NUMER. FUNCT. ANAL. AND OPTIMIZ., 14(5&6), 417–446 (1993)



# FAMILIES OF MULTIRESOLUTION AND WAVELET SPACES WITH OPTIMAL PROPERTIES

**Akram Aldroubi and Michael Unser**

*Biomedical Engineering and Instrumentation Program*
*National Institutes of Health, Bethesda, Maryland*

## ABSTRACT

Under suitable conditions, if the scaling functions $\varphi_1$ and $\varphi_2$ generate the multiresolutions $V_{(j)}(\varphi_1)$ and $V_{(j)}(\varphi_2)$, then their convolution $\varphi_1 * \varphi_2$ also generates a multiresolution $V_{(j)}(\varphi_1 * \varphi_2)$. Moreover, if $p$ is an appropriate convolution operator from $l_2$ into itself and if $\varphi$ is a scaling function generating the multiresolution $V_{(j)}(\varphi)$, then $p * \varphi$ is a scaling function generating the same multiresolution $V_{(j)}(\varphi) = V_{(j)}(p * \varphi)$. Using these two properties, we group the scaling and wavelet functions into equivalent classes and consider various equivalent basis functions of the associated function spaces. We use the $n$-fold convolution product to construct sequences of multiresolution and wavelet spaces $V_{(j)}(\varphi^n)$ and $W_{(j)}(\varphi^n)$ with increasing regularity. We discuss the link between multiresolution analysis and Shannon's sampling theory. We then show that the interpolating and orthogonal pre- and post-filters associated with the multiresolution sequence $V_{(0)}(\varphi^n)$ asymptotically converge to the ideal lowpass filter of Shannon. We also prove that the filters associated with the sequence of wavelet spaces $W_{(0)}(\varphi^n)$ converge to the ideal bandpass filter. Finally, we construct the basic wavelet sequences $\psi_b^n$ and show that they tend to Gabor functions. This provides wavelets that are nearly time-frequency optimal. The theory is illustrated with the example of polynomial splines.

Correspondence to: Akram Aldroubi, Rm. 3W13, Bldg. 13, BEIP, National Institutes of Health, Bethesda, MD 20892.

## 1. INTRODUCTION

Digital signal processing requires a means by which to represent functions defined on $\mathbb{R}$ in terms of discrete sequences and vice versa [57]. This is one of the reasons for which the classical Shannon-Whittaker Sampling Theorem [48, 55] has become popular in applied mathematics and engineering. It states that bandlimited $L_2$ functions (i.e., $L_2$ functions that have compactly supported Fourier transforms) are completely characterized by their samples, as long as the sampling rate is sufficiently fast. In particular, if the Fourier transform $S(f)$ of the real valued function $s$ is such that Support$(S(f)) \subset [-\frac{1}{2}, \frac{1}{2}]$, then $s$ can be recovered from its samples $\{s(k)\}_{k\in\mathbb{Z}}$ by the interpolation formula

$$s(x) = \sum_{i=-\infty}^{i=+\infty} s(i)\mathrm{sinc}(x - i), \tag{1}$$

where $\mathrm{sinc}(x) = \sin(\pi x)/\pi x$.

Another method to discretize signals utilizes the canonical Gabor transforms which represent a signal by a weighted sum of modulated Gaussian functions [9, 26]. These functions are optimally localized in both time (space) and frequency (see Section 6). This implies that the weighting coefficients, which form the discrete representation, are the "best" description of how the different portions of the signal contribute to the various frequency bands (more details in Section 6).

More recently, the concepts of multiresolution analysis and discrete wavelet transforms have offered an infinite variety of discrete representations [4, 9, 15, 19, 21, 24, 27, 34, 37, 40, 45, 49, 53]. Specifically, a signal $s \in L_2$ is represented in terms of dilations and translations of a single wavelet function $\psi$

$$s(x) = \sum_{l\in\mathbb{Z}} \sum_{k\in\mathbb{Z}} 2^{-l/2}\, d_{(l)}(k)\psi(2^{-l}x - k). \tag{2}$$

The coefficients $d_{(l)}(k)$ form the discrete representation. Roughly speaking, each coefficient $d_{(l)}(k)$ corresponds to a time interval of about $[2^l(k - \frac{1}{2}), 2^l(k + \frac{1}{2})]$ and to a frequency band of about $[-2^{-l}, -2^{-l-1}] \cup [2^{-l-1}, 2^{-l}]$. Thus, this representation has the desirable property of being localized in both the time (space) and the frequency domains. However, no wavelet representation can achieve the optimal time-frequency localization of a canonical Gabor transform [8, 10, 22].

In this paper, we present a unified view of digital signal representation theories. First, we construct a set of scaling functions in which the convolution product is an internal law of composition. This set and the set of corresponding wavelets are then partitioned using convolution operators associated with the discrete space $l_2$. Second, we construct sequences of scaling functions $\varphi^n$ and wavelets $\psi^n$ whose regularity increases with $n$. In particular, we study the properties of the basic, orthogonal, interpolating, and dual-interpolating sequences. We then use them to create "Shannon-type" and "Gabor-type" scaling and wavelet functions. A simplified version of these results appears, without proofs, in our Chapter [3].

This paper is organized as follows. In Section 2, we introduce some definitions and notation. In Section 3, we construct a class of scaling functions and wavelets in which the convolution is an internal law of composition for the scaling functions. We then study the algebraic properties induced by this internal convolution product and by the external convolution product between a function from the class and a discrete sequence $\{p(k)\}_{k\in\mathbb{Z}}$ from $l_2$. In Section 4, we introduce the $n$-fold convolution sequences of scaling functions and their associated wavelets. These are the natural extensions of the B-spline scaling functions and the B-spline wavelets of order $n$ [17, 50]. In Section 5, we review the approximation-sampling procedures, which are extensions of the classical sampling procedure of Shannon [2]. We then relate these techniques to the multiresolution analysis of $L_2$. In particular, we show that the pre- and post-filters associated with the orthogonal, interpolating, and dual-interpolating sequences of scaling functions (resp., wavelets) converge to the ideal lowpass filter (resp., bandpass filter). In Section 6, we show that the basic sequences converge to the real part of the canonical Gabor functions as $n$ goes to infinity. This further emphasizes the link between the different representations and allows us to construct wavelets that are nearly time-frequency optimal (see Section 6).

## 2. DEFINITIONS AND NOTATION

The signals considered here are real valued functions defined on $\mathbb{R}$. They belong to the space of measurable, square-integrable functions $L_2$ with the usual norm $\|\cdot\|_{L_2}$. We also consider the Sobolev spaces $W^{r,p}$ which consist of $L_p$ functions with $r$ distributional derivatives in $L_p$ [1]. The space of real-valued square-summable sequences (discrete functions) is denoted by $l_2$.

The symbol "$*$" will be used for three slightly different binary operations that are defined below: the convolution, the mixed convolution, and the discrete convolution. The ambiguity should be easily resolved from the context.
For two functions $f$ and $g$ defined on $\mathbb{R}$, $*$ denotes the usual convolution

$$(f*g)(x) = \int_{-\infty}^{+\infty} f(\xi)g(x - \xi)\, d\xi \qquad x \in \mathbb{R} \tag{3}$$

The mixed convolution between a sequence $\{b(k)\}_{k\in\mathbb{Z}}$ and a function $f$ defined on $\mathbb{R}$ is the function $b*f$ defined on $\mathbb{R}$ that is given by

$$(b*f)(x) = \sum_{k=-\infty}^{k=+\infty} b(k)f(x - k) \qquad x \in \mathbb{R} \tag{4}$$

The discrete convolution between two sequences $a$ and $b$ is the sequence $a*b$ given by

$$(a*b)(l) = \sum_{k=-\infty}^{k=+\infty} a(k)b(l - k) \qquad l \in \mathbb{Z} \tag{5}$$

A filter $\hat{\lambda}(f)$ is the Fourier transform of a function $\lambda$ that defines a bounded convolution operator on $L_2$:

$$\lambda : g \in L_2 \rightarrow \lambda * g \in L_2 \tag{6}$$

Since the convolution product $\lambda * g$ becomes a multiplication product $\hat{\lambda}\hat{g}$ in Fourier space, the filter $\hat{\lambda}$ selectively alters the frequency components of $\hat{g}$.
Whenever it exits, the convolution inverse $(b)^{-1}$ of a sequence $b$ is defined to be

$$\big((b)^{-1}*b\big)(k) = \delta_0(k) \tag{7}$$

where $\delta_i$ is the unit impulse located at $i$, i.e., $\delta_i(i) = 1$, and $\delta_i(k) = 0$ for $k \neq i$. The reflection of a function $f$ (resp., a sequence $b$) is the function $f^\vee$ (resp., the sequence $b^\vee$) given by

$$f^\vee(x) = f(-x) \qquad \forall x \in \mathbb{R} \tag{8}$$

$$b^\vee(k) = b(-k) \qquad \forall k \in \mathbb{Z} \tag{9}$$

The modulation $\tilde{b}(k)$ of a sequence $b$ is obtained by changing the signs of the odd components of $b$:

$$\tilde{b}(k) = (-1)^k b(k) \tag{10}$$

The decimation (or down-sampling) operator by a factor of two $[\cdot]_{\downarrow 2}$ assigns to a sequence $b$ the sequence $[b]_{\downarrow 2}$ which consists of the even components of $b$ only:

$$[b]_{\downarrow 2}(k) = b(2k) \qquad \forall k \in \mathbb{Z} \tag{11}$$



## 3. MULTIRESOLUTIONS OF $L_2$

A multiresolution of $L_2$, introduced by S. Mallat and Y. Meyer [36, 39], is a set of subspaces $\{V_{(j)}\}_{j\in\mathbb{Z}}$ of $L_2$ that are generated by dilating and translating a single function $\varphi_0$:

$$V_{(j)}(\varphi_0) = \left\{ v : v(x) = \sum_{k\in\mathbb{Z}} 2^{-j/2} c_{(j)}(k)\varphi_0(2^{-j}x - k),\ c_{(j)} \in l_2 \right\}. \tag{12}$$

The spaces $V_{(j)}$ induced by the orthonormal basis $\{2^{-j/2}\varphi_0(2^{-j}x - k)\}_{k\in\mathbb{Z}}$ must satisfy the conditions:

i) $\text{Clos}\,(V_{(0)}) = V_{(0)} \subset L_2$

ii) $V_{(j)} \subset V_{(j-1)}$

iii) $\text{Clos}\left( \bigcup_{j\in\mathbb{Z}} V_{(j)} \right) = L_2$

iv) $\bigcap_{j\in\mathbb{Z}} V_{(j)} = \{0\}$

The function $\varphi_0$ in (12) is called an orthogonal scaling function.

Obviously, if $V_{(i)} \subset V_{(j)}$, then the orthogonal projections $s_{(i)}$ of a function $s$ in $L_2$ is a coarser approximation of $s$ than the projection $s_{(j)}$. Thus, Properties (i), (ii) and (iii) allow us to approximate a function to any desired accuracy by taking an appropriate approximation space $V_{(j)}$. In contrast, Property (iv) states that coarser and coarser approximations will eventually lose all the information about the original function.

The orthogonal complement of $V_{(j)}$ relative to $V_{(j-1)}$ is the associated wavelet space $W_{(j)}$:

$$W_{(j)} \oplus V_{(j)} = V_{(j-1)} \tag{13}$$

It is generated by the orthonormal basis $\{2^{-j/2}\psi_0(2^{-j}x - k)\}_{k\in\mathbb{Z}}$ induced by the orthogonal wavelet function $\psi_0$ [36, 40]:

$$W_{(j)}(\psi_0) = \left\{ w: w(x) = \sum_{k\in\mathbb{Z}} 2^{-j/2}\, d_{(j)}(k)\psi_0(2^{-j}x - k),\ d_{(j)} \in l_2 \right\} \tag{14}$$

The difference between the two approximations $s_{(j)}$ and $s_{(j-1)}$ of a function $s$ is equal to the orthogonal projection of $s$ on $W_{(j)}$. From Property (iii), it follows that they constitute a direct sum of $L_2$:

$$L_2 = \bigoplus_j W_{(j)} := \cdots \oplus W_{(1)} \oplus W_{(0)} \oplus W_{(-1)} \oplus \cdots \tag{15}$$

Thus, the discrete wavelet representation of a function $s$ consists of the sequences $\{\cdots d_{(-1)}, d_{(0)}, d_{(1)}, \cdots\}$ (in Eq. (14)) which are obtained from the orthogonal projection of $s$ on $\{W_{(j)}\}_{j\in\mathbb{Z}}$.

*Remark 3.1*: The only wavelet functions that we will consider here are those that are associated with multiresolution structures. Another weak definition can be found in [35].



