

Sampling Theory and Wavelets

John J. Benedetto

*Department of Mathematics
University of Maryland
College Park, Maryland 20742*

ABSTRACT. We introduce the theory of frames and develop wavelet theory as a natural extension of the Classical Sampling Theorem. This material serves as background for our applications to periodicity detection, noise reduction, and multidimensional irregular sampling.

1. Introduction

As a part of the mathematical base for Signal Processing for Multimedia, I shall present the fundamentals of wavelet theory. This material is highly developed. The references [26], [14], [11], [4], [24], [31], [29], [17] systematically develop and extend the original work on wavelets by Coifman, Daubechies, Feichtinger, Frazier, Gröchenig, Grossman, Jawerth, Mallat, Meyer, Morlet, et al., and integrate all of the important, antecedent, and sometimes equivalent notions from speech and image processing and harmonic analysis.

In the brief exposition that follows, I have chosen a path to wavelets beginning with the Classical Sampling Theorem. This approach is fully developed in the forthcoming second edition of [2], and is in the spirit of the general multimedia signal processing theme.

Because of space constraints, I have not included the following material that I was able to give in my oral presentation: Gabor theory, the theory of frames, and my applications to periodicity detection, noise reduction, and multidimensional irregular sampling, e.g., [4], Chapters 3 and 7, [6], [7], and [8].

I also want to mention the lifting scheme which is largely due to Wim Sweldens, e.g., [30], [15]. This approach, which is sometimes referred to as second generation wavelet theory, allows the implementation of the multirate system component of wavelet theory to be applied in cases where translation invariance and the Fourier transform are not available. These cases include wavelet theory on finite blocks, useful in image processing, and on spheres.

Our notation is standard, and our approach to wavelet theory is from the point of view of harmonic analysis as found in [2].

In particular, the *Fourier transform* of a function $f : \mathbb{R}^d \rightarrow \mathbb{C}$ is the function $\hat{f} : \hat{\mathbb{R}}^d \rightarrow \mathbb{C}$ defined as

$$\forall \gamma \in \hat{\mathbb{R}}^d, \quad \hat{f}(\gamma) = \int f(t) e^{-2\pi i t \cdot \gamma} dt,$$

where “ \int ” designates integration over Euclidean space \mathbb{R}^d and where $\hat{\mathbb{R}}^d$ is \mathbb{R}^d considered as the spectral or frequency domain of temporal or spatial functions defined on \mathbb{R}^d . The support of a function f is designated by $\text{supp } f$; $\delta(m, n)$ is 0 if $m \neq n$ and 1 if $m = n$; $\mathbf{1}_X$ is the characteristic function of the set X ; and $e_\gamma(t) = e^{-2\pi i t \cdot \gamma}$.

1. Gabor and Wavelet Systems

DEFINITION 1.1 (Gabor and Wavelet Systems).

- a. Let $g \in L^2(\mathbb{R})$ and let $a, b > 0$. The Gabor or Weyl-Heisenberg system (of coherent states) is the sequence

$$\{g_{m,n} : (m, n) \in \mathbb{Z} \times \mathbb{Z}\}$$

where

$$g_{m,n}(t) = e^{2\pi i t m b} g(t - na) = e_{mb}(t) \tau_{na} g(t).$$

Clearly,

$$(g_{m,n})^\wedge(\gamma) = e^{2\pi i n a m b} e^{-2\pi i n a \gamma} \widehat{g}(\gamma - mb) = \tau_{mb}(e_{-na} \widehat{g})(\gamma).$$

- b. Let $\psi \in L^2(\mathbb{R})$. The wavelet or affine system corresponding to ψ is the sequence $\{\psi_{m,n} : (m, n) \in \mathbb{Z} \times \mathbb{Z}\}$, where

$$\psi_{m,n}(t) = 2^{m/2} \psi(2^m t - n).$$

Clearly,

$$(\psi_{m,n})^\wedge(\gamma) = 2^{-m/2} e^{-2\pi i n(\gamma/2^m)} \widehat{\psi}(\gamma/2^m) = 2^{-m/2} e_{-n} \widehat{\psi}(\gamma/2^m).$$

DEFINITION 1.2 (Bases and Frames). Let H be a Hilbert space and let $\{x_n : n \in \mathbb{Z}\} \subseteq H$ be a sequence in H .

- a. The sequence $\{x_n\}$ is a basis or Schauder basis for H if each $x \in H$ has a unique decomposition

$$x = \sum_{n \in \mathbb{Z}} c_n(x) x_n \quad \text{in } H.$$

A basis $\{x_n\}$ for H is an orthonormal basis (ONB) for H if it is orthonormal, that is, if

$$\forall m, n \in \mathbb{Z}, \quad \langle x_m, x_n \rangle = \delta(m, n).$$

- b. A basis $\{x_n\}$ for H is an unconditional basis for H if $\exists C > 0$ such that $\forall F \subset \mathbb{Z}$, where $\text{card } F < \infty$, and $\forall b_n, c_n \in \mathbb{C}$, where $n \in F$ and $|b_n| < |c_n|$,

$$\left\| \sum_{n \in F} b_n x_n \right\| \leq C \left\| \sum_{n \in F} c_n x_n \right\|.$$

An unconditional basis is a bounded unconditional basis for H if

$$\exists A, B > 0 \quad \text{such that} \quad \forall n \in \mathbb{Z}, \quad A \leq \|x_n\| \leq B.$$

- c. The sequence $\{x_n\}$ is a frame for H if there are $A, B > 0$ such that

$$\forall x \in H, \quad A \|x\|^2 \leq \sum_{n \in \mathbb{Z}} |\langle x, x_n \rangle|^2 \leq B \|x\|^2.$$

The constants A and B are frame bounds, and a frame is tight if $A = B$. A frame is an exact frame if it is no longer a frame whenever any of its elements is removed.

The frame operator of a given frame $\{x_n\}$ for H is the mapping $S : H \rightarrow H$ defined by

$$Sx = \sum_{n \in \mathbb{Z}} \langle x, x_n \rangle x_n.$$

- d. If $\{x_n\}$ is a frame for H and S is the corresponding frame operator, then it can be shown that S is a bijective continuous linear map, and, hence, by the Open Mapping Theorem, S^{-1} is also continuous, e.g., [4], Theorem 3.2.

Assuming this result, we can deduce that $\{S^{-1}x_n\}$ is a frame for H with frame bounds B^{-1} and A^{-1} , and that

$$(1.1) \quad \forall x \in H, \quad x = \sum_{n \in \mathbb{Z}} \langle x, S^{-1}x_n \rangle x_n = \sum_{n \in \mathbb{Z}} \langle x, x_n \rangle S^{-1}x_n.$$

Using the continuity of S and S^{-1} , the verification of (1.1) is elementary. In fact,

$$x = SS^{-1}x = \sum_{n \in \mathbb{Z}} \langle S^{-1}x, x_n \rangle x_n$$

and

$$x = S^{-1}Sx = S^{-1} \left(\sum_{n \in \mathbb{Z}} \langle x, x_n \rangle x_n \right) = \sum_{n \in \mathbb{Z}} \langle x, x_n \rangle S^{-1}x_n.$$

- e. The notions of bounded unconditional bases and exact frames are equivalent, e.g., [32], [1], Theorem 17, [4], Theorem 3.7.

