system (1.30). Unlike the system (1.30), however, whose elements where characteristic functions of sets in the frequency domain, we will show that in this case we obtain a system of well-localized functions.

To do that, we will adapt some ideas from the continuous case. Namely, for any $\xi = (\xi_1, \xi_2) \in \widehat{\mathbb{R}}^2$, $\xi_1 \neq 0$, let

$$\hat{\psi}^{(h)}(\xi) = \hat{\psi}^{(h)}(\xi_1, \xi_2) = \hat{\psi}_1(\xi_1) \, \hat{\psi}_2 \left(\frac{\xi_2}{\xi_1} \right),$$

where $\hat{\psi}_1, \hat{\psi}_2 \in C^\infty(\widehat{\mathbb{R}})$, supp $\hat{\psi}_1 \subset [-\frac{1}{2}, -\frac{1}{16}] \cup [\frac{1}{16}, \frac{1}{2}]$ and supp $\hat{\psi}_2 \subset [-1,1]$. This implies that $\hat{\psi}^{(h)}$ is a compactly-supported C^∞ function with support contained in $[-\frac{1}{2}, \frac{1}{2}]^2$. In addition, we assume that

$$\sum_{j\geq 0} |\hat{\psi}_1(2^{-2j}\omega)|^2 = 1 \quad \text{ for } |\omega| \geq \frac{1}{8}, \tag{1.41}$$

and, for each $j \geq 0$,

$$\sum_{\ell=-2^j}^{2^j-1} |\hat{\psi}_2(2^j \omega - \ell)|^2 = 1 \quad \text{for } |\omega| \leq 1. \tag{1.42}$$

From the conditions on the support of $\hat{\psi}_1$ and $\hat{\psi}_2$, one can easily deduce that the functions $\psi_{j,\ell,k}$ have frequency support contained in the set

$$\{(\xi_1,\xi_2) : \xi_1 \in [-2^{2j-1}, -2^{2j-4}] \cup [2^{2j-4}, 2^{2j-1}], |\tfrac{\xi_2}{\xi_1} + \ell 2^{-j}| \leq 2^{-j}\}.$$

Thus, each element $\hat{\psi}_{j,\ell,k}$ is supported on a pair of trapezoids of approximate size $2^{2j} \times 2^j$, oriented along lines of slope $\ell 2^{-j}$ (see Figure 1.12(b)).

From equations (1.41) and (1.42) it follows that the functions $\{\hat{\psi}_{j,\ell,k}\}$ form a tiling of the set

$$\mathscr{D}_h = \{(\xi_1,\xi_2) \in \widehat{\mathbb{R}}^2 : |\xi_1| \geq \tfrac{1}{8}, |\frac{\xi_2}{\xi_1}| \leq 1\}.$$

Indeed, for $(\xi_1,\xi_2) \in \mathscr{D}_h$

$$\sum_{j\geq 0} \sum_{\ell=-2^j}^{2^j-1} |\hat{\psi}^{(h)}(\xi A_4^{-j} B_1^{-\ell})|^2 = \sum_{j\geq 0} \sum_{\ell=-2^j}^{2^j-1} |\hat{\psi}_1(2^{-2j}\xi_1)|^2 \, |\hat{\psi}_2(2^j \frac{\xi_2}{\xi_1} - \ell)|^2 = 1. \tag{1.43}$$

An illustration of this frequency tiling is shown in Figure 1.12(a).