### 3.1. Equivalent Classes of Scaling Functions and Wavelets

A function $\varphi$ is a scaling function if $\{\varphi(x - k)\}_{k\in\mathbb{Z}}$ is an unconditional Riesz-Schauder basis of $V_{(0)}$. The set $\{\varphi(x - k)\}_{k\in\mathbb{Z}}$ is not required to be orthogonal. The only requirement is that $\varphi$ must generate the spaces $V_{(j)}$ as in (12) by replacing $\varphi_0$ by $\varphi$ [15]. Clearly, two nonidentical scaling functions $\varphi_1 \neq \varphi_2$ may generate the same multiresolution $V_{(j)}(\varphi_1) = V_{(j)}(\varphi_2)$. Depending on the application, a particular choice of $\varphi$ may have advantages over others (e.g., axially symmetric, interpolating, time-frequency localized, . . .).

The set of scaling functions can be partitioned into equivalence classes by the relation $\approx$ that associates two functions whenever they generate the same space. In fact, given a scaling function $\varphi$, we can use the mixed convolution (4) to construct an equivalent scaling function $\varphi_\approx$. This is done by convolving $\varphi$ with a sequence $p$:

$$\varphi_\approx(x) = (p * \varphi)(x) \tag{16}$$

where $p$ is an invertible convolution operator from $l_2$ into itself. A possible sequence $p$ is one which has a Fourier transform $P(f)$ satisfying $P(f) \in L_\infty$, and $\operatorname{essinf}|P(f)| = m > 0$. The first condition implies that $p$ is in $l_2$. The second condition implies that $(P(f))^{-1}$ is in $L_2$ [0, 1]. Therefore, it is the Fourier transform of a sequence $(p)^{-1}$ which is the convolution inverse of $p$ (i.e., $p * (p)^{-1} = \delta_0$). Using Plancherel Theorem, it then follows that if $c \in l_2$, then

$$m^2\|c\|_{l_2}^2 \le \|p * c\|_{l_2}^2 = \int_0^1 |P(f)C(f)|^2\, df \le \|P\|_\infty^2\, \|c\|_{l_2}^2 \tag{17}$$

Thus we have the simple proposition:

**Proposition 1:** Let $p(k)$ be a sequence and $P(f)$ its Fourier transform. If $P(f) \in L_\infty$ and if $\operatorname{essinf}|P(f)| > 0$ then $p$ defines a convolution operator on $l_2$ which is bounded and which has a bounded convolution inverse $(p)^{-1}$.

*Remark 3.2*: Clearly, if $P(f)$ is continuous and nonzero on [0, 1], then $p$ satisfies the conditions of the proposition above. The converse of Proposition 1 is also true [2].

Equation (16) allows us (by choosing a bounded invertible convolution operator on $l_2$) to create different scaling functions generating the same spaces $V_{(j)}$. In fact, if $\varphi_1$ and $\varphi_2$ are two scaling functions for the same multiresolution $V_{(j)}$, then there exist two sequences $p \in l_2$ and $q \in l_2$ such that $\varphi_2(x)$ and $(p * \varphi_1)(x)$, and that $\varphi_1(x) = (q * \varphi_2)(x)$. It follows that $\varphi_1(x) = (q * p * \varphi_1)(x)$. This implies that $p * q = \delta_0$. Thus, $q = (p)^{-1}$ is the convolution inverse of $p$. Moreover, since $\{\varphi_1(x - k)\}_{k\in\mathbb{Z}}$ and $\{\varphi_2(x - k)\}_{k\in\mathbb{Z}}$ are unconditional bases of $V_{(0)}$ we have that

$$m_1\|p * c\|_{l_2} \leq \|(c * p * \varphi_1)(x)\|_{L_2} = \|(c * \varphi_2)(x)\|_{L_2} \leq M_2\,\|c\|_{l_2} \tag{18}$$

where $m_1$ and $M_2$ are two positive constants. Thus we have proven:

**Proposition 2:** The functions $\varphi_1$ and $\varphi_2$ are two scaling functions generating the same multiresolution structure $\{V_{(j)}\}_{j\in\mathbb{Z}}$, if and only if there exist two sequences, $p \in l_2$ and its convolution inverse $(p)^{-1} \in l_2$, defining bounded convolution operators on $l_2$ such that

$$\varphi_2(x) = (p * \varphi_1)(x) \tag{19}$$

*Remark 3.3*: If the scaling functions (resp., wavelets) are required to have compact support or to have exponential decay, then $p$ must be restricted. For this case, necessary and sufficient conditions on $p$ have been established by C. K. Chui and J. Z. Wang in [16].

All of our results for the scaling functions are also valid for the wavelets. Thus, they can be partitioned into equivalent classes. Moreover, given a wavelet $\psi$ generating the spaces $W_{(j)}$, we can construct an equivalent wavelet $\psi_{\approx}$ simply by replacing $\varphi$ in Equation (16) by $\psi$. Clearly, Proposition 2 is also valid for wavelet functions.

### 3.2. The Basic Scaling Function

A constructive method to obtain a multiresolution is to start with a scaling function $\varphi$ and use Equation (12) to define the subspaces $V_{(j)}(\varphi)$. Obviously, the scaling function cannot be chosen arbitrarily. In fact, since $V_{(1)} \subset V_{(0)}$, there must be a sequence $u(k)$ that relates the dilated function $\varphi(x/2)$ and the scaling function $\varphi$:

$$2^{-1/2}\varphi\left(\frac{x}{2}\right) = (u * \varphi)(x) \tag{20}$$

where "$*$" is the mixed convolution defined by (4). The sequence $u(k)$ in (20) is called the generating sequence and Relation (20) is usually referred to as the two-scale relation.

Sufficient conditions on $u(k)$ and its Fourier transform $U(f)$, as well as a method of constructing the orthogonal scaling function, can be found in a theorem by S. Mallat [36, Theorem 2]. In order to obtain a class of scaling functions in which the convolution is an internal law of composition, we have to devise a variation.

**Theorem 3:** Let $u(k)$ be a sequence such that $u(k) = O(k^{-2})$ and let $U(f) = \Sigma_{k\in\mathbb{Z}}\, u(k)e^{-i2\pi kf}$ denotes its Fourier transform. If $U(f)$ satisfies the conditions

$$2^{-1/2}|U(0)| = 1 \tag{21}$$

$$U(f) \neq 0 \qquad \forall f \in \left[-\frac{1}{4}, \frac{1}{4}\right] \tag{22}$$

$$L(f) := \prod_{i=1}^{\infty} 2^{-1/2} U\left(\frac{f}{2^i}\right) = O(|f|^{-r}) \qquad r > \frac{1}{2} \tag{23}$$

then $L(f)$ is the Fourier transform of a function $\varphi$ that generates a closed subspace $V_{(0)}(\varphi)$ of $L_2$. Furthermore, if $L(f) \in C^2(\mathbb{R})$ and if $D^{(i)}L(f) = O(|f|^{-s})$ $i = 0, 1, 2$ with $s > 2$, then $\varphi$ is a scaling function generating the multiresolution $V_{(j)}(\varphi)$.

The conditions in Theorem 3 are the same as those given by Mallat in [36, Theorem 2] except that we have removed the Quadrature Mirror Filter requirement, added the growth Condition (23), and replaced his regularity conditions by some smoothness and growth in the Fourier domain. The fact that the infinite product in (23) makes sense is a direct consequence of a result by I. Daubechies [20, Lemma 3.1]. Let $D^{(n)}$ denote the $n$th derivative. Condition (23) and the other smoothness and growth conditions on the Fourier transform $L(f)$ are regularity requirements. They imply that $\varphi$ satisfies the decay conditions $|\varphi(x)| \leq C_1(1 + x^2)^{-1}$ and $|D^{(1)}\varphi(x)| \leq C_2(1 + x^2)^{-1}$. For faster decay, further smoothness and decay conditions must be imposed on $L(f)$. Sufficient conditions on the sequence $u$ to obtain a desired decay on $L(f)$ have been established by I. Daubechies [20].

We will call the scaling function $\varphi$ obtained from Theorem 3, the basic scaling function (sometimes we also denote it $\varphi_b$). The set of functions $\{\varphi(x - k)\}_{k\in\mathbb{Z}}$ are not necessarily orthogonal. If $U(f)$ is chosen to be real symmetric then $\varphi$ will also be real and symmetric. Moreover, if $u(k)$ has finitely many nonzero values then $\varphi$ will have compact support, i.e., $\varphi(x)$ will be zero outside a closed bounded set [20].

### 3.2.1. Proof of Theorem 3

The proof of Theorem 3 is based on the lemma and the proposition below.

**Lemma 4:** If the Fourier transform $L(f)$ of a real function is such that $L(f) \in C^n(\mathbb{R})$, $L(f) \neq 0\ \forall f \in [-\frac{1}{2}, \frac{1}{2}]$, and $D^{(j)}L(f) = O(|f|^{-r})$ $j = 0, 1, \ldots n$ with $r > \frac{1}{2}$, then the periodic function $(\Sigma_k\, |L(f - k)|^2)^{-1/2}$ belongs to $C^n(\mathbb{R})$.

*Proof*: Let $L(f) = X(f) + iY(f)$. The decay conditions $D^{(j)}L(f) = O(|f|^{-r})$ $j = 0, 1, \ldots n$ with $r > \frac{1}{2}$ imply that $D^{(j)}X(f) = O(|f|^{-r})$ and that $D^{(j)}Y(f) = O(|f|^{-r})$ $j = 0, 1, \ldots n$.
The derivative $D^{(1)}|L(f)|^2$ is given by

$$D^{(1)}|L(f)|^2 = 2X(f)D^{(1)}X(f) + 2Y(f)D^{(1)}Y(f). \tag{24}$$

From the equality above, we conclude that $D^{(1)}|L(f)|^2 = O(|f|^{-2r})$. Similarly, by taking the derivatives of Eq. (24), we get that $D^{(j)}|L(f)|^2 = O(|f|^{-2r})$ $j = 0, 1, \ldots n$. It follows that the series

$$\sum_{k\in\mathbb{Z}} D^{(j)}|L(f-k)|^2 \leq \text{Const} \sum_{k=1}^{\infty} k^{-2r} \tag{25}$$

converges uniformly for $j = 0, 1, \ldots n$. Therefore, the periodic function $\Sigma_k |L(f-k)|^2$ belongs to $C^n(\mathbb{R})$. Since $L(f) \neq 0$ $\forall f \in [-\tfrac{1}{2}/\tfrac{1}{2}]$ we conclude that $\Sigma_k |L(f-k)|^2 > 0$ and the result follows.

**Proposition 5:** The function $L(f)$ in Theorem 3 satisfies

$$\begin{aligned} 2\sum_i |L(f-i)|^2 &= \left|U\left(\frac{f}{2}\right)\right|^2 \sum_i \left|L\left(\frac{f}{2}-i\right)\right|^2 \\ &+ \left|U\left(\frac{f-1}{2}\right)\right|^2 \sum_i \left|L\left(\frac{f-1}{2}-i\right)\right|^2 \end{aligned} \tag{26}$$

*Proof*: The change of variable $x$ to $x/2$ gives that for any function $\lambda \in L_2$, we have the identity

$$\frac{1}{2}\left\langle \lambda\left(\frac{x}{2}\right), \lambda\left(\frac{x}{2}-k\right)\right\rangle_{L_2} = \langle\lambda(x), \lambda(x-k)\rangle_{L_2} \tag{27}$$

Clearly, Equation (23) implies that $L(2f) = 2^{-1/2}U(f)L(f)$. From ths equality, and Equality (27), we obtain using the Fourier transform and simple substitutions

$$\int_{\mathbb{R}} |U(f)L(f)|^2 e^{-i2\pi 2kf}\, df = \int_{\mathbb{R}} |L(f)|^2 e^{-i2\pi kf}\, df \tag{28}$$

We make the change of variable $f' = 2f$ in the integral on the left hand side of (28) and then rewrite (28) by decomposing $\mathbb{R}$ into intervals of length one to get

$$\frac{1}{2}\int_0^1 \sum_{j\in\mathbb{Z}} |U((f-j)/2)L((f-j)/2)|^2 e^{-i2\pi kf}\, df = \int_0^1 \sum_{j\in\mathbb{Z}} |L(f-j)|^2 e^{-i2\pi kf}\, df \tag{29}$$

Equality (26) then follows from (29), the fact that $U(f) = U(f+1)$ and the fact that $\{e^{-i2\pi kf}\}_{k\in\mathbb{Z}}$ is a basis of $L_2[0, 1]$.