2. Gabor's and Morlet's Ideas

EXAMPLE 2.1 (Gabor's idea). Gabor's goal was to make the "best utilization of the information area" [18]. In order to formulate this idea we recall that the Classical Uncertainty Principle inequality is

$$(UP) \quad \|f\|_{L^2(\mathbb{R})}^2 \leq 4\pi \|(t - t_0)f(t)\|_{L^2(\mathbb{R})} \|(\gamma - \gamma_0)\widehat{f}(\gamma)\|_{L^2(\widehat{\mathbb{R}})}.$$

For $s > 0$ and $(t_0, \gamma_0) \in \mathbb{R} \times \widehat{\mathbb{R}}$ define the modulated and translated Gaussian,

$$g(t) = \sqrt{\frac{2s}{\pi}} e^{-s(t-t_0)^2} e^{2\pi i t \gamma_0}.$$

It is routine to check that g provides equality in (UP).

Setting

$$\sigma^2 t_s = 4\pi \|(t - t_0)g(t)\|_{L^2(\mathbb{R})}^2$$

and

$$\sigma^2 \gamma_s = 4\pi \|(\gamma - \gamma_0)\widehat{g}(\gamma)\|_{L^2(\widehat{\mathbb{R}})}^2,$$

we compute $\sigma^2 t_s = \pi/s$ and $\sigma^2 \gamma_s = s/\pi$, so that $\sigma^2 t_s \sigma^2 \gamma_s = 1$.

If $f \in L^2(\mathbb{R})$ and $\|f\|_{L^2(\mathbb{R})} = 1$, then the expected values associated with $|f|^2$ and $|\widehat{f}|^2$ are

$$\bar{t} = \int t |f(t)|^2 dt \quad \text{and} \quad \bar{\gamma} = \int \gamma |\widehat{f}(\gamma)|^2 d\gamma;$$

and the variances associated with $|f|^2$ and $|\widehat{f}|^2$ are

$$\sigma^2 t = \int (t - \bar{t})^2 |f(t)|^2 dt \quad \text{and} \quad \sigma^2 \gamma = \int (\gamma - \bar{\gamma})^2 |\widehat{f}(\gamma)|^2 d\gamma.$$

If $|f|^2$ is a probability density function of a random variable X , then $\sigma^2 t$ is the usual notion of the variance σ_X^2 of X , e.g., [2], Section 2.8.

Because of (UP) and the equality $\sigma^2 t_s \sigma^2 \gamma_s = 1$ for g , Gabor argued that decomposing the time-frequency plane by means of non-overlapping rectangles whose sides each have length and height $\sigma^2 t_s$ and $\sigma^2 \gamma_s$, respectively, would lead to signal decompositions with optimal time and frequency localization.

Gabor's decomposition is of the form

$$f(t) = \sum_{m,n \in \mathbb{Z}} c_{m,n} \sqrt{\frac{2s}{\pi}} e^{-s(t-n\sigma^2 t_s)^2} e^{2\pi i m \sigma^2 \gamma_s},$$

cf., [18] and [3] for the original decomposition and recent formulations, respectively.

EXAMPLE 2.2 (Morlet's idea). *Jean Morlet is a geophysicist who introduced wavelet theory in the course of his analysis of seismic traces s [27], [28], [20], [21]. These traces can be considered as the real part of finite energy signals f for which \hat{f} is causal.*

Morlet's idea was to analyze the trace s by a family of functions having a fixed shape, in the sense that each element of the family should be a lattice translation and (dyadic) dilation of one function ψ . Thus, in the dyadic case, the elements $\psi_{m,n}$ of the family are indexed by pairs (m, n) of integers, and each $\psi_{m,n}$ is defined as

$$\psi_{m,n}(t) = 2^{m/2} \psi(2^m t - n)$$

as in Definition 1.1b. *The aforementioned analysis is the computation of the sequence*

$$\{\langle s, \psi_{m,n} \rangle\}$$

where s and ψ are assumed to be elements of $L^2(\mathbb{R})$, and $\langle s, \psi_{m,n} \rangle = \int s(t) \overline{\psi_{m,n}(t)} dt$. *The goal is to reconstruct the trace s from the computable data set $\{\langle s, \psi_{m,n} \rangle\}$. The point of dealing with $\{\psi_{m,n}\}$ is that the operations of dilation and translation preserve the same number of cycles for high, medium or low frequencies.*

The sampled values $\langle s, \psi_{m,n} \rangle$ are analogous to the Fourier coefficients of s , in that we want to reconstruct s in terms of them. There are, however, fundamental differences between the wavelet system $\{\psi_{m,n}\}$ and Fourier analysis.

Before discussing these differences, and even before explicitly dealing with Morlet's original function ψ , we want to make a remark about terminology. The term wavelet often means a function ψ for which the family $\{\psi_{m,n}\}$ is an ONB for $L^2(\mathbb{R})$. On the other hand, Morlet's original work, which inspired so much of wavelet theory in the 1980s, was well-understood by Morlet and other geophysicists at the time to be a redundant wavelet system. As reinforced to us by Goupillaud [19], Morlet's family $\{\psi_{m,n}\}$ could not be orthonormal or independent in order to achieve the noise reduction required to solve the geophysical problems at hand. In essence, Morlet's wavelet systems were frames.

It should be pointed out that essentially equivalent "wavelet ideas" existed in the mathematical and engineering literature prior to Morlet's work, e.g., [22], [10], [12], [16], [9].

EXAMPLE 2.3 (Plots of Gabor and Morlet systems). *In Figure 1, the function g , for the system $\{g_{m,n}\}$, is the Gaussian defined in Example 2.1. In Figure 2, the function ψ for the system $\{\psi_{m,n}\}$ is*

$$\psi(t) = e^{-\pi t^2} \left(e^{2\pi i t \gamma_0} - e^{-\pi \gamma_0^2} \right).$$

These plots of the real parts of the Gabor and wavelet systems follow.

3. Sampling in terms of Wavelet and Gabor Systems

The Paley-Wiener space PW_Ω is defined as $PW_\Omega = \{f \in L^2(\mathbb{R}) : \text{supp } \hat{f} \subseteq [-\Omega, \Omega]\}$. The following result is proved in [2], pages 256–257.

THEOREM 3.1 (Classical sampling theorem). *Let $T, \Omega > 0$ satisfy the condition that $0 < 2T\Omega \leq 1$, and let $s \in PW_{1/(2T)}$ satisfy the condition that $\hat{s} = 1$ on $[-\Omega, \Omega]$ and $\hat{s} \in L^\infty(\hat{\mathbb{R}})$. Then*