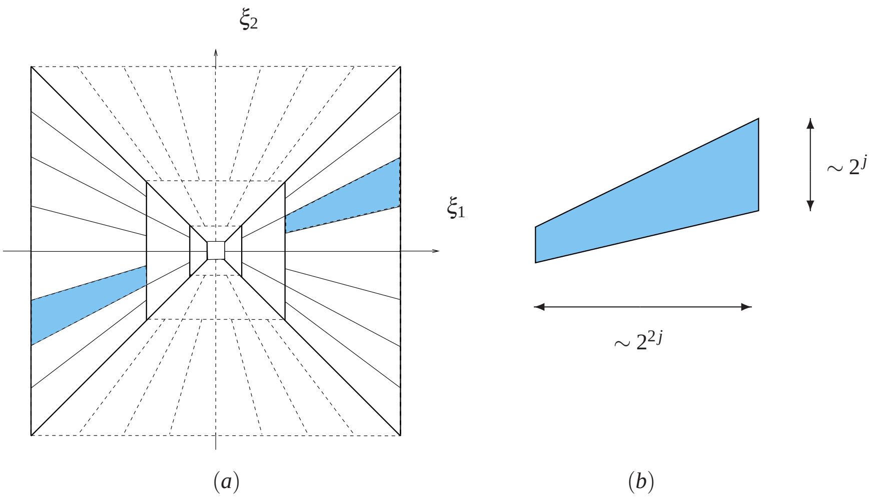


Fig. 1.12 (a) The tiling of the frequency plane $\widehat{\mathbb{R}}^2$ induced by the shearlets. The tiling of $\mathscr{D}_h$ is illustrated in solid line, the tiling of $\mathscr{D}_v$ is in dashed line. (b) The frequency support of a shearlet $\psi_{j,\ell,k}$ satisfies parabolic scaling. The figure shows only the support for $\xi_1 > 0$; the other half of the support, for $\xi_1 < 0$, is symmetrical.

Letting $L^2(\mathscr{D}_h)^\vee = \{f \in L^2(\mathbb{R}^2) : \text{supp}\,\hat{f} \subset \mathscr{D}_h\}$, property (1.43) and the fact that $\hat{\psi}^{(h)}$ is supported inside $[-\frac{1}{2},\frac{1}{2}]^2$ imply that the discrete shearlet system

$$\Psi_d^{(h)} = \{\psi_{j,\ell,k} : j \geq 0, -2^j \leq \ell \leq 2^j - 1, k \in \mathbb{Z}^2\}$$

is a Parseval frame for $L^2(\mathscr{D}_h)^\vee$. Similarly, we can construct a Parseval frame for $L^2(\mathscr{D}_v)^\vee$, where $\mathscr{D}_v$ is the vertical cone $\mathscr{D}_v = \{(\xi_1,\xi_2) \in \widehat{\mathbb{R}}^2 : |\xi_2| \geq \frac{1}{8}, |\frac{\xi_1}{\xi_2}| \leq 1\}$. Specifically, let

$$\tilde{A} = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}, \quad \tilde{B} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix},$$

and $\psi^{(v)}$ be given by

$$\hat{\psi}^{(v)}(\xi) = \hat{\psi}^{(v)}(\xi_1,\xi_2) = \hat{\psi}_1(\xi_2)\,\hat{\psi}_2\left(\frac{\xi_1}{\xi_2}\right).$$

Then the collection

$$\Psi_d^{(v)} = \{\psi_{j,\ell,k}^{(v)} : j \geq 0, -2^j \leq \ell \leq 2^j - 1, k \in \mathbb{Z}^2\}$$

where $\psi_{j,\ell,k}^{(v)} = D_{\tilde{A}}^j D_{\tilde{B}}^\ell T_k \psi^{(v)}$ is a Parseval frame for $L^2(\mathscr{D}_v)^\vee$.

Finally, let $\hat{\varphi} \in C_0^\infty(\mathbb{R}^2)$ be chosen to satisfy

$$\begin{aligned} &|\hat{\varphi}(\xi)|^2 + \sum_{j\geq 0}\sum_{\ell=-2^j}^{2^j-1} |\hat{\psi}^{(h)}(\xi A_4^{-j} B_1^{-\ell})|^2\,\chi_{\mathscr{D}_h}(\xi) \\ &+ \sum_{j\geq 0}\sum_{\ell=-2^j}^{2^j-1} |\hat{\psi}^{(v)}(\xi \tilde{A}^{-j}\tilde{B}^{-\ell})|^2\,\chi_{\mathscr{D}_v}(\xi) = 1, \quad \text{for } \xi \in \widehat{\mathbb{R}}^2, \end{aligned}$$

where $\chi_{\mathscr{D}}$ is the indicator function of the set $\mathscr{D}$. This implies that $\text{supp}\,\hat{\varphi} \subset [-\frac{1}{8},\frac{1}{8}]^2$, $|\hat{\varphi}(\xi)| = 1$ for $\xi \in [-\frac{1}{16},\frac{1}{16}]^2$, and the collection $\{\varphi_k : k \in \mathbb{Z}^2\}$ defined by $\varphi_k(x) = \varphi(x-k)$ is a Parseval frame for $L^2([-\frac{1}{16},\frac{1}{16}]^2)^\vee$.

Thus, letting $\widehat{\tilde{\psi}}_{j,\ell,k}^{(\omega)}(\xi) = \widehat{\psi}_{j,\ell,k}^{(\omega)}(\xi)\,\chi_{\mathscr{D}_\omega}(\xi)$, for $\omega = h$ or $\omega = v$, we have the following result.