*Proof of Theorem 3*: Using Lemma 4, the decay conditions on $L(f)$ imply that the periodic function $P(f)$ given by

$$P(f) = \left(\sum_i |L(f-i)|^2\right)^{-1/2} \tag{30}$$

is $C^2(\mathbb{R})$. Therefore, $P(f)$ is the Fourier transform of an $l_2$ sequence $p$. We consider the function

$$\hat{\phi}(f) = P(f)L(f) \tag{31}$$

From the expression of $L(f)$, $\hat{\phi}(f)$ can be rewritten as

$$\hat{\phi}(f) = \prod_{i=1}^{n} 2^{-1/2} U_\phi\left(\frac{f}{2^i}\right) \tag{32}$$

where $U_\phi(f) = U(f)P(2f)/P(f)$. Using (21) and (22), it follows that $|U_\phi(0)| = 1$ and $U_\phi(f) \neq 0$, $\forall f \in [-\frac{1}{4}, \frac{1}{4}]$. Moreover, using the expression of $P(f)$ given by (30) and Proposition 5 we get that

$$|U_\phi(f)|^2 + \left|U_\phi\left(f - \frac{1}{2}\right)\right|^2 = 2 \tag{33}$$

The smoothness of $P(f)$ together with the smoothness and decay properties of $L(f)$ imply that $\hat{\phi}(f) = P(f)L(f)$ is $C^2$ smooth and has the same decay properties as $L(f)$. The decay of $\hat{\phi}(f)$ and its two derivatives implies that $D^{(j)}\hat{\phi}$ and $D^{(j)}(if\hat{\phi})$ are in $L_1$ for $j = 0, 1, 2$. We take the inverse Fourier transform of $\hat{\phi}$ and integrate by parts twice to get

$$\phi(x) = -(2\pi x)^{-2} \int_{\mathbb{R}} \hat{\phi}''(x) e^{-i2\pi f x}\, df \tag{34}$$

Since $\hat{\phi}$ belongs to $L_1$, $\phi$ is continuous. From this, the fact that $\phi''$ is in $L_1$, and (34) we conclude that $|\phi(x)| \leq C_1(1 + x^2)^{-1}$. A similar argument implies that $|D^{(1)}\phi(x)| \leq C_2(1 + x^2)^{-1}$. These properties of $\phi$ and the properties of $U_\phi$ that we have derived are precisely the conditions needed in the Construction Theorem of Mallat to obtain an orthogonal scaling function $\phi$ [36, Theorem 2]. Thus $\phi$ generates a multiresolution. Proposition 2 and Equation (31) then imply that $\varphi$ generates the same multiresolution.

### 3.3. The Basic Wavelet

From the results of Subsection 3.1, we know that there are infinitely many wavelets that generate the wavelet spaces $W_{(j)}$ associated with the multiresolution $V_{(j)}(\varphi)$. One such wavelet is the basic wavelet $\psi_b$ which we define in the following proposition:

**Proposition 6:** If $\varphi$ is a scaling function generating $V_{(j)}(\varphi)$, then the wavelet spaces $W_{(j)}$ can be generated by the wavelet function $\psi_b$ given by

$$2^{-1/2}\psi_b\left(\frac{x}{2}\right) = \sum_{k=-\infty}^{k=-\infty} (\delta_1 * \tilde{u}^\vee * \tilde{a})(k)\varphi(x - k) \tag{35}$$

where $a$ is the sampled autocorrelation function of $\varphi$:

$$a(k) = (\varphi * \varphi^\vee)(k) \qquad k \in \mathbb{Z} \tag{36}$$

The symbol "~" denotes the modulation operator as defined by (10), and the symbol "∨" denotes the reflection operators as defined by (8) and by (9).

The well-known results of Daubechies [20] imply that $\psi_b$ has compact support whenever the generating sequence $u$ has finitely many nonzero values. Moreover, if the generating sequence $u$ has a vertical axis of symmetry (e.g., $u(k) = u(-k)$) then $\psi_b$ also has a vertical axis of symmetry. In particular, if $\varphi$ is the B-spline of order $n$, then we obtain the B-spline wavelet [17, 50]. Another special case is the generalized B-spline wavelet of Chui and Wang which corresponds to the situations in which $u$ has finitely many nonzero values [16].

*Proof*: From the proof of Theorem 3, $\hat{\phi}$ as given by Equations (30) and (31) is the Fourier transform of an orthogonal scaling function $\phi$ corresponding to

$V_{(j)}(\phi) = V_{(j)}(\varphi)$. Thus, the Fourier transform $\hat{\phi}$ of an orthogonal wavelet $\psi_0$ is given by

$$\hat{\psi}_0(f) = 2^{-1/2} e^{-i\pi f}\, \overline{U_\phi\left(\frac{f}{2} - \frac{1}{2}\right)}\, \hat{\phi}\left(\frac{f}{2}\right) \tag{37}$$

where as before, $U_\phi(f) = U(f)P(2f)/P(f)$, and where $\overline{U_\phi}$ is the complex conjugate of $U_\phi$ [36]. Taking the Fourier transform of (36) and using Poisson's formula we find that $\hat{a}(f) = (P(f))^{-2}$. Hence, the Fourier transform of the basic wavelet $\psi_b(x)$ defined in (35) is given by

$$\hat{\psi}_b(f) = 2^{-1/2} e^{-i\pi f}\, \overline{U\left(\frac{f}{2} - \frac{1}{2}\right)} \left(P\left(\frac{f}{2} - \frac{1}{2}\right)\right)^{-2} L\left(\frac{f}{2}\right) \tag{38}$$

The ratio between $\hat{\psi}_b(f)$ and $\hat{\psi}_0(f)$ is then given by

$$Q(f) = \left(P\left(\frac{f}{2}\right) P(f) P\left(\frac{f}{2} - \frac{1}{2}\right)\right)^{-1} \tag{39}$$

Since $P(f)$ is continuous and positive, $Q(f)$ is well defined. Moreover, since $P(f) = P(f + 1)$ we have that $P(f/2 + \frac{1}{2}) = P(f/2 - \frac{1}{2})$, from which we deduce that $Q(f) = Q(f + 1)$. Thus $Q(f)$ is periodic with period 1 and, by Proposition 1, it is the Fourier transform of an invertible convolution operator $q$. We have

$$\psi_b(x) = (q * \psi_0)(x). \tag{40}$$

### 3.4. Special Bases of Scaling Functions and Wavelets

In all the constructions below, $u$ will denote the generating sequence of $\varphi$ as in (20). The sequence $\{a(k)\}_{k\in\mathbb{Z}}$ is the sampled autocorrelation of $\varphi$ as defined by Equation (36).

#### 3.4.1. Orthogonal Scaling Functions

An orthogonal scaling function $\varphi_0$ must satisfy

$$\langle \varphi_0(x), \varphi_0(x - k)\rangle = (\varphi_0 * \varphi_0^\vee)(k) = \begin{cases} 1 & k = 0 \\ 0 & k = \pm 1, \pm 2, \ldots \end{cases} \tag{41}$$

To obtain such a function from an arbitrary scaling function $\varphi$, we can choose $p = p_0$ in (19) to be the inverse of the operator-square-root of the sampled autocorrelation function $a(k)$ defined by (36):

$$p_0 = (a)^{-1/2} \tag{42}$$

This is the inverse Fourier transform of

$$P_0(f) = \left(\sum_i |L(f - i)|^2\right)^{-1/2} \tag{43}$$

Of course, there are many other orthogonal generating function associated with $V_{(j)}(\varphi)$.

### 3.4.2. Orthogonal Wavelets

Similarly, an orthogonal wavelet $\psi_0$ must also satisfy the orthogonality condition (41). Using the sampled autocorrelation function $\{a(k)\}_{k\in\mathbb{Z}}$, we obtain $\psi_0$ from an arbitrary basic wavelet $\psi_b$ as follows:

$$\psi_0(x) = q_0 * \psi_b \tag{44}$$

where

$$q_0 = (a * [\tilde{a} * a]_{\downarrow 2})^{-1/2} \tag{45}$$

The operator $[\cdot]_{\downarrow 2}$ is defined by (11) and $\tilde{a}$ is the modulation of $a$ as defined by (10). Equation (45) is derived using (41), the fact that $[\hat{h}]_{\downarrow 2} = [h]_{\downarrow 2}$ for all sequences $h$, and the identity

$$[u * a * u^{\vee}]_{\downarrow 2} = a \tag{46}$$

Identity (46) can be deduced by taking the Fourier transform of (26) in Proposition 5.

### 3.4.3. Interpolating Scaling Functions

A function $\varphi_I$ which interpolates between samples on $\mathbb{Z}$ must have a value of zero at all integers except at the origin where it must be 1:

$$\varphi_I(k) = \delta_0(k) = \begin{cases} 1 & k = 0 \\ 0 & k = \pm 1, \pm 2, \ldots \end{cases} \tag{47}$$

In the literature, such a function is sometimes called *fundamental* or *cardinal*. If $\varphi$ is a scaling function satisfying the conditions of Theorem 3, and if $P(f) = |\Sigma_i L(f - i)| \neq 0$ for all $f \in [0, 1]$, then we can choose a sequence $p_I$ so that $\varphi_I = p_I * \varphi$ is an interpolating function. The sequence $p_I$ can be specified by its Fourier transform as

$$P_I(f) = \left(\sum_i L(f - i)\right)^{-1} \tag{48}$$

From Poisson's formula, the sequence $p_I(k)$ is simply the convolution inverse of the sequence $b(k)$ given by

$$b(k) = \varphi(k) \tag{49}$$

### 3.4.4. Interpolating Wavelets

The interpolating wavelet $\psi_I$ must be such that $\psi_I(k + \frac{1}{2}) = \delta_0(k)$ i.e., $\psi_I(\frac{1}{2}) = 1$, $\psi_I(k + \frac{1}{2}) = 0$ for all $k \in \mathbb{Z}\backslash\{0\}$. We have chosen a slightly different characterization from the one for $\varphi_I$. Specifically, a function $w \in W_{(0)}$ can be expanded in terms of its samples $\{w(k + \frac{1}{2})\}_{k\in\mathbb{Z}}$ as

$$w(x) = \sum_{k\in\mathbb{Z}} w\left(k + \frac{1}{2}\right) \psi_I(x - k) \tag{50}$$

This definition of the interpolating wavelet $\psi_I$ allows us to obtain interpolating wavelets with axial symmetry (see end of this section). To obtain $\psi_I$ from an arbitrary basic wavelet $\psi_b$ we choose an appropriate sequence $q_I$ as follows:

$$q_I = 2^{-1/2} ([\tilde{u}^{\vee} * \tilde{a} * b]_{\downarrow 2})^{-1} \tag{51}$$

where $b$ is as in (49). The interpolating wavelet $\psi_I$ is then given by

$$\psi_I(x) = q_I * \psi_b. \tag{52}$$

If the generating sequence $u$ is symmetric, then the associated basic function $\varphi$ is symmetric. From its expression (35), the corresponding basic wavelet $\psi_b(x)$ has an axis of symmetry at $x = \frac{1}{2}$. For this case, the wavelet $\psi_I(x)$ defined by (52) has an axis of symmetry at $x = \frac{1}{2}$ since, from the expression (51), $q_I$ is clearly symmetrical.

## 4. SEQUENCES OF SCALING FUNCTIONS AND WAVELETS

The convolution product is an internal law of composition for the set of functions in Theorem 3. In particular, if $u_1$ and $u_2$ are two generating sequences corresponding to the scaling functions $\varphi_1$ and $\varphi_2$, then $u = 2^{-1/2} u_1 * u_2$ corresponds to $\varphi_1 * \varphi_2$. To see this, we simply note that the product $\hat{\varphi}_1(f)\hat{\varphi}_2(f)$ satisfies the decay and smoothness conditions of Theorem 3. Accordingly, we have:

**Proposition 7:** If $\varphi_1$ and $\varphi_2$ are two scaling functions generated by the procedure in Theorem 3, then $\varphi_1 * \varphi_2$ is also a scaling function.

If the Fourier transforms $L(f)$ of $\varphi$ are such that $L(f) = O(|f|^{-r})$, then $\varphi$ belongs to the Sobolev spaces $W^{s,2}$ for any $s < r - \frac{1}{2}$. Thus, if the Fourier transforms $L_1(f)$ and $L_2(f)$ of $\varphi_1$ and $\varphi_2$ are such that $L_1(f) = O(|f|^{-r_1})$, and $L_2(f) = O(|f|^{-r_2})$, then $\varphi_1 * \varphi_2$ belongs to the Sobolev space $W^{s,2}$ for any $s < r_1 + r_2 - \frac{1}{2}$. It follows that the convolution $\varphi_1 * \varphi_2$ is more regular than its individual constitutive atoms $\varphi_1$ and $\varphi_2$. This fact combined with Propositions 1 and 2 allows us to obtain sequences of increasingly regular scaling functions and wavelets with other prescribed properties.