$$(3.1) \quad \forall f \in PW_\Omega, \quad f = T \sum f(nT) \tau_{nT} s,$$

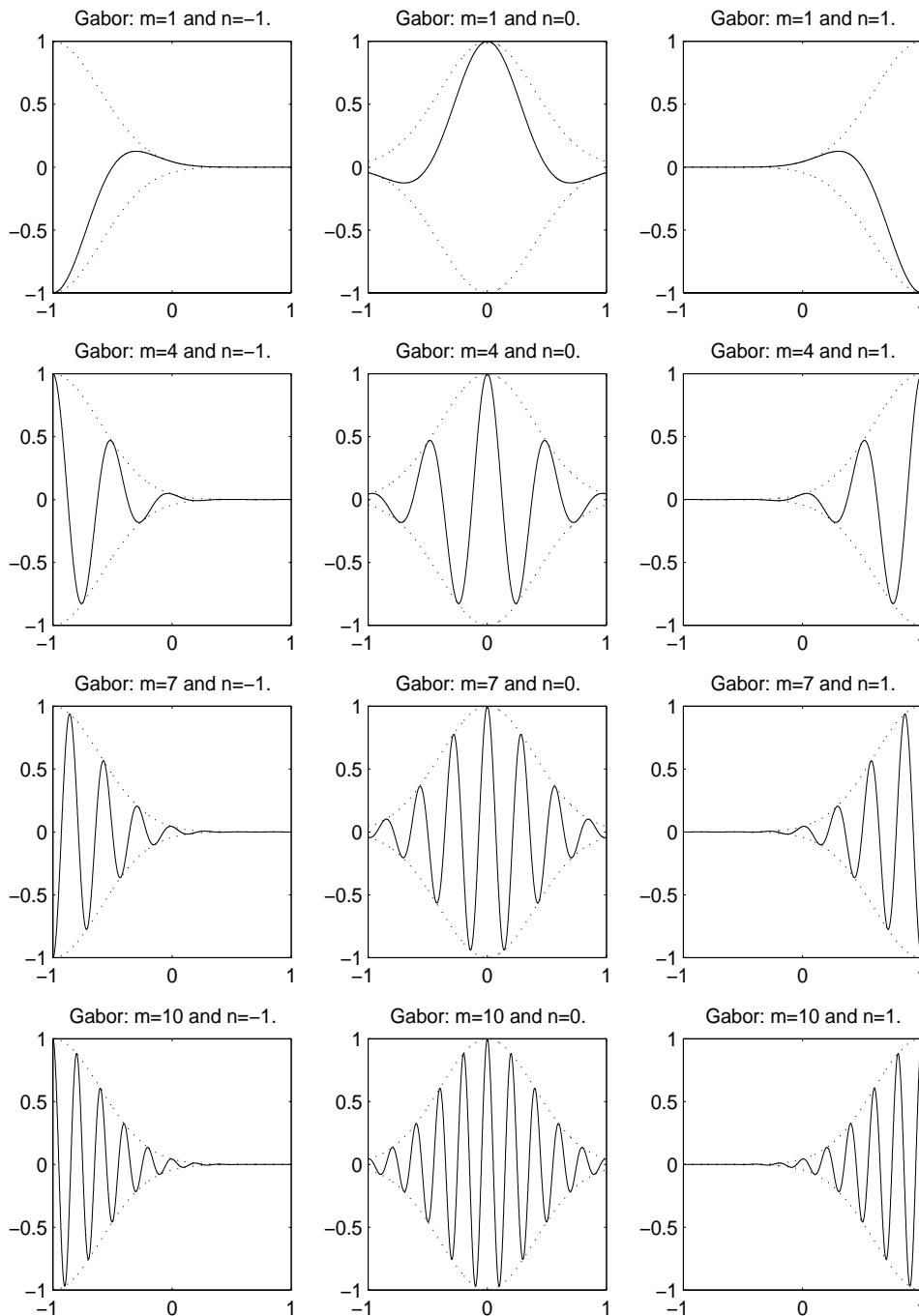


FIGURE 1. Gabor system

where $\tau_{nT}s(t) = s(t - nT)$, and where the convergence in (3.1) is in $L^2(\mathbb{R})$ norm and uniformly in \mathbb{R} . A possible sampling function s is the Dirichlet kernel

$$d_{2\pi\Omega}(t) = \frac{\sin 2\pi\Omega t}{\pi t}.$$

EXAMPLE 3.2 (The Shannon wavelet system). Let $\Omega > 0$. To be compatible with standard wavelet notation, let

$$\phi = \phi_{(\Omega)} = \frac{1}{\sqrt{2\Omega}} d_{2\pi\Omega}$$

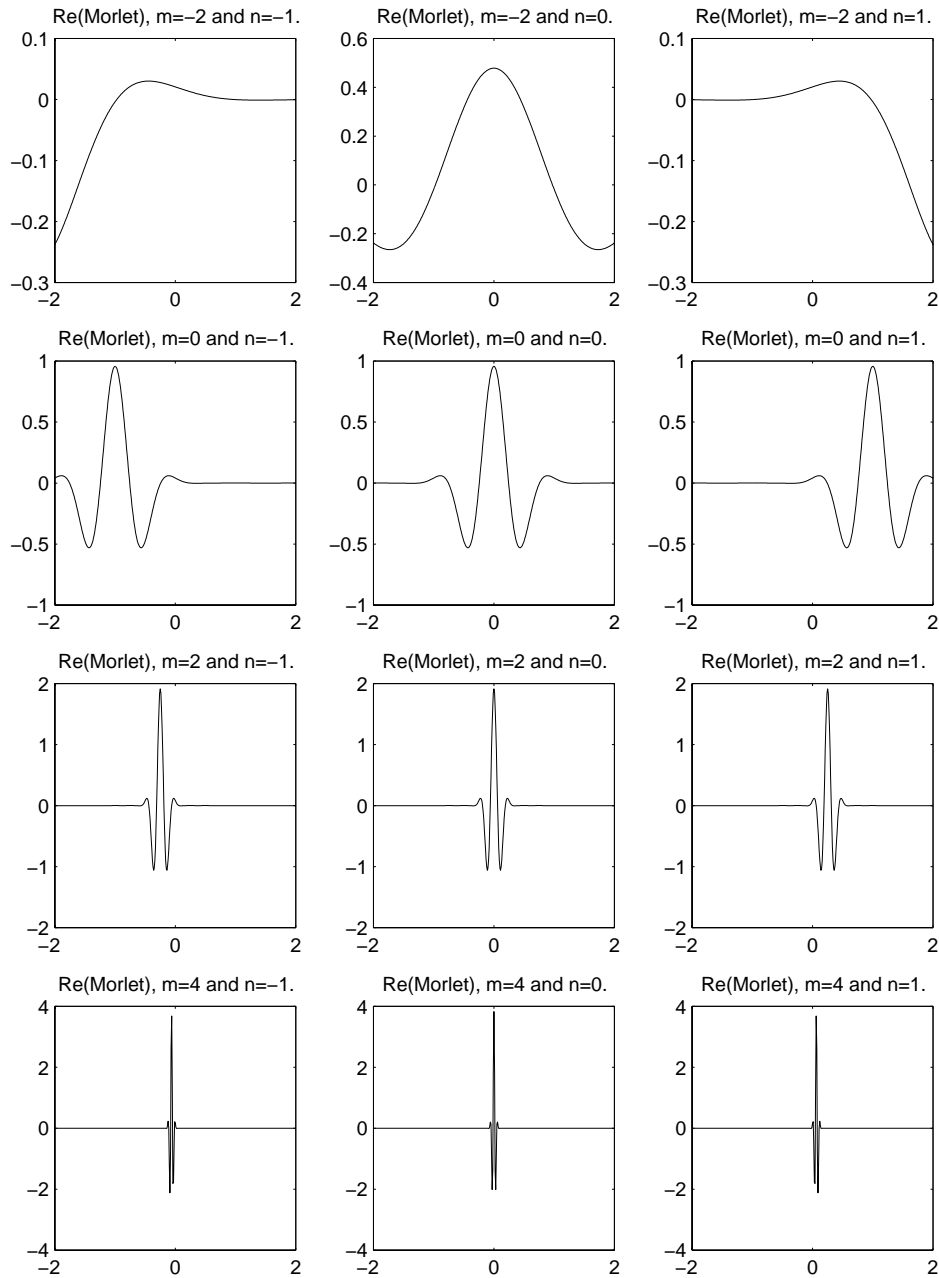


FIGURE 2. Morlet system

and let

$$\psi = \psi_{(\Omega)} = \frac{1}{\sqrt{2\Omega}}(d_{2\pi(2\Omega)} - d_{2\pi\Omega}).$$

$\psi_{(\Omega)}$ is the Shannon wavelet.

The following result is proved by periodizing $\sqrt{2\Omega}F(\gamma)\widehat{\psi}(\gamma/2^m)$ and calculating its Fourier series. This is the same method used to prove the Classical Sampling Theorem, and, in fact, *Theorem 3.1* is a special case of *Theorem 3.3* for $f \in PW_{\Omega}$.