Theorem 9. *The discrete shearlet system*

$$\begin{aligned} \{\varphi_k : k \in \mathbb{Z}^2\} \bigcup \{\tilde{\psi}_{j,\ell,k}^{(\omega)}(x) : j \geq 0, \ell = -2^j, 2^j - 1, k \in \mathbb{Z}^2, \omega = h,v\} \\ \bigcup \{\psi_{j,\ell,k}^{(\omega)}(x) : j \geq 0, -2^j + 1 \leq \ell \leq 2^j - 2, k \in \mathbb{Z}^2, \omega = h,v\}, \end{aligned}$$

is a Parseval frame for $L^2(\mathbb{R}^2)$.

The "corner" elements $\tilde{\psi}_{j,\ell,k}^{(\omega)}(x)$, $\ell = -2^j, 2^j - 1$, are simply obtained by truncation on the cones $\chi_{\mathscr{D}_\omega}$ in the frequency domain. Notice that the corner elements in the horizontal cone $\mathscr{D}_v$ match nicely with those in the vertical cone $\mathscr{D}_h$. We refer to [13, 23] for additional details on this construction.

1.5.4 Optimal Representations using Shearlets

One major feature of shearlet system is that, if f is a compactly supported function which is C^2 away from a C^2 curve, then the sequence of discrete shearlet coefficients $\{\langle f, \psi_{j,\ell,k}\rangle\}$ has (essentially) optimally fast decay. To make this more precise, let f_N^S be the N–term approximation of f obtained from the N largest coefficients of its shearlet expansion, namely

$$f_N^S = \sum_{\mu \in I_N} \langle f, \psi_\mu \rangle \, \psi_\mu,$$

where $I_N \subset M$ is the set of indices corresponding to the N largest entries of the sequence $\{|\langle f, \psi_\mu \rangle|^2 : \mu \in M\}$. Also, we follow [2] and introduce $STAR^2(A)$, a class of indicator functions of sets B with C^2 boundaries ∂B. In polar coordinates, let $\rho(\theta) : [0, 2\pi) \to [0,1]^2$ be a radius function and define B by $x \in B$ if and only if $|x| \le \rho(\theta)$. In particular, the boundary ∂B is given by the curve in $\mathbb{R}^2$:

$$\beta(\theta) = \begin{pmatrix} \rho(\theta)\cos(\theta) \\ \rho(\theta)\sin(\theta) \end{pmatrix}. \tag{1.44}$$

The class of boundaries of interest to us are defined by

$$\sup |\rho''(\theta)| \le A, \quad \rho \le \rho_0 < 1. \tag{1.45}$$

We say that a set $B \in STAR^2(A)$ if $B \subset [0,1]^2$ and B is a translate of a set obeying (1.44) and (1.45). Finally, we define the set $\mathscr{E}^2(A)$ of *functions which are C^2 away from a C^2 edge* as the collection of functions of the form

$$f = f_0 + f_1 \, \chi_B,$$

where $f_0, f_1 \in C_0^2([0,1]^2)$, $B \in STAR^2(A)$ and $\|f\|_{C^2} = \sum_{|\alpha| \le 2} \|D^\alpha f\|_\infty \le 1$. We can now state the following result from [19].

Theorem 10. *Let $f \in \mathscr{E}^2(A)$ and f_N^S be the approximation to f defined above. Then*

$$\|f - f_N^S\|_2^2 \le C N^{-2} (\log N)^3.$$

Notice that the approximation error of shearlet systems significantly outperforms wavelets, in which case the approximation error $\|f - f_N^W\|_2^2$ decays at most as fast as $O(N^{-1})$ [34], where f_N^W is the N–term approximation of f obtained from the N largest coefficients in the wavelet expansion. Indeed, the shearlet representation is essentially optimal for the kind of functions considered here, since the optimal theoretical approximation rate (cf. [9]) satisfies

$$\|f - f_N\|_2^2 \asymp N^{-2}, \qquad N \to \infty.$$

Only the curvelet system of Candès and Donoho are known to satisfy similar approximation properties [2]. However, the curvelet construction has a number of important differences, including the fact that the curvelet system is not associated with a fixed translation lattice, and, unlike the shearlet system, is not an affine-like system since it is not generated from the action of a family of operators on a single or finite family of functions.

The optimal sparsity of the shearlet system plays a fundamental role in a number of applications. For example, the shearlet system can be applied to provide a sparse representation of Fourier Integral Operators, a very important class of operators which appear in problems from Partial Differential Equations [17, 18]. Another class of applications comes from image processing, where the sparsity of the shearlet representation is closely related to the ability to efficiently separate the relevant features of an image from noise. A number of results in this direction are described in [11, 12, 13].