*Remark 4.1*: Note that, in general, the convolution of two wavelet functions is not a wavelet (see Remark 3.1).

### 4.1. The Basic Scaling Sequence

Starting from a scaling function $\varphi$, we define the basic sequence $\varphi^n$ (or $\varphi_b^n$) by

$$\varphi^n(x) = \varphi * \varphi * \varphi * \cdots * \varphi \qquad (n - 1 \text{ convolutions}) \tag{53}$$

The corresponding generating sequence $u^n$ is given by

$$u^n = 2^{-(n-1)/2}\, u * u * \cdots * u \qquad (n - 1 \text{ discrete convolutions}) \tag{54}$$

where $u$ is the generating sequence associated with $\varphi$ and where $u^n$ in (54) is the generating sequence for $\varphi^n$ (see Theorem 3).

### 4.2. The Basic Wavelet Sequence

For each function $\varphi^n$, there is a corresponding basic wavelet function $\psi_b^n$ obtained from the construction (35):

$$2^{-1/2}\psi_b^n\left(\frac{x}{2}\right) = \sum_{k=-\infty}^{k=-\infty} (\delta_1 * \tilde{u}^{n\vee} * \tilde{a}^n)(k)\varphi^n(x - k) \tag{55}$$

$$a^n(k) = (\varphi^n * \varphi^{n\vee})(k) \qquad \forall k \in \mathbb{Z} \tag{56}$$

where as before, the symbol "$\tilde{}$" denotes the modulation operator as defined by (10), and the symbol "$\vee$" denotes the reflection operators as defined by (8) and by (9).

*Remark 4.2*: It should be noted that the sampled autocorrelation function $a^n$ is not equal to the $(n - 1)$-fold convolution of $a(k) = (\varphi * \varphi^\vee)(k)$.

### 4.3. The Orthogonal Scaling Sequence

Using the result of Subsection 3.4.1, we obtain the orthogonal scaling sequence

$$\varphi_0^n(x) = ((a^n)^{-1/2} * \varphi^n)(x) \tag{57}$$

$$a^n(k) = (\varphi^n * \varphi^{n\vee})(k) \qquad \forall k \in \mathbb{Z} \tag{58}$$

### 4.4. The Orthogonal Wavelet Sequence

The orthogonal wavelet $\psi_0^n$ is obtained from the basic wavelet $\psi_b^n$ by (cf., Subsection 3.4.2)

$$\psi_0^n(x) = (q_0^n * \psi_b^n)(x) \tag{59}$$

where

$$q_0^n(k) = (a^n * [\tilde{a}^n * a^n]_{\downarrow 2})^{-1/2}\,(k) \qquad \forall k \in \mathbb{Z} \tag{60}$$

### 4.5. The Interpolating Scaling Sequence

Starting from a scaling function $\varphi^n$, we define the discrete function $b^n(k)$ by

$$b^n(k) = \varphi^n(k) \qquad \forall k \in \mathbb{Z} \tag{61}$$

From Subsection 3.4.3, an interpolating sequence $\varphi_I^n$ exists as long as $b^n$ is invertible. The interpolating sequence $\varphi_I^n$ is then given by

$$\varphi_I^n(x) = ((b^n)^{-1} * \varphi^n)(x) \tag{62}$$

### 4.6. The Interpolating Wavelet Sequence

Using the modulation, reflection and down sampling operators of Section 2, we define $q_I^n(k)$ to be

$$q_I^n(k) = 2^{-1/2}([b^n * \tilde{u}^{n\vee} * \tilde{a}^n]_{\downarrow 2})^{-1}(k) \qquad \forall k \in \mathbb{Z} \tag{63}$$

The interpolating wavelet sequence is then given by

$$\psi_I^n(x) = (q_I^n * \psi_b^n)(x) \tag{64}$$

where $\psi_b^n$ is the basic wavelet (see Equation (55)).

### 4.7. Symmetrical Sequences of Scaling and Wavelet Functions

If $\varphi$ is symmetrical, then $a^n(k) = b^{2n}(k) = \varphi^{2n}(k)$ and there is a direct relation between the orthogonal and interpolating scaling sequences $\varphi_I^{2n} = \varphi_0^n * \varphi_0^n$. In this case, the even sequence $b^{2n}$ has an inverse. This is because its Fourier transform is always positive:

$$\hat{b}^{2n}(f) = \left(\sum_i |L(f - i)|^{2n}\right) > 0 \tag{65}$$

Thus, if $\varphi$ is symmetrical, then the existence of an interpolating scaling sequence is guaranteed for even values of $n$. This is also true for the interpolating wavelet sequence $\psi_I^n$. Moreover, the basic wavelets $\psi_b^n(x)$ defined by (55) have an axis of symmetry about $x = \frac{1}{2}$. This also holds for the interpolating wavelets $\psi_I^n(x)$. Finally, it is easy to construct symmetrical scaling functions. For instance, we can start from a symmetrical generating sequence $u(k) = u(-k)$. Alternatively, we can start from any scaling function $\tau$ and obtain by convolution the symmetrical function $\varphi = \tau * \tau^\vee$.

## 5. MULTIRESOLUTION AND WAVELET SEQUENCES IN CONNECTION WITH SHANNON'S SAMPLING PROCEDURE

The classical signal processing paradigm for the discretization of signals is shown in Figure 1A. This scheme is widely used in connection with Shannon's Sampling Theorem when the signals are not bandlimited [13, 48]. In particular, when a function is not bandlimited, it is first prefiltered with the ideal lowpass filter (Equation (66) below) before sampling:

$$\text{rect}(f) = \begin{cases} 1 & |f| \leq \dfrac{1}{2} \\ 0 & \text{elsewhere} \end{cases} \tag{66}$$

The whole procedure prevents aliasing errors. It is equivalent to finding the least squares approximation of the signal in the space $V_{(0)}(\text{sinc})$ generated by $\text{sinc}(x) = \sin(\pi x)/\pi x$:

$$V_{(0)}(\text{sinc}) = \left\{ v : v(x) = \sum_{k=-\infty}^{k=+\infty} c(k)\text{sinc}(x - k) \qquad c \in l_2 \right\} \tag{67}$$

The least squares approximation of a function $s$ in $V_{(0)}(\text{sinc})$ is given by

$$s_{(0)}(x) = \sum_{k=-\infty}^{k=+\infty} (s * \text{sinc})(k)\text{sinc}(x - k) \tag{68}$$

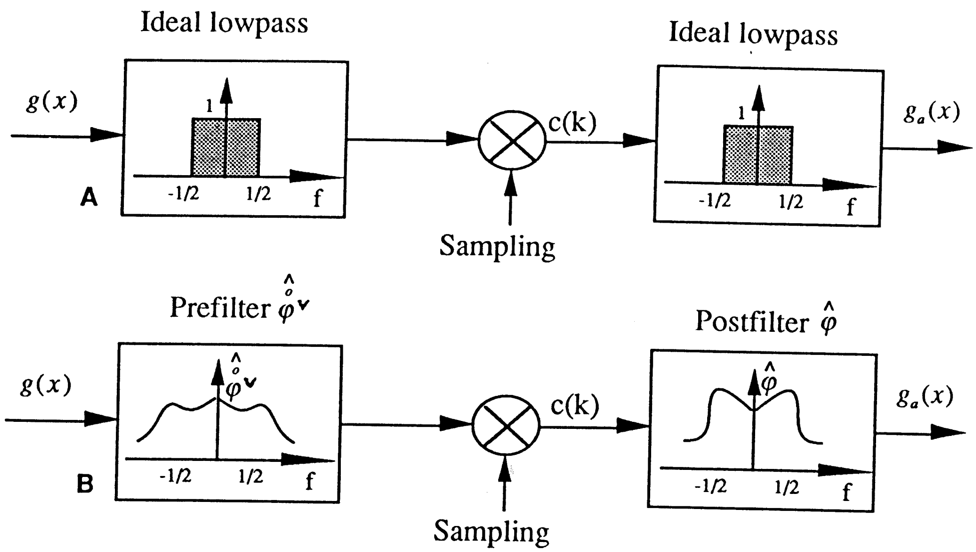


**Figure 1**

We call this procedure approximation-sampling to distinguish it from the interpolation-sampling scheme and its extensions [12, 30, 32, 41, 54, 56, 58] (for a survey see [14, 28, 31]).

The function $(s * \mathrm{sinc})(x)$, $x \in \mathbb{R}$, corresponds to prefiltering $s$ with the ideal lowpass filter. Thus, its Fourier transform has a compact support which lies in $[-\frac{1}{2}, \frac{1}{2}]$. By the Paley-Wiener Theorem, it must be an entire function. Therefore, the samples $c(k) = (s * \mathrm{sinc})(k)$, $k \in \mathbb{Z}$, are well defined. Moreover, a sampling $c_h(k)$ of a translation $s(x + h)$ of $s$ has the property that $c_h(k)$ converges to $c(k)$ as $h$ goes to zero. Because of this property, we say that the sampling scheme is *jitter-stable*. In fact, the convergence of $c_h(k)$ to $c(k)$ is in the $l_2$-norm, and we say that the sampling is *strongly jitter-stable*.

It is not difficult to show that the sinc function, which acts as a convolution operator in this sampling theory, is an orthogonal scaling function. A corresponding orthogonal wavelet is

$$\psi_0 = \cos(1.5\pi x)\mathrm{sinc}(0.5x) \tag{69}$$

It is the inverse Fourier transform of the ideal bandpass filter

$$\mathrm{BP}(f) = \mathrm{rect}(2f - 1.5) + \mathrm{rect}(2f + 1.5) \tag{70}$$

### 5.1. Generalized Sampling Procedures for Multiresolutions of $L_2$

The approximation-sampling procedure described in the last section has been extended to finite dimensional signals [42], to polynomial splines [29, 52], and to infinite dimensional "shift invariant" function spaces [2]. In this latter extension, we show that under suitable conditions on a function $\lambda \in L_1$ called the generating function, the least squares approximation on the space $V = \{v{:}v(x) = (c * \lambda)(x),\ c \in l_2\}$ produces a well-defined sampling procedure. This means that $V$ must be a closed subspace of $L_2$, and that the orthogonal projection on $V$ must be implementable as:

i) a shift invariant prefilter
ii) a strongly jitter-stable sampling
iii) a shift invariant postfilter

where the notion of filter is defined in Section 2 (see Equation (6)). The full procedure is described in Figure 1B. It is very similar to the classical scheme shown in Figure 1A. The main difference is that the ideal lowpass filters are replaced by the appropriate pre- and post-filters: $\mathring{\lambda}^{\vee}$ and $\hat{\lambda}$ respectively. The role of the optimal prefilter $\mathring{\lambda}^{\vee}$ is analogous to the anti-aliasing lowpass filter required in conventional sampling theory.

We have shown in [2] that in order to obtain a well defined sampling procedure, the Fourier transform $L(f)$ of $\lambda \in L_1$ must satisfy the decay condition

$$L(f) = O(|f|^{-r}) \qquad r > \frac{1}{2} \tag{71}$$

Moreover, the zeros of $L(f)$ must not be "structured." Specifically, the intersection of the sets $A_i := \{f \in [0, 1]: L(f - i) = 0\}$ must be empty:

$$\bigcap_i A_i = \emptyset \tag{72}$$

A scaling function obtained as in Theorem 3 satisfies all the above properties for $\lambda$. First, the smoothness and decay of $\hat{\varphi}(f)$ imply that $\varphi$ is in $L_1$. Specifically, since $\hat{\varphi} \in C^2(\mathbb{R})$, $\hat{\varphi}(f) = O|f|^{-s})$, and $D^{(1)}\hat{\varphi}(f) = O(|f|^{-s})$ with $s > 2$, we conclude that $\varphi(x) \in L_1$. Also, Conditions (21) and (22) on $U(f)$ imply that $\hat{\varphi}(f) \neq 0$ for all $f \in [-\frac{1}{2}, \frac{1}{2}]$. This in turn implies that the intersection $\bigcap_i A_i$ of the sets $A_i$ associated with $\hat{\varphi}(f)$ satisfies Condition (72): $\bigcap_i A_i = \emptyset$. A similar argument also holds for the basic wavelet function $\psi_b$. We state these simple facts in the following proposition:

**Proposition 8:** A scaling function $\varphi$ satisfying the conditions of Theorem 3 belongs to $L_1(\mathbb{R})$ and satisfies Conditions (71) and (72). The associated basic wavelet $\psi_b$ also belongs to $L_1(\mathbb{R})$ and satisfies Conditions (71) and (72).