THEOREM 3.3 (A Shannon wavelet decomposition using Fourier series). *Let $f \in L^2(\mathbb{R})$ have Fourier transform F , and let $\Omega > 0$. For the function $\phi = \frac{1}{\sqrt{2\Omega}}d_{2\pi\Omega}$ and the Shannon*

wavelet $\psi = \frac{1}{\sqrt{2\Omega}}(d_{2\pi(2\Omega)} - d_{2\pi\Omega})$, there is the decomposition

$$\begin{aligned}
 (3.2) \quad f &= \sqrt{2\Omega} f * \phi + \sum_{m=0}^{\infty} \sum_n d_{m,n} \psi_{m,n/(4\Omega)} \\
 &= \sum_n f * \phi\left(\frac{n}{2\Omega}\right) \tau_{n/(2\Omega)} \phi + \sum_{m=0}^{\infty} \sum_n d_{m,n} \psi_{m,n/(4\Omega)} \\
 &= \sum_m \sum_n d_{m,n} \psi_{m,n/(4\Omega)},
 \end{aligned}$$

where

$$d_{m,n} = \frac{1}{\sqrt{2\Omega} 2^{(m/2)+1}} \int_{-2^{m+1}\Omega}^{2^{m+1}\Omega} F(\gamma) \left(\mathbf{1}_{[-2^{m+1}\Omega, -2^m\Omega]}(\gamma) + \mathbf{1}_{[2^m\Omega, 2^{m+1}\Omega]}(\gamma) \right) e^{2\pi i n \gamma / (2^{m+2}\Omega)} d\gamma$$

and where convergence in (3.2) is in $L^2(\mathbb{R})$ norm.

With minor modifications we can then prove the following result.

THEOREM 3.4 (Shannon wavelet systems: ONBs and tight frames). *Let $\psi = \psi_{(\Omega)}$ be the Shannon wavelet.*

- a. $\{\psi_{m,n/(2\Omega)}\}$ is a wavelet ONB for $L^2(\mathbb{R})$.
- b. $\{\psi_{m,n/(4\Omega)}\}$ is a tight frame for $L^2(\mathbb{R})$ with frame constants $A = B = 2$.

Gabor systems also provide a generalization of the Classical Sampling Theorem, as illustrated by the following result [5].

THEOREM 3.5 (Gabor decomposition). *Let $T, \Omega > 0$ be constants for which $0 < 2T\Omega \leq 1$, and let $g \in PW_{1/(2T)}$ have the properties that $\hat{g} \in L^\infty(\widehat{\mathbb{R}})$,*

$$\hat{g} = 1 \quad \text{on} \quad [-\Omega, \Omega],$$

and, in case $2T\Omega < 1$, \hat{g} is continuous and

$$|\hat{g}| > 0 \quad \text{on} \quad \left(-\frac{1}{2T}, -\Omega\right] \cup \left[\Omega, \frac{1}{2T}\right).$$

Set

$$G(\gamma) = \sum |\hat{g}(\gamma - mb)|^2 \quad \text{and} \quad s(t) = \left(\frac{\hat{g}}{G}\right)^\vee(t),$$

where $\Omega + \frac{1}{2T} \leq b < \frac{1}{T}$ in case $2T\Omega < 1$ and $\Omega + \frac{1}{2T} = b$ if $2T\Omega = 1$. Then

$$\exists A, B > 0 \quad \text{such that} \quad A \leq G(\gamma) \leq B \quad \text{a.e.},$$

$s \in PW_{1/(2T)}$, $\hat{s} \in L^\infty(\widehat{\mathbb{R}})$, $\hat{s} = 1$ on $[-\Omega, \Omega]$,

$$\forall f \in L^2(\mathbb{R}), \quad f = T \sum \langle \hat{f}, e_{nT} \tau_{mb} \hat{g} \rangle \tau_{-nT} (e_{mb} s) \quad \text{in} \quad L^2(\mathbb{R}),$$

and

$$\forall f \in PW_\Omega, \quad f = T \sum f(nT) \tau_{nT} s \quad \text{in} \quad L^2(\mathbb{R}).$$

The generalizations, *Theorems 3.3–3.5*, of *Theorem 3.1* provide a mathematical explanation of aliasing, e.g., [1], [4], Chapter 7, and [2], Sections 3.10.12.

4. Multiresolution Analysis Wavelet ONBs

DEFINITION 4.1 (Orthonormal wavelets). *Let $\psi \in L^2(\mathbb{R})$. For each $m, n \in \mathbb{Z}$, $\psi_{m,n}$ is the function defined by $\psi_{m,n}(t) = 2^{m/2}\psi(2^m t - n)$. The function ψ is an orthonormal wavelet if $\{\psi_{m,n}\}$ is an ONB for $L^2(\mathbb{R})$.*

THEOREM 4.2 (Daubechies 1987). *Let $r \geq 1$. There are compactly supported r -times continuously differentiable orthonormal wavelets.*

EXAMPLE 4.3 (Deblurring and multiresolution). *Let ψ be an orthonormal wavelet. The concept of multiresolution is motivated and understood by the following nomenclature and intuition. Let $f \in L^2(\mathbb{R})$ and suppose $\text{supp } \psi \subseteq [-1/2, 1/2]$. Then $\text{supp } \psi_{m,n} \subseteq I_{m,n}$, where*

$$I_{m,n} = \left[\frac{n}{2^m} - \frac{1}{2^{m+1}}, \frac{n}{2^m} + \frac{1}{2^{m+1}} \right].$$

The length of $I_{m,n}$ is $|I_{m,n}| = 2^{-m}$. Let

$$f_M = \sum_{m \leq M} \sum_n \langle f, \psi_{m,n} \rangle \psi_{m,n}.$$

Then, $f_{M+1} = f_M + \sum_n \langle f, \psi_{M+1,n} \rangle \psi_{M+1,n}$ can be thought of as deblurring f_M by adding to f_M the behavior of f on intervals of length $2^{-(M+1)}$.

DEFINITION 4.4 (Multiresolution analysis (MRA)). *An MRA of $L^2(\mathbb{R})$ is a sequence*

$$\{V_j : j \in \mathbb{Z}\}$$

of closed subspaces of $L^2(\mathbb{R})$ satisfying the following properties:

- a. $\forall j \in \mathbb{Z}, V_j \subseteq V_{j+1}$;
- b. $f \in V_0$ if and only if $\tau_k f \in V_0$ for all $k \in \mathbb{Z}$;
- c. $f(t) \in V_j$ if and only if $f(2t) \in V_{j+1}$;
- d. $\bigcap V_j = \{0\}$ and $\overline{\bigcup V_j} = L^2(\mathbb{R})$;
- e. $\exists \phi \in V_0$ such that $\{\tau_k \phi\}$ is an ONB for V_0 .

The function ϕ is a scaling function for the $\{V_j\}$.

The proof of the following fundamental theorem can be found in [25], [26], [14]. As in Section 3, there is an interplay between Fourier series and Fourier transforms. In fact, 1-periodic Fourier series H_0 , defined by the property

$$|H_0(\gamma)|^2 + |H_0(\gamma + \frac{1}{2})|^2 = 2 \text{ a.e.},$$

are essential. These series are *quadrature mirror filters (QMFs)*.