Exercises

1. Prove Proposition 3 by modifying the argument of Proposition 3.
2. Let ψ be a Schwarz class function and $\mathscr{S}_\psi$ be the fine-scale continuous shearlet transform (for $\alpha = 1/2$), as defined in this section. Show that, for any $s \in \mathbb{R}$, the continuous shearlet transform of the Dirac delta distribution satisfies:

$$\mathscr{S}_\psi \delta(a,s,(0,0)) \sim a^{-\frac{3}{4}},$$

   asymptotically as $a \to 0$. Show that if $t \neq (0,0)$, then, for any $N \in \mathbb{N}$ there is a constant $C_N > 0$ such that

$$\mathscr{S}_\psi \delta(a,s,(0,0)) \leq C_N\, a^N,$$

   asymptotically as $a \to 0$.
3. Let ψ and $\mathscr{S}_\psi$ be as in Exercise 2. For $p \in \mathbb{R}$, consider the distribution $\nu_p(x_1,x_2)$ defined by

$$\int_{\mathbb{R}^2} \nu_p(x_1,x_2)\, f(x_1,x_2)\, dx_1\, dx_2 = \int_{\mathbb{R}} f(px_2,x_2)\, dx_2.$$

   Show that, for $s = -p$ and $t_1 = pt_2$, we have:

$$\mathscr{S}_\psi \delta(a,s,(t_1,t_2)) \sim a^{-\frac{1}{4}},$$

   asymptotically as $a \to 0$. Show that for all other values of $t = (t_1,t_2)$ or s, then, for any $N \in \mathbb{N}$ there is a constant $C_N > 0$ such that

$$\mathscr{S}_\psi \delta(a,s,(0,0)) \le C_N\, a^N,$$

asymptotically as $a \to 0$.

References

1. J. Blanchard, *Minimally supported frequency composite dilation Parseval frame wavelets*, to appear in J. Geom Anal.
2. E.J. Candès and D.L. Donoho, *New tight frames of curvelets and optimal representations of objects with C^2 singularities*, Comm. Pure Appl. Math. **56** (2004), 219–266.
3. E. J. Candès and D. L. Donoho, *Continuous curvelet transform: I. Resolution of the wavefront set*, Appl. Comput. Harmon. Anal. **19** (2005), 162–197.
4. P. G. Casazza, *The art of frame theory*, Taiwanese J. Math., **4** (2000), 129–201.
5. A. Córdoba and C. Fefferman, *Wave packets and Fourier integral operators*, Comm. Partial Diff. Eq. **3** (1978), 979–1005.
6. O. Christensen, *An introduction to frames and Riesz bases*, Birkhäuser, Boston, 2003.
7. W. Czaja, *Characterizations of Gabor Systems via the Fourier Transform*, Collect. Math. **51** 2 (2000), 205–224.
8. E. Davies, *Machine Vision: Theory, Algorithms and Practicalities*, Academic Press, 1990.
9. R. A. DeVore, *Nonlinear approximation*, in: Acta Numerica, A. Iserles (Ed.), Cambridge University Press, Cambridge, 1998, pp. 51-150.
10. M. N. Do, M. Vetterli, The contourlet transform: an efficient directional multiresolution image representation, IEEE Trans. Image Proc. **14** (2005), 2091–2106.
11. G. R. Easley, D. Labate, and F. Colonna, *Shearlet Based Total Variation for Denoising*, to appear in IEEE Trans. Image Process (2008)
12. G. Easley, W. Lim, and D. Labate, *Optimally Sparse Image Representations using Shearlets*, Proc. 40th Asilomar Conf. on Signals, Systems and Computers, Monterey (2006),
13. G. Easley, W. Lim, and D. Labate, *Sparse Directional Image Representations using the Discrete Shearlet Transform*, Appl. Comput. Harmon. Anal., **25** (2008), 25–46.
14. G. Gripenberg, *A necessary and sufficient condition for the existence of a father wavelet*, Studia Math. **114** (1995), 297–226.
15. K. Guo, G. Kutyniok, and D. Labate, *Sparse multidimensional representations using anisotropic dilation and shear operators*, in: Wavelets and Splines, G. Chen and M. Lai (eds.), Nashboro Press, Nashville, TN (2006), 189–201.
16. K. Guo and D. Labate, *Some Remarks on the Unified Characterization of Reproducing Systems*, Collect. Math. **57** (2006), 279–293.
17. K. Guo and D. Labate, *Sparse Shearlet Representation of Fourier Integral Operators*, ERA-MS **14** (2007), 7–19.
18. K. Guo and D. Labate, *Representation of Fourier Integral Operators using Shearlets*, J. Fourier Anal. Appl. **14** (2008), 327–371.
19. K. Guo and D. Labate, *Optimally sparse multidimensional representation using shearlets*, SIAM J. Math. Anal. **39** (2007), 298–318.
20. K. Guo, W. Lim and D. Labate, *Characterization and Analysis of Edges using the Continuous Shearlet Transform*, preprint (2008).
21. K. Guo, W. Lim and D. Labate, *Edge Analysis and Identification using the Continuous Shearlet Transform*, to appear in Appl. Comput. Harmon. Anal. (2008).
22. K. Guo, W. Lim, D. Labate, G. Weiss and E. Wilson, *Wavelets with composite dilations*, Electron. Res. Announc. Amer. Math. Soc. **10** (2004), 78–87.
23. K. Guo, W. Lim, D. Labate, G. Weiss and E. Wilson, *Wavelets with composite dilations and their MRA properties*, Appl. Comput. Harmon. Anal. **20** (2006), 220–236.
24. P. Gressman, D. Labate, Weiss, G., and E.N. Wilson, *Affine, quasi-affine and co-affine and wavelets*, in Beyond Wavelets, J. Stoeckler and G. Welland (Eds.), 2002.