As a corollary of this proposition and the results in [2] we immediately obtain:

**Theorem 9:** If $\varphi$ (resp., $\psi$) is a scaling function (resp., wavelet) obtained as in Theorem 3, then the orthogonal projection in $V_{(0)}(\varphi)$ (resp., $W_{(0)}(\psi)$) is a well defined sampling procedure which consists of a prefiltering with $\mathring{\varphi}^{\vee}$ (resp., $\mathring{\psi}^{\vee}$), followed by a strongly jitter-stable sampling, and finally a postfiltering with $\hat{\varphi}$ (resp., $\hat{\psi}$). The function $\mathring{\varphi}$ is given by

$$\mathring{\varphi}(x) = ((a)^{-1} * \varphi)(x) \qquad \forall x \in \mathbb{R} \tag{73}$$

where $a$ is the autocorrelation of $\varphi$ given by

$$a(k) = (\varphi * \varphi^{\vee})(k) \qquad \forall k \in \mathbb{Z} \tag{74}$$

**Remark 5.1:** The function $\mathring{\varphi}$, which is the reflection of the optimal prefilter, is simply the dual scaling function (resp., dual wavelet function) [17]. Particular examples for polynomial splines can be found in [52]. It is not difficult to verify that the functions $\varphi$ and $\mathring{\varphi}$ are biorthogonal, i.e., $\langle\varphi(x), \varphi(x - k)\rangle = \delta_0(k)$ [18].

### 5.2. Asymptotic Results

The sequence of scaling functions and the sequence of basic wavelets defined by (53) and (55) provide us with a set of increasingly regular multiresolution structures $V_{(j)}(\varphi^n)$ and wavelet spaces $W_{(j)}(\psi^n)$. As discussed earlier, there are different choices of basis functions that will characterize the spaces $V_{(j)}(\varphi^n)$ and $W_{(j)}(\psi^n)$. The orthogonal projection on $V_{(0)}(\varphi^n)$ or $W_{(0)}(\psi^n)$ does not change. However, the filters associated with the sampling procedure depend on the choice of the basis. These are related by a discrete convolution operator as described by Proposition 2 and their choice depends on the application. A case of special interest is the orthogonal scaling sequence (57) and its associated orthogonal wavelet sequence (59). Another case of interest is the interpolating scaling sequence (62) and the wavelet sequence (64).

As $n$ increases, the sequences of filters become more regular. The Central Limit Theorem implies that the basic sequence $\varphi^n$ converges to a Gaussian function [25]. In Section 6, we will show that the basic wavelet sequence $\psi_b^n$ tend to Gabor functions (modulated Gaussians) as $n$ tends to infinity. This latter result, however, is not a consequence of the Central Limit Theorem. For instance, the orthogonal and interpolating wavelets $\psi_o^n$ and $\psi_I^n$ generate the same spaces as $\psi_b^n$ but converge to the ideal bandpass filters, as stated in Theorems 13 and 14 below.

We will start by showing that the orthogonal, interpolating, and dual-interpolating scaling sequences (see Remark 5.1), converge to the ideal lowpass filters. These asymptotic properties are consistent with the results of de Boor, Höllig, and Riemenschneider [23, 43, 44]; a detailed discussion of this issue can be found in [2].

#### 5.2.1. Convergence Results for the Scaling Sequences

From Proposition 8 and [2, Theorems 14 and 16], we obtain the asymptotic convergence results of the orthogonal, interpolating, and dual-interpolating scaling sequences.

**Theorem 10:** If $\varphi$ is a symmetrical scaling function satisfying the conditions of Theorem 3, and if its Fourier transform $L(f)$ satisfies

$$\min_{f\in I} |L(f)| > |L(f)| \qquad \forall f \notin I = \left[-\frac{1}{2}, \frac{1}{2}\right] \tag{75}$$

then the Fourier transforms $\hat{\varphi}_I^{2n}(f)$ of the interpolating functions $\varphi_I^{2n}$, and the Fourier transforms $\hat{\mathring{\varphi}}_I^{2n}(f)$ of their duals $\mathring{\varphi}_I^{2n}$ converge pointwise a.e. and in $L_p$-norms, $p \in [1, \infty)$, to the ideal lowpass filter as $n$ tends to infinity:

$$L_p - \lim_{n\to\infty} \hat{\varphi}_I^{2n}(f) = \text{rect}(f) \tag{76}$$

$$L_p - \lim_{n\to\infty} \hat{\mathring{\varphi}}_I^{2n}(f) = \text{rect}(f) \tag{77}$$

**Theorem 11:** If $\varphi$ is a scaling function satisfying the conditions of Theorem 3, and if its Fourier transform $L(f)$ satisfies

$$\min_{f\in I}|L(f)| > |L(f)| \qquad \forall f \notin I = \left[-\frac{1}{2}, \frac{1}{2}\right] \tag{78}$$

then the modulus $|\hat{\varphi}_0^n(f)|$ of the Fourier transform of the orthogonal functions and its dual $\mathring{\varphi}_0^n = (\varphi_0^n)^\vee$ converge pointwise a.e. and in $L_p$-norms, $p \in [1, \infty)$, to the ideal lowpass filter as the order $n$ tends to infinity:

$$L_p - \lim_{n\to\infty} |\hat{\varphi}_0^n(f)| = \text{rect}(f) \tag{79}$$

Using the property that the Fourier transform is a bounded operator from $L_p$ into $L_q$ with $p \in [1, 2]$ and $p^{-1} + q^{-1} = 1$, we immediately get:

**Corollary 12:** If $\varphi$ is symmetrical then the interpolating, dual-interpolating, and orthogonal sequences $\varphi_I^{2n}$, $\mathring{\varphi}_I^{2n}$, and $\varphi_0^n$ converge, in $L_q$-norms, $q \in [2, \infty]$, to the ideal sinc interpolator of Shannon as the order $n$ tends to infinity.

**Remark 5.2:** Since $L(f)$ is a continuous function, Condition (75) (or (78)) implies that the minimum of $|L(f)| = |L(-f)|$ in $I = [-\frac{1}{2}, \frac{1}{2}]$ is achieved at $f = \pm\frac{1}{2}$. Clearly, this minimum can be achieved elsewhere in $I$, and $|L(f)|$ can have other local minima in $[-\frac{1}{2}, \frac{1}{2}]$. An equivalent statement is: $|L(f_1)| > |L(f_2)| \forall f_1 \in I$ and $\forall f_2 \notin I$. Thus, Condition (75) (or (78)), essentially means that $L(f)$ is a nonideal lowpass filter in the frequency band $[-\frac{1}{2}, \frac{1}{2}]$. The theorems can then be viewed as stating that the ideal lowpass filter can be approximated as closely as needed by the sequences $\hat{\varphi}_I^{2n}$, $\mathring{\varphi}_I^{2n}$ or $\hat{\varphi}_0^{2n}$.

### 5.2.2. Convergence of the Orthogonal and Interpolating Wavelet Sequences

As mentioned in Section 5, the bandpass filter BP$(f)$ defined by (70) is the Fourier transform of the orthogonal (also interpolating) wavelet associated with the sinc function. This observation, and the results of the previous subsection, suggest that the interpolating and the orthogonal wavelet sequences converge to the ideal bandpass filter. In fact, this is known for orthogonal spline wavelets [33]. A general result is given by the following theorem:

**Theorem 13:** The orthogonal wavelet sequence $|\hat{\psi}_0^n(f)|$ associated with a scaling function satisfying Condition (78) of Theorem 11 converges pointwise and in $L_p$, $1 \le p < \infty$, to the ideal bandpass filter BP$(f)$ as the order $n$ tends to infinity:

$$L_p - \lim_{n\to\infty} |\hat{\psi}_0^n(f)| = \text{BP}(f) \tag{80}$$

*Proof*: The Fourier transform of the orthogonal scaling function $\varphi_0^n$ satisfies

$$\hat{\varphi}_0^n(2f) = 2^{-1/2} U_{\varphi_0^n}(f)\hat{\varphi}_0^n(f) \tag{81}$$

where $U_{\varphi_0^n}$ is the generating function of the orthogonal sequence $\varphi_0^n(f)$. Since $\hat{\varphi}_0^n(f) \neq 0 \; \forall f \in [-\frac{1}{2}, \frac{1}{2}]$, Equation (81) in conjunction with Theorem 11 implies that

$$\lim_{n\to\infty} |2^{-1/2} U_{\varphi_0^n}(f)| = \text{rect}(2f) \qquad \forall f \in \left(-\frac{1}{2}, \frac{1}{2}\right) \tag{82}$$

The Fourier transform of the orthogonal wavelet sequence $\psi_0^n(f)$ is given by

$$\hat{\psi}_0^n(f) = 2^{-1/2} e^{-i\pi f}\, \overline{U_{\varphi_0^n}\left(\frac{f}{2} - \frac{1}{2}\right)}\, \hat{\varphi}_0^n\left(\frac{f}{2}\right) \tag{83}$$

For $f \in [0, \infty)$, Equations (79), (82) and (83) imply that $|\hat{\psi}_0^n(f)| \rightarrow \text{rect}(2f - 1.5)$ $\forall f \in [0, \infty)$ pointwise as $n \rightarrow \infty$. Since $U_{\varphi_0^n}(f)$ is periodic with period 1 we also have that $|\hat{\psi}_0^n(f)| \rightarrow \text{rect}(2f + 1.5)$ $\forall f \in (-\infty, 0]$. To prove the convergence in $L_p$, we note that since $2^{-1/2}|U_{\varphi_0^n}(f)| \leq 1$ (see Equation (33)), we have

$$R^n(f) := |\hat{\psi}_0^n(f) - \text{BP}(f)|^p \leq \text{Const}\left(\left|\hat{\varphi}_0^n\left(\frac{f}{2}\right)\right|^p + |\text{BP}(f)|^p\right) =: T^n(f) \qquad (84)$$

From Theorem 11 the function $T^n(f)$ on the right side of Inequality (84) converges pointwise and in $L_1$ to $T(f) = \text{rect}(f/2) + \text{BP}(f)$ as $n$ goes to infinity. To finish the proof, we use a standard argument which is a generalization of the Lebesgue Dominated Convergence Theorem. Since the left side of Inequality (84), $R^n(f)$, converges pointwise to zero, Fatou's Lemma then yields

$$\begin{aligned} 0 \leq \int_{\mathbb{R}} T(f)\, df &\leq \liminf_{n\rightarrow\infty} \int_{\mathbb{R}} (T^n(f) - R^n(f))\, df \\ &\leq \int_{\mathbb{R}} T(f)\, df + \liminf_{n\rightarrow\infty} \int_{\mathbb{R}} -R^n(f)\, df \end{aligned} \qquad (85)$$

From (85), we immediately obtain that

$$\limsup_{n\rightarrow\infty} \int_{\mathbb{R}} R^n(f)\, df \leq 0 \qquad (86)$$

from which the $L_p$ convergence of the orthogonal wavelets follows.

If the scaling function $\varphi$ is symmetrical, then the functions $\varphi^{2n}$ are symmetrical and the interpolating wavelet functions $\psi_I^{2n}(x)$ have an axis of symmetry at $x = \frac{1}{2}$. The functions $\psi_I^{2n}$ can be written in terms of the orthogonal functions $\psi_0^{2n}$ using the invertible sequences $b^{2n}(k) = \psi_0^{2n}(k + \frac{1}{2})$. Specifically, $\psi_I^{2n} = q_I^{2n} * \psi_0^{2n}$ where $q_I^{2n} = (b^{2n})^{-1}$. In Fourier space, we have

$$\hat{\psi}_I^{2n}(f) = \left(\sum_i |\hat{\psi}_0^{2n}(f - i)|\right)^{-1} \hat{\psi}_0^{2n}(f) \qquad (87)$$

Since, by Theorem 13, the functions $|\hat{\psi}_0^{2n}| \rightarrow \text{BP}(f)$, we conclude that the series in (87) tends to 1. Thus, $|\hat{\psi}_I^{2n}(f)|$ converges pointwise to the bandpass filter $\text{BP}(f)$ as $n \rightarrow \infty$. Furthermore, if there exists a constant *Const* independent of $n$ such that

$$\left(\sum_i |\hat{\psi}_0^{2n}(f - i)|\right)^{-1} \leq \text{Const} \qquad (88)$$

then the convergence is in $L_p$, $\forall p \in [1, \infty)$. This last assertion follows from an argument similar to the one in the proof of Theorem 13, so that we can state:

**Theorem 14:** If $\varphi$ has an axis of symmetry, then the shifted cardinal wavelet sequence $|\psi_I^{2n}(f)|$ converges pointwise, to the ideal bandpass filter $\text{BP}(f)$ as the order $n$ tends to infinity. Moreover, if Condition (88) is satisfied, then the convergence is in $L_p$, $\forall p \in [1, \infty)$.