THEOREM 4.5. *Let $\{V_j : j \in \mathbb{Z}\}$ be an MRA of $L^2(\mathbb{R})$ with scaling function ϕ . There is a constructible orthonormal wavelet ψ , depending on ϕ . In fact, a choice for the function ψ is*

$$\psi(t) = \sqrt{2} \sum_n h_1[n] \phi(2t - n),$$

where convergence is in $L^2(\mathbb{R})$, where

$$\forall n \in \mathbb{Z}, h_1[n] = (-1)^n \overline{h_0[-n+1]},$$

and where $\{h_0[n]\}$ is the sequence of Fourier coefficients of the QMF $H_0 \in L^\infty(\mathbb{T})$, which in turn is a solution of the frequency scaling equation

$$\sqrt{2} \widehat{\phi}(2\gamma) = H_0(\gamma) \widehat{\phi}(\gamma) \text{ in } L^2(\mathbb{R}).$$

H_1 is the Fourier series with Fourier coefficients $\{h_1[n]\}$.

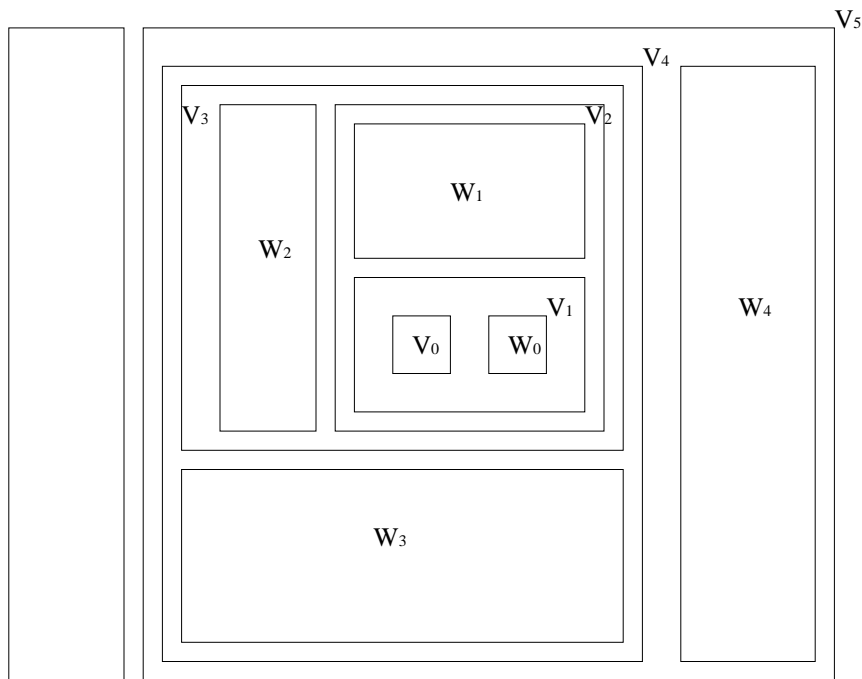


FIGURE 3

REMARK 4.6. a. Let $\{V_j\}$ be an MRA and let W_j be the orthogonal complement of V_j in V_{j+1} . In light of Definition 4.4 and Theorem 4.5, we have

$$(4.1) \quad L^2(\mathbb{R}) = \bigoplus_{j \in \mathbb{Z}} W_j.$$

The wavelet ψ of Theorem 4.5 is an element of W_0 and $\{\tau_n \psi\}$ is an ONB for W_0 . The inclusions, $V_j \subseteq V_{j+1}$, and the direct sum of (4.1) are illustrated in Figure 3.

b. Let $\phi = \mathbf{1}_{[0,1)}$, and let $V_0 = \overline{\text{span}}\{\tau_n \phi\}$. If we set $V_j = \{f \in L^2(\mathbb{R}) : 2^{j/2} f(2^j t) \in V_0\}$ then we obtain the Haar MRA of $L^2(\mathbb{R})$; and the function ψ of Theorem 4.5 is the Haar orthonormal wavelet. Similarly, if $p \in [1, \infty)$ and

$$V_0 = \left\{ \sum a_n \tau_n \phi : \{a_n\} \in \ell^p(\mathbb{Z}) \right\},$$

then we obtain the Haar MRA of $L^p(\mathbb{R})$.

EXAMPLE 4.7 (MRA and change of basis). Let $\{V_j\}$ be an MRA of $L^2(\mathbb{R})$ with scaling function $\phi \in V_0$ and orthonormal wavelet $\psi \in W_0$, constructed as in Theorem 4.5. Thus, $\{\tau_n \phi\}$ is an ONB for V_0 and $\{\tau_n \psi\}$ is an ONB for W_0 .

a. Besides $\{\tau_n \phi\}$, the sequence of functions

$$(4.2) \quad \frac{1}{\sqrt{2}} \phi\left(\frac{t}{2} - n\right), \frac{1}{\sqrt{2}} \psi\left(\frac{t}{2} - n\right),$$

where $n \in \mathbb{Z}$, forms an ONB for V_0 .

To see this, first note that

$$(4.3) \quad \forall n \in \mathbb{Z}, \quad \frac{1}{\sqrt{2}} \phi\left(\frac{t}{2} - n\right) \in V_{-1}.$$

In fact, if $f(t) = \phi(t - n)$, then $f \in V_0$ so that $f(\frac{t}{2}) \in V_{-1}$ by definition of an MRA. Thus, (4.3) is obtained since $f(\frac{t}{2}) = \phi(\frac{t}{2} - n)$ and since V_{-1} is a linear subspace of

$L^2(\mathbb{R})$. Next, we can assert that

$$(4.4) \quad \forall n \in \mathbb{Z}, \quad \frac{1}{\sqrt{2}}\psi\left(\frac{t}{2} - n\right) \in W_{-1}.$$

In fact, if $g(t) = \psi(t - n)$, then $g \in W_0$ so that $g(\frac{t}{2}) \in W_{-1}$. Thus (4.4) is obtained since $g(\frac{t}{2}) = \psi(\frac{t}{2} - n)$ and since W_{-1} is a vector subspace of $L^2(\mathbb{R})$. The orthonormality of the functions in (4.2) is a consequence of the orthonormality of $\{\tau_n\phi\}$ in V_0 , the orthonormality of $\{\tau_n\psi\}$ in W_0 , and the fact that $V_0 = V_{-1} \oplus W_{-1}$. For example,

$$\int \frac{1}{\sqrt{2}}\phi\left(\frac{t}{2} - n\right) \frac{1}{\sqrt{2}}\overline{\phi\left(\frac{t}{2} - q\right)} dt = \int \phi(u - n)\overline{\phi(u - q)} du = \delta(n, q).$$

Further, the fact that the functions in (4.2) form an ONB for V_0 is a consequence of the just proved orthonormality and the following calculations for arbitrary $f \in V_0$:

$$f = f_V + f_W, \text{ where } f_V \in V_{-1}, f_W \in W_{-1};$$

$f_V(2t) \in V_0$ so that

$$f_V(2t) = \sum \langle f_V(2u), \tau_n\phi(u) \rangle \tau_n\phi(t) \text{ in } L^2(\mathbb{R}),$$

and thus,

$$f_V(t) = \sum \langle \sqrt{2}f_V(2u), \tau_n\phi(u) \rangle \frac{1}{\sqrt{2}}\phi\left(\frac{t}{2} - n\right) \in V_{-1}.$$

and finally $f_W(2t) \in W_0$ so that

$$f_W(2t) = \sum \langle f_W(2u), \tau_n\psi(u) \rangle \tau_n\psi(t) \text{ in } L^2(\mathbb{R}),$$

and thus

$$f_W(t) = \sum \langle \sqrt{2}f_W(2u), \tau_n\psi(u) \rangle \frac{1}{\sqrt{2}}\psi\left(\frac{t}{2} - n\right) \in W_{-1}.$$

- b. The procedure of part a can be used to make other ONBs for V_0 . In fact, for any fixed $m \in \mathbb{N}$ we have the orthogonal complement direct sum

$$V_0 = V_{-m} \oplus W_{-m} \oplus W_{-m+1} \oplus W_{-m+2} \oplus \dots \oplus W_{-1};$$

and hence the sequence of functions

$$\phi_{-m,n}, \psi_{-m,n}, \psi_{-m+1,n}, \dots, \psi_{-1,n},$$

where $n \in \mathbb{Z}$, forms an ONB for V_0 .