25. E. Hernández, D. Labate, and G. Weiss, *A unified characterization of reproducing systems generated by a finite family, II*, J. Geom. Anal. **12(4)** (2002), 615–662.
26. E. Hernández, D. Labate, G. Weiss, and E. Wilson *Oversampling, quasi affine frames and wave packets*, Appl. Comput. Harmon. Anal. **16** (2004), 111–147.
27. E. Ionascu, Y. Wang, *Simultaneous translations and multiplicative tiling and wavelet sets in* $\mathbb{R}^n$, preprint.
28. A.J.E.M. Janssen, *Signal analytic proof of two basic results on lattice expansions*, Appl. Comp. Harm. Anal., **1**(4) (1994), 350–354.
29. I. Kryshtal, B. Robinson, G. Weiss, and E. Wilson, *Compactly supported wavelets with composite dilations*, J. Geom. Anal. **17** (2006), 87–96.
30. G. Kutyniok and D. Labate, *The Construction of Regular and Irregular Shearlet Frames*, J. Wavelet Theory Appl., **1** (2007), 1–10.
31. G. Kutyniok and D. Labate, *Resolution of the Wavefront Set using Continuous Shearlets*, to appear in Trans AMS.
32. D. Labate, *A unified characterization of reproducing systems generated by a finite family*, J. Geom. Anal. **12(3)** (2002), 469–491.
33. R. S. Laugesen, N. Weaver, G. Weiss, and E. Wilson, *A characterization of the higher dimensional groups associated with continuous wavelets*, J. Geom. Anal. **12** (2001), 89–102.
34. S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, San Diego, CA, 1998.
35. S. Mallat and W. L. Hwang, *Singularity detection and processing with wavelets*, IEEE Trans. Inf. Theory **38**(2) (1992), 617-643.
36. S. Mallat and S. Zhong, *Characterization of signals from multiscale edges*, IEEE Trans. Pattern Anal. Mach. Intell. **14**(7) (1992), 710-732.
37. A. Ron and Shen, Z. *Frames and Stable Bases for Shift-Invariant Subspaces of* $L_2(\mathbb{R}^d)$, Can. J. Math. **47** (1995), 1051–1094.
38. A. Ron and Shen, Z. *Affine systems in* $L_2(\mathbb{R}^d)$*: the analysis of the analysis operator*, J. Funct. Anal. **148** (1997), 408–447.
39. A. Ron and Shen, Z. *Weyl-Heisenberg frames and Riesz bases in* $L^2(\mathbb{R}^d)$, Duke Math. J. **89** (1997), 237–282.
40. D.M. Speegle, *On the existence of wavelets for non–expansive dilation matrices*. Collect. Math. **54** (2003), 163–179.
41. X. Wang, *The study of wavelets from the properties of their Fourier transforms*, Ph.D. Thesis, Washington University in St. Louis, 1995.
42. G. Weiss and E. Wilson, *The mathematical theory of wavelets*, Proceeding of the NATO–ASI Meeting. Harmonic Analysis 2000 – A Celebration. Kluwer Publisher, 2001.
43. S. Yi, D. Labate, G. R. Easley, and H. Krim, *Edge Detection and Processing using Shearlets*, to appear in IEEE Int. Conf. on Image Processing (ICIP), (2008).
44. S. Yi, D. Labate, G. R. Easley, and H. Krim, *A Shearlet approach to Edge Analysis and Detection*, preprint 2008.