## 6. MULTIRESOLUTION SEQUENCES IN CONNECTION WITH GABOR TRANSFORMS

When it is sufficiently regular, a wavelet is localized in time (space) and frequency. This means that the standard deviation of its squared modulus $\sigma_{|\psi|^2}$, and the standard deviation $\sigma_{|\hat{\psi}|^2_+}$ for the positive frequencies in Fourier space, are both finite. Moreover, the Fourier transform of a wavelet $\psi$ has the property that $\hat{\psi}(0) = 0$. Thus, it is essentially a bandpass filter concentrated around some frequencies $\pm f_0$. For definiteness, if $f_0 = \frac{3}{4}$, $\sigma_{|\psi|^2} = 2^{-1}$, and $\sigma_{|\hat{\psi}|^2_+} = 2^{-2}$, then the coefficients $d_{(l)}(k)$, of the wavelet decomposition of a signal $s$ (see Eq. (2)), correspond to a time interval of about $[2^l(k - \frac{1}{2}), 2^l(k + \frac{1}{2})]$ and to a frequency band of about $[-2^{-l}, -2^{-l-1}] \cup [2^{-l-1}, 2^{-l}]$. Thus, a discrete wavelet representation has the desirable property of being localized in both the space (time) and the frequency domains.

A measure of the time and frequency localization of a function is given by the product $\chi = \sigma_{|\hat{\psi}|^2}\sigma_{|\psi|^2}$. It is bounded below by the optimal value $(4\pi)^{-2}$. The only functions for which the equality holds are the canonical Gabor functions [26]:

$$g(x) = \exp(i\Omega(x - x_0) - i\theta)\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x - x_0)^2}{2\sigma^2}\right) \tag{89}$$

These are modulated Gaussian functions with four parameters: the offset $x_0$, the standard deviation $\sigma$, the modulation frequency $\Omega$ and the phase shift $\theta$. Because of their time-frequency optimality, they have led to the Gabor transforms used for signal representations [6, 9, 26]. However, there are no values of the parameters in (89) that can force the corresponding function, its real part or complex part to form a wavelet basis of $L_2$ [8, 11, 22] (see Remark 3.1). Nevertheless, we will show that, under mild conditions, the basic wavelet sequence $\psi_b^n$ defined by (55) converges to the real part of Gabor functions as $n$ goes to infinity. This may appear surprising since the orthogonal and interpolating wavelets $\psi_0^n$ and $\psi_I^n$ generate the same spaces as $\psi_b^n$, yet we have proven in the last section that they converge to modulated sinc functions instead.

We start from a symmetrical scaling function $\varphi$ with Fourier transform $\hat{\varphi}(f) = L(f)$. The Fourier transform $G^{2n}(f)$ of the shifted basic sequence $2^{-1/2}\delta_{-1} * \psi_b^{2n}(x/2)$ generated by $\varphi$ (see Equation (55)) is given by

$$G^{2n}(f) = 2^{1/2}\left|2^{-1/2}U\left(f - \frac{1}{2}\right)L(f)\right|^{2n}\sum_i\left|L\left(f - \frac{1}{2} - i\right)\right|^{4n} \tag{90}$$

Using the fact that $2^{1/2}L(2f) = U(f)L(f)$, and the periodicity of $U(f)$ (i.e., $U(f) = U(f + 1)$), we rewrite (90) to obtain

$$\begin{aligned} 2^{-1/2}G^{2n}(f) = {} & \left|L(f)L\left(f - \frac{1}{2}\right)L(2f - 1)\right|^{2n} \\ & + \sum_{i\neq 0}\left|L(f)L\left(f - i - \frac{1}{2}\right)L(2f - 2i - 1)\right|^{2n} \end{aligned} \tag{91}$$

We consider the function $Z(f)$ given by

$$\begin{aligned} Z(f) &= 2^{-1/2}\left|L\left(f-\frac{1}{2}\right)\right|^2\left|U\left(f-\frac{1}{2}\right)L(f)\right| \\ &= \left|L(f)L\left(f-\frac{1}{2}\right)L(2f-1)\right| \end{aligned} \tag{92}$$

which appears in Equality (91). Since $u(k)$ is a real-valued sequence, and since $\varphi$ is a real-valued function, we have that $Z(f) = Z(-f)$. Thus, we will only consider the positive frequencies. For $f \in [0, \infty)$, Lemma 16 (Subsection 6.1 below) asserts the existence of a global maximum of $Z(f)$ which must occur in the frequency interval $[\frac{1}{4}, \frac{1}{2}]$. Let $f_0$ be the frequency at which this maximum occurs (if more than one, then $f_0$ will denote the smallest frequency) and define the number $\sigma_0$ by

$$\left.\frac{d^2Z}{df^2}\right|_{f=f_0} = -(2\pi\sigma_0)^2 Z(f_0) \tag{93}$$

We will assume that $\sigma_0 \neq 0$. We are now ready to state the convergence theorem:

**Theorem 15:** Let $\varphi$ be a symmetrical scaling function, and let $L(f)$ be its Fourier transform. If

$$\min_{f\in I} |L(f)| > |L(f)| \qquad \forall f \notin I = \left[-\frac{1}{2}, \frac{1}{2}\right] \tag{94}$$

then the Fourier transforms $G^{2n}(f)$ of the wavelets $2^{-1/2}\delta_{-1} * \psi_b^{2n}(x/2)$ have the convergence property given by

$$\lim_{n\to+\infty}\left\{\frac{1}{(Z(f_0))^{2n}}\, G^{2n}\left(\frac{f}{\sigma_n} \pm f_0\right)\right\} = 2^{1/2}\exp\left(-\frac{f^2}{2}\right) \tag{95}$$

where $\sigma_n = 2\pi\sigma_0(2n)^{1/2}$. If, in the interval $[0, \infty)$, $f_0$ is unique, then

$$\lim_{n\to+\infty}\left\|\frac{1}{(Z(f_0))^{2n}}\, G^{2n}\left(\frac{f}{\sigma_n}\right) - 2^{1/2}\exp\left(-\frac{(f - f_0\sigma_n)^2}{2}\right)\right\|_{L_p(0,+\infty)} = 0 \tag{96}$$

and for $q \in [2, \infty)$, we have in $L_q(\mathbb{R})$

$$\begin{aligned} \lim_{n\to+\infty}\Bigg\| &\frac{2^{-1/2}\sigma_n}{(Z(f_0))^{2n}}\,\delta_{-1} * \psi_b^{2n}\left(\frac{\sigma_n x}{2}\right) \\ &- 4\pi^{1/2}\cos(2\pi f_0\sigma_n x)\exp\left(-\frac{(2\pi x)^2}{2}\right)\Bigg\|_{L_q} = 0 \end{aligned} \tag{97}$$

Roughly speaking, the theorem states that the basic wavelet $\psi_b^{2n}(x/2)$ is essentially the real part of a Gabor function. It is centered at $x = 1$, its standard deviation is $\sigma_n = 2\pi\sigma(2n)^{1/2}$ and its modulation frequency is $\Omega = 2\pi f_0$. The functions $\psi_b^{2n}$ have axial symmetry, and their regularity increases with increasing values of $n$. Moreover, if the generating sequence $u(k)$ is finite, then the wavelets have compact support (see Section 4.1).

*Remark 6.1*: For a detailed discussion of Condition (94), we refer the reader to Remark 5.2.

### 6.1. Proof of Theorem 15

The proof of Theorem 15 relies on the following five lemmas:

**Lemma 16:** For $f \in [0, \infty)$, the function $Z(f)$ defined by (92) has a global maximum which can only occur at a frequency $f_0 \in [\frac{1}{4}, \frac{1}{2}]$.

*Proof*: The regularity conditions of Theorem 3 imply that $U(\frac{1}{2}) = 0$. This in turn implies that $L(i) = 0$, $\forall i \in \mathbb{Z}\backslash\{0\}$. Hence $Z(0) = 0$. Moreover, because $L(f)$ is continuous and $L(f) = O(|f|^{-r})$ with $r > 2$, $Z(f)$ is continuous and $Z(f) \to 0$ as $f \to \infty$. Thus, $Z(f)$ has a global maximum $Z(f_0)$ which occurs at some frequency $f_0$. Let $m$ be the minimum of $|L(f)|$ for $f \in [0, \frac{1}{2}]$; $m = \min_{f\in[0,1/2]}|L(f)|$. Condition (94) then implies that $Z(f) \geq m^3$, $\forall f \in [\frac{1}{4}, \frac{1}{2}]$, from which we get that $\max_{f\geq 0} Z(f) \geq m^3$. For $f > 1$, Condition (94) implies that $Z(f) < m^3$. Thus the global maximum of $Z(f)$ cannot occur for $f > 1$. For $f \in (\frac{1}{2}, 1]$, we use the symmetry $|L(f)| = |L(-f)|$, the fact that $f' = 1 - f$ is in $[0, \frac{1}{2}]$, and Condition (94) to get

$$\begin{aligned} Z(1-f) &= \left| L(1-f)L\left(f - \frac{1}{2}\right) L(2f-1) \right| \\ &> \left| L(f)L\left(f - \frac{1}{2}\right) L(2f-1) \right| \end{aligned} \tag{98}$$

Since the right hand side of Inequality (98) is precisely $Z(f)$, the maximum of $Z(f)$ cannot occur in $(\frac{1}{2}, 1]$. A similar argument in which we use $f' = \frac{1}{2} - f$ will show that the maximum must occur in the interval $[\frac{1}{4}, \frac{1}{2}]$.

**Lemma 17:** If a sequence $e_i$ is such that $|e_i| < 1$ $\forall i \in \mathbb{Z}$, and $|e_i| \leq \mathrm{Const}|i|^{-\alpha}$ for some $\alpha > 1$, then

$$\lim_{n\to+\infty} \sum_{i\in\mathbb{Z}} (e_i)^n = 0 \tag{99}$$

*Proof*: Since $|e_i| \leq \mathrm{Const}|i|^{-\alpha}$ with $\alpha > 1$, there exists an integer $i_0$ such that $|e_{i_0}| \geq |e_i|$, $\forall i \in \mathbb{Z}$. We multiply and divide the series by $|e_{i_0}|^n$ to get

$$S_n = \left|\sum_i e_i^n\right| \leq \sum_i |e_i|^n = |e_{i_0}|^n \sum_i |e_i/e_{i_0}|^n \tag{100}$$

The decay of $e_i$ for large values of $i$ and the fact that $|e_i/e_{i_0}| \leq 1$ imply that we can bound the series on the right hand of (100) by substituting $n = 2$. Thus we obtain

$$S_n \leq |e_{i_0}|^n \sum_i |e_i/e_{i_0}|^2 \leq \mathrm{Const}\, |e_{i_0}|^n \tag{101}$$

Since $|e_{i_0}| < 1$, $S_n$ tends to zero as $n$ tends to infinity.

**Lemma 18:** If a function $g(x)$ satisfies

$$|g(x)| \leq 1 - \varepsilon x^2 \qquad |x| < \varepsilon^{-1/2} \tag{102}$$

for some $\varepsilon > 0$, then

$$|g(n^{-1/2}x)|^n \leq \exp(-\varepsilon x^2) \qquad |x| < n^{1/2}\varepsilon^{-1/2} \tag{103}$$

*Proof*: Using the inequality $\ln(1 + u) \leq u$ for $|u| < 1$, we get

$$n \ln |g(n^{-1/2}x)| \leq n \ln (1 - \varepsilon n^{-1}x^2) \leq -\varepsilon x^2 \qquad |x| < n^{1/2}\varepsilon^{-1/2} \tag{104}$$

from which the proof immediately follows.

**Lemma 19:** If a function $g(x)$ satisfies

$$|g(x)| \leq |cx|^{-2} \tag{105}$$

then there exists a constant Const independent of $n$ such that $\forall n > 2$

$$|g(n^{-1/2}x)|^n \leq \text{Const}|x|^{-2} \qquad |cx|n^{-1/2} \geq E > 1 \tag{106}$$

*Proof:* Let $t(x)$ be the right side of Inequality (105). The function $|t(xn^{-1/2})|^n$ can be written as

$$|t(xn^{-1/2})|^n = c^{-2}n(cxn^{-1/2})^{2-2n}|x|^{-2} \tag{107}$$

For $c|x|n^{-1/2} \geq E > 1$ and for $n > 2$, we can estimate the right side of (107) to obtain

$$|t(xn^{-1/2})|^n \leq c^{-2}E^2nE^{-2n}|x|^{-2} \tag{108}$$

Since the term $c^{-2}E^2nE^{-2n}$ in (108) is bounded above by a constant Const independent of $n > 0$, the proof of the lemma follows.