This procedure can be generalized significantly by introducing the idea of wavelet-packets.

5. Waveletpackets

Let $\{V_j\}$ be an MRA of $L^2(\mathbb{R})$ with scaling function $\phi \in V_0$ and orthonormal wavelet $\psi \in W_0$ constructed as in *Theorem 4.5*. We have

$$\phi(t) = \sum h_0[n]\phi_{1,n}(t) \in V_1 \quad \text{in } L^2(\mathbb{R})$$

and

$$\psi(t) = \sum h_1[n]\phi_{1,n}(t) \in V_1 \quad \text{in } L^2(\mathbb{R}).$$

Thus,

$$\widehat{\phi}(\gamma) = \frac{1}{\sqrt{2}}\widehat{\phi}\left(\frac{\gamma}{2}\right)H_0\left(\frac{\gamma}{2}\right) \quad \text{and} \quad \widehat{\psi}(\gamma) = \frac{1}{\sqrt{2}}\widehat{\phi}\left(\frac{\gamma}{2}\right)H_1\left(\frac{\gamma}{2}\right) \quad \text{in } L^2(\widehat{\mathbb{R}}).$$

The ONB splitting of V_0 for a given MRA, reflected by (4.2), can be effected at every node in the dyadic tree of *Figure 4*. (At this point we have only labeled the nodes $V_0, V_{-1}, W_{-1}, V_{-2}, W_{-2}, V_{-3}, W_{-3}, \dots$) To see how this is done, we begin with the ONB $\{\tau_n\phi\}$ for

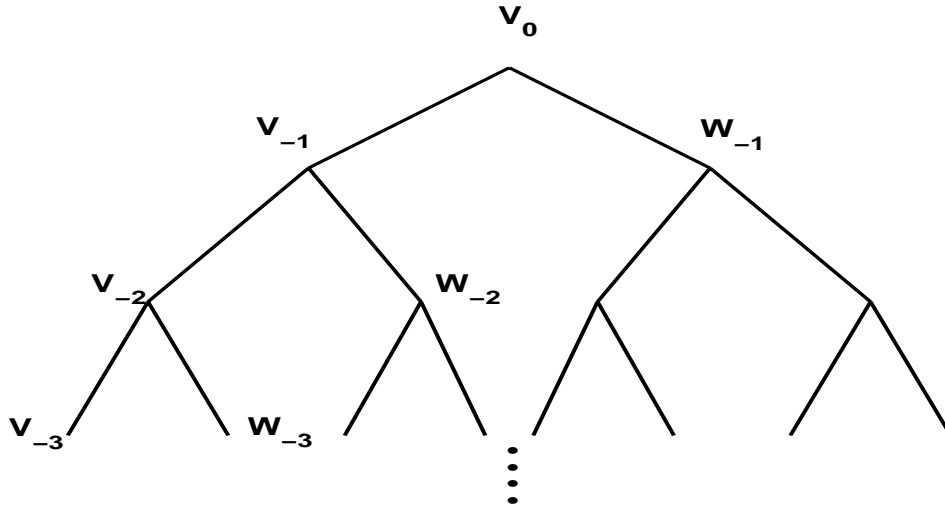


FIGURE 4

V_0 , or, equivalently, the ONB $\{e_{-n}\widehat{\phi}\}$ for $\widehat{V}_0 \subseteq L^2(\widehat{\mathbb{R}})$. Then the nodes designated V_{-1} and W_{-1} can be characterized by the formulas written in *Figure 5*.

$$\begin{array}{cc}
 V_{-1} & W_{-1} \\
 \phi_{-1,n}(t) = \frac{1}{\sqrt{2}}\phi\left(\frac{t}{2} - n\right) \leftrightarrow e_{-2n}(\gamma)\sqrt{2}\widehat{\phi}(2\gamma), & \psi_{-1,n}(t) = \frac{1}{\sqrt{2}}\psi\left(\frac{t}{2} - n\right) \leftrightarrow e_{-2n}(\gamma)\sqrt{2}\widehat{\psi}(2\gamma) \\
 \sqrt{2}\widehat{\phi}(2\gamma) = \widehat{\phi}(\gamma)H_0(\gamma) & \sqrt{2}\widehat{\psi}(2\gamma) = \widehat{\psi}(\gamma)H_1(\gamma)
 \end{array}$$

Figure 5

In particular, $\{e_{-2n}(\gamma)\sqrt{2}\widehat{\phi}(2\gamma)\}$ is an ONB for \widehat{V}_{-1} and $\{e_{-2n}(\gamma)\sqrt{2}\widehat{\psi}(2\gamma)\}$ is an ONB for \widehat{W}_{-1} . We obtain an ONB splitting of V_{-1} by the sequence

$$(5.1) \quad \left\{ \left(\frac{1}{\sqrt{2}}\right)^2 \phi\left(\frac{t}{2^2} - n\right), \left(\frac{1}{\sqrt{2}}\right)^2 \psi\left(\frac{t}{2^2} - n\right) : n \in \mathbb{Z} \right\},$$

noting that

$$(5.2) \quad \left\{ \left(\frac{1}{\sqrt{2}}\right)^2 \phi\left(\frac{t}{2^2} - n\right) \right\} \subseteq V_{-2} \quad \text{and} \quad \left\{ \left(\frac{1}{\sqrt{2}}\right)^2 \psi\left(\frac{t}{2^2} - n\right) \right\} \subseteq W_{-2}.$$

In the case of (5.1) and (5.2), the analogue of *Figure 5*, for the nodes designated V_{-2} and W_{-2} coming from the splitting of V_{-1} , is given by the formulas written in *Figure 6*.

$$\begin{array}{cc}
 V_{-2} & W_{-2} \\
 \phi_{-2,n}(t) = \left(\frac{1}{\sqrt{2}}\right)^2 \phi\left(\frac{t}{4} - n\right) & \psi_{-2,n}(t) = \left(\frac{1}{\sqrt{2}}\right)^2 \psi\left(\frac{t}{4} - n\right) \\
 \longleftrightarrow & \longleftrightarrow \\
 e_{-4n}(\gamma)(\sqrt{2})^2 \widehat{\phi}(4\gamma), & e_{-4n}(\gamma)(\sqrt{2})^2 \widehat{\psi}(4\gamma), \\
 (\sqrt{2})^2 \widehat{\phi}(4\gamma) = \widehat{\phi}(\gamma)H_0(\gamma)H_0(2\gamma) & (\sqrt{2})^2 \widehat{\psi}(4\gamma) = \widehat{\psi}(\gamma)H_0(\gamma)H_1(2\gamma)
 \end{array}$$

Figure 6

In particular, $\{e_{-4n}(\gamma)(\sqrt{2})^2 \widehat{\phi}(4\gamma)\}$ is an ONB for \widehat{V}_{-2} and $\{e_{-4n}(\gamma)(\sqrt{2})^2 \widehat{\psi}(4\gamma)\}$ is an ONB for \widehat{W}_{-2} . Clearly, the last formulas in *Figure 6* are consequences of equation (5.1), e.g.,

$$(\sqrt{2})^2 \widehat{\phi}(4\gamma) = \sqrt{2}\widehat{\phi}(2\gamma)H_0(2\gamma) = \widehat{\phi}(\gamma)H_0(\gamma)H_0(2\gamma).$$

At this point it is reasonable to ask if there is a natural splitting of W_{-1} in *Figure 4* to complete the remaining two nodes which are on the same level as V_{-2} and W_{-2} .