The next lemma is well-known. It appears as part of the proof of the Central Limit Theorem [25]. In this latter context, $A(x)$ is the characteristic function of a random variable and has a maximum at $x = 0$.

**Lemma 20:** If a function $A(x)$ is such that

i) $A(x_0) = 1$

ii) $\left.\dfrac{dA(x)}{dx}\right|_{x=x_0} = 0$

iii) $-\infty < \left.\dfrac{d^2A(x)}{dx^2}\right|_{x=x_0} = -\alpha^2 < 0$

then

$$\lim_{n\to+\infty}\left(A\left(\frac{x}{\alpha\sqrt{n}} + x_0\right)\right)^n = \exp\left(-\frac{x^2}{2}\right) \tag{109}$$

*Proof:* Without loss of generality, we assume that $x_0 = 0$, and that $\alpha = 1$. For a fixed value of $x$, we consider the function

$$L_n(x) = n\ln\left(A\left(\frac{x}{\sqrt{n}}\right)\right) \tag{110}$$

To evaluate the limit, we apply l'Hospital's rule and differentiate twice to get

$$\begin{aligned}\lim_{n\to+\infty} L_n(x) &= \lim_{n\to+\infty}\left(\frac{1}{2}\frac{n^{-3/2}xA'(n^{-1/2}x)}{A(n^{-1/2}x)n^{-2}}\right)\\ &= \lim_{n\to+\infty}\left(\frac{x^2}{2}A''\left(n^{-1/2}x\right)\right) = -\frac{x^2}{2}\end{aligned} \tag{111}$$

which immediately yields (109).

*Proof of Theorem 15:* (i) *Pointwise convergence*. The function $2^{-1/2}G^{2n}(f)$

is given by (91). We divide (91) by $(Z(f_0))^{2n}$, where $Z(f)$ is defined by (92) and where $Z(f_0)$ is a global maximum as in Lemma 16. We get

$$2^{-1/2}\frac{G^{2n}(f)}{(Z(f_0))^{2n}} = \left|(Z(f_0))^{-1}L(f)L\left(f-\frac{1}{2}\right)L(2f-1)\right|^{2n} + \sum_{i\neq 0}\left|(Z(f_0))^{-1}L(f)L(f-i-\frac{1}{2}) L(2f-2i-1)\right|^{2n} \tag{112}$$

To simplify the notation, we define the following quantities:

$$A(f) = \left|(Z(f_0))^{-1}L(f+f_0)L\left(f+f_0-\frac{1}{2}\right)L(2f+2f_0-1)\right| \tag{113}$$

$$c_i(f) = \left|(Z(f_0))^{-1}L(f)L\left(f-i-\frac{1}{2}\right)L(2f-2i-1)\right| \tag{114}$$

$$R_{2n}(f) = \sum_{i\neq 0}|c_i(f)|^{2n} \tag{115}$$

With these definitions, the function $2^{-1/2}G^{2n}(\sigma_n^{-1}f+f_0)(Z(f_0))^{-2n}$ can be written as

$$2^{-1/2}\frac{G^{2n}(\sigma_n^{-1}f+f_0)}{(Z(f_0))^{2n}} = |A(\sigma_n^{-1}f)|^{2n} + R_{2n}(\sigma_n^{-1}f+f_0) \tag{116}$$

where $\sigma_n = 2\pi\sigma_0(2n)^{1/2}$.
By Lemma 20, it follows that

$$|A(\sigma_n^{-1}f)^{2n} = (Z(\sigma_n^{-1}f+f_0)/Z(f_0))^{2n} \tag{117}$$

converges to $\exp(-f^2/2)$ as $n$ tends to infinity.
The decay condition $L(f)$ for large values of $f$ implies that there exists a constant Const independent of $n$ such that the terms $c_i(\sigma_n^{-1}f+f_0)$ of the remaining series $R_{2n}(\sigma_n^{-1}f+f_0)$ are bounded by

$$|c_i(\sigma_n^{-1}f+f_0)| \le \text{Const}|i|^{-4} \qquad \forall n \ge N_1(f) \tag{118}$$

Moreover, using Condition (94), an argument similar to the one in Lemma 16 allows us to conclude that $|c_i(f_0)| < 1$ for $i \neq 0$. It follows that for $f \geq 0$, we have

$$\lim_{n\to\infty}|c_i(\sigma_n^{-1}f+f_0)| < 1 \tag{119}$$

The estimate (118) and the limit (119) allow us to find a sequence $e_i$ which does not depend on $n$, satisfying $e_i < 1\ \forall i \neq 0$, $e_i = O(|i|^{-4})$, and such that

$$|c_i(\sigma_n^{-1}f+f_0)| \le e_i < 1 \qquad \forall n \ge N(f) \tag{120}$$

By Lemma 17, this last inequality implies that the remainder series $R_{2n}(\sigma_n^{-1}f + f_0)$ in (116) converges to zero as $n$ goes to infinity. This finishes the proof of the pointwise convergence (95).

(ii) *Convergence in $L_p$*. For the remaining part of the theorem, we have that for $f \geq 0$ the global maximum occurs at a unique location $f_0$. From this, and because $\sigma_0 \neq 0$ and $L(f) = O(|f|^{-2})$, we have that the function $A_+(f) = \chi_{[-f_0,\infty)}A(f)$, where $\chi_I$ denotes the indicator function on $I$, satisfies

$$|A_+(f)| \leq 1 - \varepsilon|f|^2 \qquad |f| \leq \varepsilon^{-1/2} \tag{121}$$

for all sufficiently small $\varepsilon$; $0 < \varepsilon < \varepsilon_0$. We also have that

$$|A_+(f)| \leq |\gamma f|^{-2} \qquad |f| \geq T \tag{122}$$

By choosing $\varepsilon$ sufficiently small, and by using the last two inequalities, Lemma 18 and Lemma 19, we get the estimate

$$|A_+(\sigma_n^{-1}f)|^{2n} \leq \text{Const}(\exp(-\varepsilon_1 f^2) + \chi_{\mathbb{R}\setminus[-1,1]}|f|^{-2}) \qquad \forall f \in \mathbb{R} \tag{123}$$

where the constants Const and $\varepsilon_1 > 0$ are independent of $n$ and where $\chi_{\mathbb{R}\setminus[-1,1]}$ is an indicator function. The right hand side of Inequality (123) is a function independent of $n$, which lies in the spaces $L_p(\mathbb{R})$ $\forall p \in [1, \infty)$. Using the pointwise convergence of $|A(\sigma_n^{-1}f)|^{2n}$ to $\exp(-f^2/2)$ established earlier and using the Lebesgue Dominated Convergence Theorem, we then get

$$\lim_{n\to\infty} \left\| |A_+(\sigma_n^{-1}f)|^{2n} - \exp(-f^2/2) \right\|_{L_p(-\infty,\infty)} = 0 \tag{124}$$

It follows that, in $L_p[0, \infty)$ we have the limit

$$\lim_{n\to\infty} \left\| |(Z(f_0))^{-1}Z(\sigma_n^{-1}f)|^{2n} - \exp(-(f - \sigma_n f_0)^2/2) \right\|_{L_p[0,\infty)} = 0 \tag{125}$$

To finish the proof, we need to estimate the quantity

$$\int_{-\sigma_n f_0}^{\infty} |R_{2n}(\sigma_n^{-1}f + f_0)|^p \, df \tag{126}$$

where $R_{2n}(f)$ is defined by (115). We first note that

$$\int_{-\sigma_n f_0}^{\infty} |R_{2n}(\sigma_n^{-1}f + f_0)|^p \, df = \int_0^{\infty} |R_{2n}(\sigma_n^{-1}f)|^p \, df = \sigma_n \int_0^{\infty} |R_{2n}(f)|^p \, df \tag{127}$$

We bound $\int_0^\infty |R_{2n}(f)|^p \, df$ from above by the sum of the three terms below:

$$\begin{aligned} \int_0^{\infty} |R_{2n}(f)|^p \, df \leq{}& C \int_0^T \left| \sum_{0<|i|\leq i_0} |c_i(f)|^{2n} \right|^p df \\ &+ C \int_0^T \left| \sum_{|i|>i_0} |c_i(f)|^{2n} \right|^p df + C \int_T^{\infty} |R_{2n}(f)|^p \, df \end{aligned} \tag{128}$$

where $C$ is a constant independent of $n$. We rewrite $R_{2n}(f)$ as

$$|R_{2n}(f)| = |(Z(f_0))^{-1}L(f)|^{2n} \sum_{i\neq 0} \left| L\left(f - i - \frac{1}{2}\right) L(2f - 2i - 1) \right|^{2n} \tag{129}$$

Since $|L(f)| = O(|f|^{-2})$, the series on the right-hand side is bounded by $C^{2n}$ where $C$ is a constant independent of $n$. Combining this fact and the fact that $|L(f)| = O(|f|^{-2})$ we conclude that

$$|R_{2n}(f)| \leq \left|\text{Const}|f|\right|^{-4n} \quad \forall |f| > m_3 \tag{130}$$

Thus by choosing $T > m_3$ sufficiently large, we use the last inequality to estimate the third terms in the right-hand side of (128) by

$$\int_T^\infty |R_{2n}(f)|^p \, df \leq \text{Const}_3 |a_3|^{4np} \qquad |a_3| < 1 \tag{131}$$

where the constant $\text{Const}_3$ is independent of $n$.

Since $|L(f)| = O(|f|^{-2})$, we can estimate $c_i(f)$ in (115) for $f \in [0, T]$ by

$$|c_i(f)| \leq \left|\text{Const}|i|\right|^{-4} \qquad \forall |i| \geq m_2 \tag{132}$$

From this inequality, we get an upper bound for the second term on the right hand side of (128)

$$\int_0^T \left| \sum_{|i|>i_0} |c_i(f)|^{2n} \right|^p df \leq 2T \left| \text{Const}|i_0| \right|^{-8np} \left| \sum_{|i|\geq 1} \left| \frac{i_0}{i_0 + i} \right|^{8n} \right|^p \tag{133}$$

The series on the right-hand side of (133) is largest when $n = 1$. It follows that by choosing $i_0 > m_2$ sufficiently large, we obtain the inequality

$$\int_0^T \left| \sum_{|i|>i_0} |c_i(f)|^{2n} \right|^p df \leq \text{Const}_2 |a_2|^{8np} \qquad |a_2| < 1 \tag{134}$$

To estimate the first term on the right-hand side of (128) we first note that, by a simple argument similar to the one in Lemma 16, Condition (94) implies that for $f \geq 0$, and $i \neq 0$

$$|c_i(f)| \leq |a_1| < 1 \tag{135}$$

From the above inequality, it follows that

$$\int_0^T \left| \sum_{0<|i|\leq i_0} |c_i(f)|^{2n} \right|^p df \leq T(2i_0)^p |a_1|^{2np} \leq \text{Const}_1 |a_1|^{2np} \tag{136}$$

Combining Equations (127), (128), (131), (134), and (136) we obtain

$$\int_0^\infty |R_{2n}(\sigma_n^{-1} f)|^p \, df \leq \text{Const}\, \sigma_n (|a_1|^{2np} + |a_2|^{8np} + |a_3|^{4np}) \tag{137}$$

where $\sigma_n = 2\pi\sigma_0(2n)^{1/2}$ and where the constant Const is independent of $n$. Since $|a_i| < 1$ for $i = 1, 2, 3$, it follows that the left hand side of (137) tends to zero as $n \to \infty$. Equations (116), (125), and (137) yield Equation (96) of the theorem. Finally, Equation (97) in the theorem is obtained using (95–96), the fact that the Fourier transform is a bounded operator from $L_p$ into the $L_q$ with $p \in [1, 2]$, $q^{-1} + p^{-1} = 1$, and from

$$\lim_{n\to\infty} \|\exp(-(f + \sigma_n f_0)^2/2)\|_{L_p[0,\infty)} = 0 \qquad \forall p \in [1, \infty) \tag{138}$$

$$\lim_{n\to\infty} \|\exp(-(f - \sigma_n f_0)^2/2)\|_{L_p(-\infty,0]} = 0 \qquad \forall p \in [1, \infty) \tag{139}$$

## 7. EXAMPLES: POLYNOMIAL SPLINES

A particular case of the present theory is provided by the example of polynomial splines. The centered B-splines of order $m$, $\beta^m$, are obtained by repeated convolution of the B-spline of order 0 [46]:

$$\beta^m(x) = (\beta^0 * \beta^0 * \cdots * \beta^0)(x) \quad (m \text{ convolutions}) \tag{140}$$

where $\beta^0(x)$ is the characteristic function in the interval $[-\frac{1}{2}, \frac{1}{2})$. Schoenberg has shown that these functions generate the polynomial splines of order $m$ with knot points at the integer ($m$ odd). In other words, any polynomial spline function $s(x)$ of order $m$ can be represented as

$$s(x) = \sum_{k=-\infty}^{k=+\infty} c(k)\beta^m(x - k) \tag{141}$$

By choosing $\varphi = \beta^0 * \beta^0$, the scaling functions $\varphi^n = \beta^{2n-1}$ generate a well defined sequence of multiresolutions of $L_2$. For instance, it can be verified that the generating sequence $u$ for $\varphi$ has the Fourier transform

$$U(f) = 2^{-1/2}(1 + \cos(2\pi f)) \tag{142}$$

Using this last expression, it is easy to verify that all the conditions of Theorem 3 are satisfied.