The answer is “yes” and the procedure results in new ONBs for V_0 . In order to describe the procedure it is convenient to use the classical dyadic tree notation, e.g., [2], Remark 3.9.4. Thus, at level 0, X_0^0 denotes V_0 ; at level 1, X_0^1 and X_1^1 denote V_{-1} and W_{-1} ; at level 2, $X_{(0,0)}^2$ and $X_{(0,1)}^2$ denote V_{-2} and W_{-2} .

In the case of level 1, the subscript 0 of X_0^1 corresponds to the subscript 0 of H_0 on the left side of *Figure 5*, and the subscript 1 of X_1^1 corresponds to the subscript 1 of H_1 on the right side of *Figure 5*.

In the case of level 2, there are the four consecutive spaces $X_{(0,0)}^2$, $X_{(0,1)}^2$, $X_{(1,0)}^2$, and $X_{(1,1)}^2$. The subscript (0, 0) of $X_{(0,0)}^2 = V_{-2}$ corresponds to the two subscripts 0 and 0 of $H_0(\gamma)$ and $H_0(2\gamma)$ on the left side of *Figure 6*. The subscript (0, 1) of $X_{(0,1)}^2 = W_{-2}$ corresponds to the two subscripts 0 and 1 of $H_0(\gamma)$ and $H_1(2\gamma)$ on the right side of *Figure 6*.

Because of this pattern we associate the function $\widehat{\phi}(\gamma)H_1(\gamma)H_0(2\gamma)$ with $X_{(1,0)}^2$ and the function $\widehat{\phi}(\gamma)H_1(\gamma)H_1(2\gamma)$ with $X_{(1,1)}^2$. In this way, level 2 of the tree in *Figure 4* can be completed. In fact, we define the functions $\theta_{1,0}$ and $\theta_{1,1}$ by the formulas

$$(\sqrt{2})^2 \widehat{\theta}_{1,0}(2^2\gamma) = \widehat{\phi}(\gamma)H_1(\gamma)H_0(2\gamma)$$

and

$$(\sqrt{2})^2 \widehat{\theta}_{1,1}(2^2\gamma) = \widehat{\phi}(\gamma)H_1(\gamma)H_1(2\gamma).$$

Then $X_{(1,0)}^2$ is defined as

$$X_{(1,0)}^2 = \overline{\text{span}} \left\{ \frac{1}{(\sqrt{2})^2} \theta_{1,0} \left(\frac{t}{2^2} - n \right) : n \in \mathbb{Z} \right\} \quad \text{in } L^2(\mathbb{R})$$

and $X_{(1,1)}^2$ is defined as

$$X_{(1,1)}^2 = \overline{\text{span}} \left\{ \frac{1}{(\sqrt{2})^2} \theta_{1,1} \left(\frac{t}{2^2} - n \right) : n \in \mathbb{Z} \right\} \quad \text{in } L^2(\mathbb{R}).$$

It is easy to prove that $\left\{ \frac{1}{(\sqrt{2})^2} \theta_{1,0} \left(\frac{t}{2^2} - n \right) \right\}$ is an ONB for $X_{(1,0)}^2$, that $\left\{ \frac{1}{(\sqrt{2})^2} \theta_{1,1} \left(\frac{t}{2^2} - n \right) \right\}$ is an ONB for $X_{(1,1)}^2$, and that

$$W_{-1} = X_1^1 = X_{(1,0)}^2 \oplus X_{(1,1)}^2,$$

an orthogonal complement direct sum.

The previous details give rise to the following general procedure for a given MRA. We form the tree $\{X_m^r : r \geq 1\}$, where $m \in \{0, 1, 2, 3, \dots, 2^r - 1\}$ and where $X_0^0 = V_0$. Each element X_m^r of the tree is a closed linear subspace of V_0 . Further, X_m^r , $r \geq 1$, is determined by the function

$$(5.3) \quad 2^{r/2} \widehat{\theta}_{\epsilon_1, \dots, \epsilon_r}(2^r\gamma) = \widehat{\phi}(\gamma)H_{\epsilon_1}(\gamma)H_{\epsilon_2}(2\gamma)H_{\epsilon_3}(2^2\gamma) \dots H_{\epsilon_r}(2^{r-1}\gamma),$$

$m = \sum_{j=1}^r \epsilon_j 2^{j-1}$ and $\epsilon_j \in \{0, 1\}$, in the sense that

$$(5.4) \quad X_m^r = \overline{\text{span}} \left\{ 2^{-r/2} \theta_{\epsilon_1, \dots, \epsilon_r} \left(\frac{t}{2^r} - n \right) : n \in \mathbb{Z} \right\},$$

$$(5.5) \quad \left\{ 2^{-r/2} \theta_{\epsilon_1, \dots, \epsilon_r} \left(\frac{t}{2^r} - n \right) : n \in \mathbb{Z} \right\} \quad \text{is an ONB for } X_m^r,$$

and

$$(5.6) \quad V_0 = \bigoplus_{m=0}^{2^r-1} X_m^r,$$

an orthogonal complement direct sum.

The *Wavelet Packet Algorithm* is the tree $\{X_m^r\}$ of subspaces of V_0 defined by (5.3), (5.4), (5.5), and (5.6). We have written the orthogonal complement direct sum in (5.6) in terms of the natural ordering $0, 1, \dots, 2^r - 1$. In fact, at level r , the spaces X_m^r are ordered, left to right on the tree, by bit reversal ordering.

6. Multidimensional Orthonormal Wavelets

EXAMPLE 6.1 (Rectilinear wavelet decompositions). *Let ψ be an orthonormal wavelet for $L^2(\mathbb{R})$. For each $(j, k) \in \mathbb{Z}^2$ and $(m, n) \in \mathbb{Z}^2$, define the tensor product*

$$\psi_{j,k} \otimes \psi_{m,n}(x_1, x_2) = \psi_{j,k}(x_1)\psi_{m,n}(x_2).$$

Then $\{\psi_{j,k} \otimes \psi_{m,n} : j, k, m, n \in \mathbb{Z}\}$ is an ONB for $L^2(\mathbb{R}^2)$. For every ordered 4-tuple $(j, k, m, n) \in \mathbb{Z}^4$, we define

$$(\psi \otimes \psi)_{j,k,m,n}(x_1, x_2) = \psi_{j,k} \otimes \psi_{m,n}(x_1, x_2).$$

As such we shall say that $\psi \otimes \psi$ is a rectilinear orthonormal wavelet for $L^2(\mathbb{R}^2)$. Notice that $(\psi \otimes \psi)_{j,k,m,n}$ provides separate dilation and translation in the x_1 and x_2 variables. Consequently, the wavelet decomposition of arbitrary $f \in L^2(\mathbb{R}^2)$ involves rectangles of all shapes, e.g., very long and skinny ones, in order to achieve reconstruction. We refer to this approach as the rectilinear tensor product wavelet decomposition.