Since the function $\varphi$ also satisfies the conditions of our convergence theorems, we can conclude that the interpolating and orthogonal spline filters $\hat{\varphi}^n_I$, $\mathring{\hat{\varphi}}^n_I$, and $\hat{\varphi}^n_0$, converge to the ideal lowpass filter as $n$ tends to infinity. More detailed convergence results can be found in [5, 52]. Related asymptotic properties of spline interpolants are also discussed in [38, 47].

We can generate the corresponding basic, orthogonal, and interpolating spline wavelets using the formulas in Section 4. This construction was performed explicitly in [50]. The basic spline wavelets are the so called B-spline wavelets of compact support [17, 51]. The corresponding orthogonal wavelets are precisely the Battle/Lemarié spline wavelets [7, 33]; see also [37]. From our convergence results, we conclude that the Fourier transforms $\hat{\psi}^n_0$, and $\hat{\psi}^n_I$ converge to the ideal bandpass filter as $n$ tends to infinity. The pointwise convergence in the special case of the orthogonal spline wavelet has been established by Lemarié [33]. Figure 2 shows the basic piecewise linear and cubic B-spline wavelets, and the Gabor limit, which is a Gaussian function modulated by the cosine. The graph shows that cubic B-spline wavelet is already a good approximation to the Gabor function. For this case, the time-frequency measure $\chi = \sigma_{|\psi|^n_b|^2}\sigma_{|\hat{\psi}|^n_b|^2_+}$ defined at the beginning of this section is found to be within 2% of the optimal number $(2\pi)^{-2}$. More details can be found in [51].

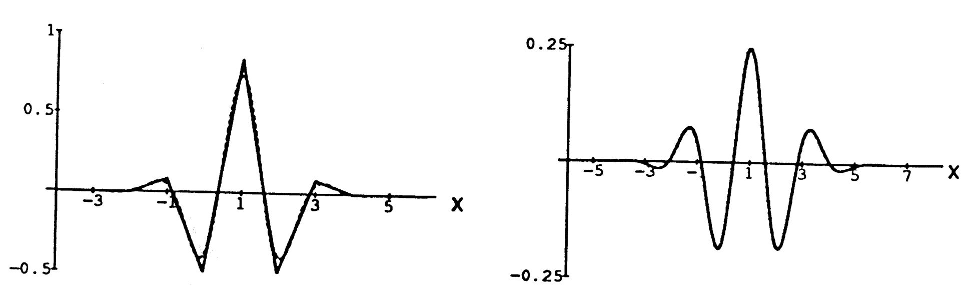


**Figure 2**

## ACKNOWLEDGMENTS

We would like to thank Barry Bowman and Jane Bradley for their editorial help. We also would like to thank the reviewers for their insights and helpful comments.

## REFERENCES

1. R. Adams, *Sobolev spaces*, Academic Press, New York, 1975.
2. A. Aldroubi and M. Unser, *Sampling procedures in function spaces and asymptotic equivalence with Shannon's sampling theory, Numer. Funct. Anal. Optimiz.*, to appear.
3. A. Aldroubi and M. Unser, *Families of wavelet transforms in connection with Shannon's sampling theory and the Gabor transform*, in Wavelets—A tutorial in theory and applications, C. K. Chui, ed., Academic Press, San Diego 1992, 509–528.
4. A. Aldroubi, M. Unser, and M. Eden, *Asymptotic properties of least square spline filters and applications to multi-scale decomposition of signals*, in Proc. international symposium on information theory and its applications, Waikiki, Hawaii, 1, 1990, 271–274.
5. A. Aldroubi, M. Unser, and M. Eden, Cardinal spline filters: stability and convergence to the ideal sinc interpolator, *Signal Processing 28* (1992), 1–12.
6. M. J. Bastiaans, On the sliding-window representation in digital signal processing, *IEEE Trans. on Acoust. Speech Sign. Proc. 33* (1985), 868–873.
7. G. Battle, A block spin construction of ondelettes. Part I: Lemarié functions, *Commun. Math. Phys. 110* (1987), 601–615.
8. G. Battle, Heisenberg proof of the Balian-Low theorem, *Lett. Math. Phys. 15* (1988), 175–177.
9. J. Benedetto, *Gabor representation and wavelets*, in Commutative harmonic analysis, D. Collela, ed., Providence American Mathematical Society, 1989, 9–27.
10. J. Benedetto, C. Heil, and D. Walnut, *Remarks on the proof of the Balian-Low Theorem*, preprint.
11. J. J. Benedetto, *Uncertainty principle inequalities and spectrum estimation, in Fourier analysis and its applications*, NATO Advanced Study Institute, IL Ciocco, Italy, 1989.
12. J. J. Benedetto, *Irregular sampling and frames*, in Wavelets—A tutorial in theory and applications, C. K. Chui, ed., Academic Press, Orlando, 1992, 445–507.
13. W. M. Brown, Optimal prefiltering of sampled data, *IRE Trans. Inform. Theory 7* (1961), 269–270.
14. P. L. Butzer, A survey of the Whittaker-Shannon sampling theorem and some of its extensions, *J. Math. Res. Exposition 3* (1983), 185–212.
15. C. K. Chui, *An overview of wavelets, in approximation theory and functional analysis, Academic Press*, New York, 47–71, 1991.
16. C. K. Chui and J. Z. Wang, *A general framework of compactly supported splines and wavelets*, CAT Report 219, Texas A&M University, 1990.
17. C. K. Chui and J. Z. Wang, On compactly supported spline wavelets and a duality principle, *Trans. Amer. Math. Soc.*, 1991.
18. A. Cohen, I. Daubechies, and J. C. Feauveau, *Biorthogonal bases of compactly supported wavelets*, AT&T Bell Laboratories, Technical Report #11217-900529-07.
19. R. R. Coifman, Y. Meyer, S. Quake, and M. V. Wickerhauser, *Signal processing and compression with wave packets*, in Proc. of the conference on wavelets, Marseilles, 1989.
20. I. Daubechies, Orthonormal bases of compactly supported wavelets, *Comm. Pure Appl. Math. 41* (1988), 909–996.

21. I. Daubechies, The wavelet transform, time-frequency localization and signal analysis, *IEEE Trans. Inform Theory 22* (1990), 592–593.
22. I. Daubechies, S. Jaffard, and J. L. Journé, A simple Wilson orthogonal basis with exponential decay, *SIAM J. Math. Anal. 22* (1991), 554–572.
23. C. de Boor, K. Höllig, and S. Riemenschneider, Convergence of cardinal series, *Proc. Amer. Math. Soc. 98* (1986), 457–460.
24. R. A. DeVore, B. Jawerth, and B. J. Lucier, Image compression through wavelet transform coding, *IEEE Trans. on Inform. Theory 38* (1992), 719–746.
25. W. Feller, *An introduction to probability theory and its application. Vol. II.* Wiley, New York, 1971.
26. D. Gabor, Theory of communication, *J. IEE 93* (1945), 429–441.
27. C. E. Heil and D. F. Walnut, Continuous and discrete wavelet transforms, *SIAM Review 31* (1989), 628–666.
28. J. R. Higgins, Five short stories about the cardinal series, *Bull. Amer. Math. Soc. 121* (1985), 45–89.
29. R. Hummel, Sampling for spline reconstruction, *SIAM J. Appl. Math. 43*, 1983, 278–288.
30. A. Jerri, Sampling expansion for Laguerre transforms, *J. Res. Nat. Bur. Standards Sec. B (80), (B) 3* 1976), 415–418.
31. A. J. Jerri, The Shannon sampling theorem-its various extensions and applications: a tutorial review, Proc. *IEEE 65* (1977), 1565–1596.
32. H. P. Kramer, A generalized sampling theorem, *J. Math. Phys. 38* (1959), 68–72.
33. P. G. Lemarié, Ondelettes à localisation exponentielles, *J. Math. Pures et Appl. 67* (1988), 227–236.
34. P. G. Lemarié and Y. Meyer, Ondelettes et base hilbertiennes, *Rev. Math. Iberoamericana 2*, 1986.
35. S. Mallat and W. L. Huang, Singularity detection and processing with wavelets, *IEEE Trans. Inform. Theory 38* (1992), 617–643.
36. S. G. Mallat, Multiresolution approximations and wavelet orthogonal bases of $L_2(\mathbb{R})$, *Trans. Amer. Math. Soc. 315* (1989), 69–87.
37. S. G. Mallat, A theory of multiresolution signal decomposition: the wavelet representation, *IEEE Trans. Pattern Anal. Machine Intell. 11* (1989), 674–693.
38. M. J. Marsden, F. B. Richards, and S. D. Riemenschneider, Cardinal spline interpolation operators on $l_p$ data, *Indiana Univ. Math. J. 24* (1975), 677–689.
39. Y. Meyer, *Ondelettes, functions splines, et analyses graduees*, Univ. of Toronto, 1986.
40. Y. Meyer, Ondelettes, Hermann Paris, Editeurs des Sciences et des Arts, 1990.
41. M. Z. Nashed and G. G. Walter, General sampling theorems for functions in reproducing kernel Hilbert spaces, *Math. of Control Signals Sys. 4* (1991), 373–412.
42. H. Ogawa, A generalized sampling theorem, *Electronics and Communications in Japan 72* (1989), 97–105.
43. S. Riemenschneider, Multivariate Cardinal Interpolation, in Approximation Theory VI, C. K. Chui, L. L. Schumaker and J. D. Ward, Ed., Academic Press, New York, 1989, 561–580.
44. S. D. Riemenschneider and Z. Shen, Box splines, cardinal series, and wavelets, in Approximation theory and functional analysis, C. K. Chui, ed., Academic Press, New York, 1991, 133–149.
45. O. Rioul and M. Vetterli, Wavelets and signal processing, *IEEE Signal Processing Magazine 8* (1991), 14–35.
46. I. J. Schoenberg, Contribution to the problem of approximation of equidistant data by analytic functions. *Quart. Appl. Math. 4* (1946), 45–99, 112–141.
47. I. J. Schoenberg, Notes on spline functions III: on the convergence of the interpolating cardinal splines as their degree tends to infinity, *Israel J. Math 16* (1973), 87–92.

48. C. E. Shannon, Communications in the presence of noise, *Proc. IRE 37* (1949), 10–21.

49. M. Unser and A. Aldroubi, Polynomial splines and wavelets—A signal processing perspective, in Wavelets—A tutorial in theory and applications, C. K. Chui, ed., Academic Press, Orlando, FL, 91–122, 1992.

50. M. Unser, A. Aldroubi, and M. Eden, A family of polynomial spline wavelet transforms, *Signal Processing*, *30* (1993), 141–162.

51. M. Unser, A. Aldroubi, and M. Eden, On the asymptotic convergence of B-spline wavelets to Gabor functions, *IEEE Trans. Inform. Theory* (special issue on wavelets) *38* (1992), 864–872.

52. M. Unser, A. Aldroubi, and M. Eden, *A sampling theory for polynomial splines*, in Proc. international symposium on information theory and its applications, Waikiki, Hawaii (1990), 279–282.

53. M. Vetterli and C. Herley, Wavelets and filter banks, *IEEE Trans. on Signal Proc.*, *40* (1992), 2207–2231.

54. G. G. Walter, Sampling theorems for wavelet subspaces, *IEEE Trans. Inform. Theory* (special issue on Wavelets), *38* (1992), 881–884.

55. J. M. Whittaker, Interpolation function theory, in Cambridge tracts in mathematics and mathematical physics, Cambridge University Press, Cambridge, 1935.

56. K. Yao, Application of Reproducing Kernel Hilbert Spaces-bandlimited signal models, *Information and Control 11* (1967), 429–444.

57. L. P. Yaroslavsky, *Digital picture processing—An introduction*, Springer-Verlag, Berlin, 1985.

58. A. Zayed, G. Hinsen and P. Butzer, On Lagrange interpolation and Kramer-type sampling theorems associated with Sturm-Liouville problems, *SIAM J. Appl. Math.* *50* (1990), 893–909.

*Received*: February 1992
*Revised*: September 1992
*Accepted*: October 1992