EXAMPLE 6.2 (MRA wavelet decompositions). *Let $\{V_j\}$ be an MRA of $L^2(\mathbb{R})$ with scaling function ϕ and orthonormal wavelet ψ constructed from Theorem 4.5. Define the tensor product*

$$\phi \otimes \phi(x_1, x_2) = \phi(x_1)\phi(x_2),$$

and let V_0^2 be the set of functions $f : \mathbb{R}^2 \rightarrow \mathbb{C}$ having the form

$$f(x_1, x_2) = \sum_{n_1, n_2 \in \mathbb{Z}} c_{n_1, n_2} \phi \otimes \phi(x_1 - n_1, x_2 - n_2),$$

where $c = \{c_{n_1, n_2}\} \in \ell^2(\mathbb{Z}^2)$. Then, $V_0^2 = \overline{\text{span}}\{f \otimes g(x_1, x_2) : f, g \in V_0\} = V_0 \hat{\otimes} V_0$, the projective tensor product. Next, set

$$V_j^2 = \{g(x_1, x_2) = f(2^j x_1, 2^j x_2) : f \in V_0^2\}.$$

It can now be proved that $\{V_j^2 : j \in \mathbb{Z}\}$ is an MRA of $L^2(\mathbb{R}^2)$ with scaling function $\phi \otimes \phi$. The orthogonal complement of V_0^2 in V_1^2 is denoted by W_0^2 .

Clearly,

$$\begin{aligned} V_{j+1}^2 &= V_{j+1} \hat{\otimes} V_{j+1} = (V_j \oplus W_j) \hat{\otimes} (V_j \oplus W_j) \\ &= V_j \hat{\otimes} V_j \oplus [(W_j \hat{\otimes} V_j) \oplus (V_j \hat{\otimes} W_j) \oplus (W_j \hat{\otimes} W_j)]. \end{aligned}$$

Thus, three wavelets, viz.,

$$\begin{aligned} \phi \otimes \psi(x_1, x_2) &= \phi(x_1)\psi(x_2), \\ \psi \otimes \phi(x_1, x_2) &= \psi(x_1)\phi(x_2), \\ \psi \otimes \psi(x_1, x_2) &= \psi(x_1)\psi(x_2), \end{aligned}$$

are required to provide an MRA wavelet decomposition of $L^2(\mathbb{R}^2)$.

In light of this background, the following is an amazing result, e.g., [13].

THEOREM 6.3. *There exists $K \subseteq \hat{\mathbb{R}}^d$ such that $\hat{\psi} = \mathbf{1}_K$ defines an orthonormal wavelet for $L^2(\mathbb{R}^d)$, i.e.,*

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

is an ONB for $L^2(\mathbb{R}^d)$.

It was elementary to construct wavelet ONBs for $L^2(\mathbb{R}^d)$ of the form

$$\psi_{j_1, k_1} \otimes \dots \otimes \psi_{j_d, k_d},$$

where ψ is an orthonormal wavelet for $L^2(\mathbb{R})$ and $j_1, \dots, j_d, k_1, \dots, k_d \in \mathbb{Z}$ (Example 6.1). Theorem 6.3 is a fundamentally deeper result since the dilation depends only on $j \in \mathbb{Z}$. Similarly, it is a deeper result than the MRA approach to \mathbb{R}^d (Example 6.2).

References

- [1] J. J. Benedetto, Irregular sampling and frames, in *Wavelets—A Tutorial in Theory and Applications*, C. K. Chui, editor, Academic Press, Boston (1992), 445–507.
- [2] J. J. Benedetto, *Harmonic Analysis and Applications*, CRC Press, Boca Raton, FL, 1997.
- [3] J. J. Benedetto, Generalized harmonic analysis in Gabor and Wavelet systems, Wiener Centennial, *Amer. Math. Soc. Proc. of Symposia in Applied Math.*, 52(1997), 85–113.
- [4] J. J. Benedetto and M. W. Frazier, editors, *Wavelets: Mathematics and Applications*, CRC Press, Boca Raton, FL, 1994.
- [5] J. J. Benedetto and W. Heller, Irregular sampling and the theory of frames. *Note di Matematica*, X, 10(1990), 103–125, Suppl.n.1.
- [6] J. Benedetto and S. Li, The theory of multiresolution frames with applications to filter banks, *Appl. and Comp. Harmonic Analysis*.
- [7] J. J. Benedetto and Goetz Pfander, Wavelet periodicity detection.
- [8] J. J. Benedetto and H.-C. Wu, Balayage and multidimensional signal reconstruction from data on curves.
- [9] P. Burt and T. Adelson, The Laplacian pyramid as a compact image code, *IEEE Trans. Commun.*, 31(1983), 532–540.
- [10] A. Calderón, Intermediate spaces and interpolation, the complex method, *Studia Math.*, 24(1964), 113–190.
- [11] C. K. Chui, editor, *Wavelets—A Tutorial in Theory and Applications*, Academic Press, Boston, 1992.
- [12] A. Croisier, D. Esteban and C. Galand, Perfect channel splitting by use of interpolation decimation tree decomposition techniques, *Int'l. Conference on Information Sciences and Systems*, Patras, 1976, pages 443–446.
- [13] X. Dai, D. R. Larson and D. M. Speegle, Wavelet sets in \mathbb{R}^n , *J. Fourier Analysis and Applications*, 3(1997), 451–456.
- [14] I. Daubechies, *Ten Lectures on Wavelets*, volume 61 of CBMS-NSF Conf. Lect. Notes Ser. Appl. Math., Soc. Ind. Appl. Math., Philadelphia, PA, 1992.
- [15] I. Daubechies and W. Sweldons, Factoring wavelet transforms into lifting steps, *J. Fourier Analysis and Applications*, 4(1998).
- [16] D. Esteban and C. Galand Applications of quadrature mirror filters to split-band voice coding schemes, *Proc. ICASSP*, Hartford, 1977, pages 191–195.
- [17] H. G. Feichtinger and T. Strohmer, editors, *Gabor Analysis and Algorithms*, Birkhauser, Boston 1998.
- [18] D. Gabor, Theory of communication, *J.IEE London*, 93(1946), 429–457.
- [19] P. Goupillaud, Oral communication, 1997, *AMS Special Session on Wavelet Theory*, San Diego.
- [20] P. Goupillaud, A. Grossmann, and J. Morlet, Cycle-octave and related transforms in seismic signal analysis, *Geoexploration*, e.g., 23(1984/1985), 85–102.
- [21] A. Grossmann and J. Morlet, Decomposition of Hardy functions into square integrable wavelets of constant shape, *SIAM J. Math. Analysis*, 15(1984), 723–736.
- [22] A. Haar, Zur Theorie der orthogonale Funktionensysteme, *Math. Annalen*, 69(1909), 331–371.
- [23] E. Hernández and G. Weiss, *A First Course on Wavelets*, CRC Press, Boca Raton, FL, 1996.
- [24] M. Holschneider, *Wavelets: An Analysis Tool*, Oxford University Press, Oxford, 1995.
- [25] S. Mallat, Multiresolution approximations and wavelet orthonormal bases of $L^2(R)$, *Trans. Amer. Math. Soc.*, 315(1989), 69–87.
- [26] Y. Meyer, *Ondelettes et Opérateurs*, Hermann, Paris, 1990.
- [27] J. Morlet, Sampling theory and wave propagation, *Proc. Soc. Exploration Geophysicists*, October 1981 (March 1982).
- [28] J. Morlet et al., Wave propagation and sampling theory, Parts I and II, *Geophysics*, 47(1982), 203–221 and 222–236.
- [29] G. Strang and T. Nguyen, *Wavelets and Filter Banks*, Wellesley-Cambridge Press, Wellesley, MA, 1996.
- [30] W. Sweldons, The lifting scheme: a custom-designed construction of biorthogonal wavelets, *Appl. and Comp. Harmonic Analysis*, 3(1996), 186–200.

- [31] M. Vetterli and J. Kovačević, *Wavelets and Subband Coding*, Prentice Hall, Inc., Englewood Cliffs, NJ, 1995.
- [32] R. M. Young, *An Introduction to Nonharmonic Fourier Series*, Academic Press, New York, 1980.

